Device-based measured Physical Activity - Technical Decisions, Preprocessing and Evaluation:

Using Accelerometers in the large scale epidemiological MoMo Study

Zur Erlangung des akademischen Grades eines
DOKTORS DER PHILOSOPHIE (Dr. phil.)

von der KIT-Fakultät für Geistes- und Sozialwissenschaften
des Karlsruher Instituts für Technologie (KIT)
angenommene

DISSERTATION

von

Alexander Burchartz

KIT-Dekan:	Prof. Dr. Michael Mäs
1. Gutachter:	Prof. Dr. Alexander Woll
2. Gutachter:	Prof. Dr. Claudio Nigg
3. Gutachterin:	Dr. Claudia Niessner

Tag der mündlichen Prüfung: 14. Juli 2023

Alexander Burchartz

Motorik-Modul Study

Social and Health Sciences in Sport

Institute of Sports and Sports Science

Karlsruhe Institute of Technology

Engler-Bunte-Ring 15, Building 40.40

76131 Karlsruhe, Germany

alexander.burchartz@kit.edu
orcid.org/0000-0003-1338-5395

Acknowledgments

I would like to thank my doctoral mentor Prof. Alexander Woll as well as my two other mentors Claudia Niessner and Doris Oriwol for their excellent supervision and support during my work. I am very grateful for all the things I could learn during your supervision and our work together. You have not only taught me a lot in the academic fields, but also helped me to grow as a person.

Furthermore, I would like to thank my colleagues at the Institute of Sports and Sports Science at the Karlsruhe Institute of Technology for the productive discussions and your valuable support in many matters. Special recognition goes to Simon Kolb for the close and constructive teamwork in the context of the MoMo-Study, who had my back in project management the whole time, but especially in the final stages and Bastian Anedda for the collaborative very deep familiarization with the sometimes rather extensive technical documentation in the selection and application of methods for quantifying physical behavior at the beginning of this journey. A huge thank you also goes to the whole MoMo team, without you this work would not have been possible. I will always remember the intense, beautiful, eventful, stressful, and shaping time and will never forget how much we also laughed together.

I would also like to mention the PhD Colloquium for the constructive and educational meetings where one could exchange ideas and experiences with the multitude of other PhD students and postdocs.

A heartfelt thank you is also due to all co-authors for their contributions, helpful discussions, and excellent feedback on our publications. A big thank you to all the research assistants, students, and countless field study assistants who helped with the data collection and processing. Without this great teamwork, this work would not have been possible.

Finally, this endeavor would not have been accomplished without my family and friends. On an emotional level, my friends have made a crucial contribution to the success of my dissertation. They have supported me in many ways and have given me a lot of strength and confirmation on my way, especially through their listening, their encouragement, and the many experiences and activities I have shared with them. I would especially like to thank Christoph Wallisch, Denis Knöpfle, Philipp Weimann, Yasha Farmani, Lisa Mai, Lisa Kohl and Oliver Herz. It is a great gift to have spent my time at university with you.

The same goes for my volleyball friends. Without the sportive counterbalance, I would not have been able to finish this work, even if I probably would have finished faster without volleyball. Nevertheless, I don't want to miss a single minute that I spent with you on and off the court. You are one of the most important motivations for me to deal with the topic of movement and physical activity and I thank Christian Schneider, David Wulf, Tom Haulitschke, Florian Salah, Konstantin Sieler, and all my other teammates from the bottom of my heart for the wonderful experiences.

The unconditional and loving support of my parents made this dissertation possible in the first place, for which I am infinitely grateful. They have always encouraged me to go my own way and have given me exemplary support in achieving my goals, even if I have sometimes not been quite at the pace I had hoped.

My deepest thanks go to my beloved wife Nicole and our wonderful daughter Jenna. Thank you for your unconditional love for your understanding, your endless patience, your trust, and your enormous support in all ways.

From an early age, I wanted to go to school as long as possible and I definitely managed that. Now I am curious where the next journey takes me, but I am sure that I am on the right path and will learn more interesting stuff in the future.

Summary

According to the Global Health Risks Report of the World Health Organization (WHO), about one-third of all deaths worldwide are attributable to a few risk factors. Inadequate physical activity ranks fourth among these risk factors. A person's physical activity is defined as any movement performed by skeletal muscles that consumes energy. In short, many people simply do not move enough in general. Worldwide trends in insufficient physical activity are regularly tracked by WHO using large pooled data analyses. One finding from its most recent analyses is that the global age-standardized prevalence of inadequate physical activity was nearly one-third at its latest count. The difference between genders was even greater, and the situation is additionally more critical for young people. Overall, only one-fifth of the children and adolescents worldwide were sufficiently physically active. Thus, a large proportion of adolescents are not adhering to the current recommendations of at least 60 minutes of moderate physical activity daily, which may impact their current and future health status.

But how much physical activity and exercise is enough and how is this being monitored? One of the first recommendations on this question was published by the American College of Sports Medicine in 1975. This publication had a major impact on sports science and has been revised several times. In 2010, the World Health Organization published the first international version of the physical activity guidelines. Building on this, many countries have also adopted these guidelines as national recommendations, including Germany. The WHO recommended amount of weekly or daily physical activity is used as a reference to determine which individuals are reaching a sufficient level of physical activity. As a result of these globally standardized recommendations, many studies have used questionnaires to determine the impact on health from adhering to these guidelines.

Population-based studies investigating adherence to these recommendations in Germany include the DEGS, KiGGS, and MoMo studies. The „German Health Interview and Examination Survey for Children and Adolescents" (KiGGS) assesses physical activity, health parameters, and social and environmental determinants of children and adolescents in Germany. The **Motorik Modul (MoMo)** longitudinal study has been a representative module of the KiGGS study since 2003 and at the same time an in-depth study to record physical activity and motor performance, which was conducted by the Karlsruhe Institute of Technology in three survey waves (wave 1 2009-2012, wave 2

2015-2017, wave 3 2018-2022) after the baseline study (2003-2006). This dissertation was also mainly conducted as part of the MoMo study, focusing on the measurement of physical activity with accelerometers.

Until 2015, only the **MoMo Physical Activity Questionnaire** was used to record physical activity in the MoMo study. The questionnaire focuses on recording the central aspects of type, intensity, frequency and duration of typical or habitual activity. In 2015, the recording of physical activity by activity monitors, so-called accelerometers, was added to the MoMo study. This had the purpose of expanding the data collection to include a continuous recording of the current activity intensity, which is difficult to record with questionnaires.

In fact, it is also evident from the sales figures that **physical activity monitors** are now a leading fitness trend. In 2017-18, more than 100 million tracking devices and accelerometer-based smartwatches were sold, and by 2020, that number was nearly 445 million units. Physical activity monitors, also technically known as accelerometers, have emerged as important monitoring tools in clinical research and health promotions, and are now being used more frequently in national research studies. ActiGraph accelerometers were among the first accelerometers to be validated and used early in the study of physical activity and health. In 2014, more than 51 % of 76 studies, each with more than 400 participants in 36 countries, used an ActiGraph accelerometer. Therefore, KiGGS and MoMo also chose to use ActiGraph accelerometers in Wave 2 to ensure comparability with other large epidemiological studies.

Today, accelerometers offer the ability to collect real-time, valid data on the intensity and extent of physical behavior over multiple days and weeks. To describe the unique aspects of physical behavior, a large number of metrics can be derived from these **multidimensional data**. When choosing and applying instruments to quantify physical behavior, researchers must manage a great number of technical details in accelerometry. The impact of these technical choices on the metrics (energy expenditure, activity intensity, body position, and activity patterns) can occur in a number of ways. The device, the wearing position (hip, wrist, thigh), and the recording parameters (epoch length, frequency, storage capacity, recording frequency, and signal filters) have a major impact on the measured activity. Different backgrounds such as study design (purpose, repeated measurements) and duration (time frame, wearing time) as well as data storage and evaluation have to be considered when deciding on the parameters. In addition, several adjusting factors (raw data, contextual information, non-wearing time, intensity classification, compliance) have to be customized in the evaluation depending on the target variables. In MoMo, for example, this results in approximately 18 million data points for each subject when wearing an accelerometer for one week. The data from all subjects must be preprocessed, transformed, and then analyzed in the most time-efficient manner possible.

The **main objectives of this dissertation** were therefore to 1) disscuss how technical decisions should be made when using accelerometers in large epidemiological studies of activity research, 2) which basic methodological aspects were therefore applied in the device-based measurement of physical activity specifically in the MoMo study, 3) what are the implications of different evaluation methods in the preprocessing of accelerometer data, 4) how accelerometer survey data differ from the representative MoMo Physical Activity Questionnaire data, and 5) what are the results of accelerometer data with respect to differences between weekdays and weekend days. These questions were investigated in five scientific articles.

In the **“Consensus Article”** we reflect the expert consensus of the authors during the 2nd International Workshop for the Center for the Assessment of Physical Activity (CAPA) on the topic of accelerometry for the assessment of movement behavior. The article clarifies the aspects that should be considered in the design and evaluation of accelerometer studies. One of the major findings of the workshop is the need to thoroughly familiarize oneself with existing validation studies and to document as many technical decisions as possible in data collection and analysis to allow for later comparison of data across studies. We suggest that based on the most valid approach, the particular behavioral metric and corresponding appropriate method should be chosen. From this perspective, the method chosen will determine the type of device and the prediction algorithm used to determine physical activity. Due to the large number of possible "subjective" decisions prior to the actual measurement, the article provides an overview of the technical details of accelerometry that need to be considered. At the same time, it outlines best practices in the selection and use of devices to quantify the three major behavioral categories of general interest to the research community: physical activity, sedentary behavior, and sleep.

To date, there have been no nationwide device-based physical activity data for children and adolescents in Germany. With KiGGS and MoMo wave 2, device-based physical activity data on children and adolescents were collected nationwide for the first time. Therefore, building on the recommendations of the CAPA workshop, in our **“Study Protocol”** we elaborated on the basic methodological aspects that were used for the device-based measurement of physical activity specifically in the MoMo study. The study protocol provides an overview of the technical details and basic decisions when using accelerometers in large-scale epidemiological studies. Limitations imposed by the specified filters and evaluation routines must be given special consideration.

To clarify which consequences different evaluation methods have in the subsequent processing of accelerometer data, the influence of three specific factors was investigated in the **“Methods Article”**. We examined the extent to which the factors "non-wearing time algorithm," "epoch length," and "intensity thresholds" influence the quantification of physical activity measured with accelerometers. The differences found were significant for all three factors. As a result, depending on the selection of the factor, the resulting discrepancies in the estimated values for sedentary behavior and physical activity can become very large. In particular, the magnitude of the divergence between epoch lengths has shown that particular consideration must be given to this when interpreting accelerometer data. Thus, epoch length should be considered decisively as an influential variable when comparing values of physical activity between studies. We also suggest that data be pooled and analyzed in a unified manner. In addition to new validation studies with short epoch lengths for younger children, we also recommend conducting more meta-analyses. This could use data from multiple studies to validate cut-off points and propose a consensual set of thresholds that can be used in different settings and studies in children and adolescents.

Because children in particular demonstrate more complex but less structured movement behavior than adults, determining their many spontaneous and impulsive movements is a particular challenge for assessing physical activity. Neither questionnaires nor accelerometers provide optimal coverage of all facets of physical activity. For this reason, a combined multimodal approach of self-report and device-based methods is recommended. In the **"Comparison Article"**, we therefore compared physical activity measured with the accelerometer and physical activity recorded with the MoMo questionnaire. Based on the number of days participants achieved the WHO recommendation for physical activity, this study examines the extent to which self-reported and device-based measured physical activity differ. It also examines whether the differences in physical activity measured by the two methods are comparable between age and gender groups. As a result, we found for accelerometers that only 1 in 25 respondents (4%) achieved the WHO recommendation of 60 minutes of daily physical activity. Self-reported physical activity by questionnaire was slightly higher (9%) but also very low. Differences between the methods were smaller in younger children than in older age groups. The older the subjects, the lower the proportion of those who meet the WHO recommendation on any given day, with girls less likely to meet the recommendation than boys in all age groups. Children and adolescents living in Germany thus meet the WHO recommendation for physical activity only to a very small extent. While younger children are much more active in their free play, children over the age of 10 and girls in particular should be the target of physical activity programs.

Finally, in a last article we analyzed a **"Typical Day"**. In this, we looked at an average day in which physical activity intensity was measured using accelerometers. The goal was to better understand how device-based physical activity data differs between school and weekend days in children and adolescents. We determined various aspects of physical activity behavior intensity, including absolute and relative levels of light, moderate, and vigorous physical activity, sedentary behavior, and wear time. We found that the wear time of the participants increases with age, presumably because the waking phase during the day increases. However, the percentage of vigorous physical activity remained constant at about 3% across all age groups, with girls consistently engaging in less vigorous physical activity than boys. This leads to an absolute increase in vigorous activity at later ages of 5 minutes on average. No significant differences were found between boys and girls in wear time. Interestingly, however, especially Friday as the longest day and Sunday as the shortest day differ when absolute waking times are considered compared to the other days of the week. Surprisingly, the percentage distributions of intensities on these days are also almost the identical as during all the rest of the days. Therefore, one of the core results is that the movement behavior should be considered from waking to waking and not as a fixed 24-hour behavior cycle.

In conclusion, the fact remains that children today do not move enough. No matter which algorithms are used to evaluate the data, the result remains basically the same. This leads to a variety of health problems, such as obesity and an increased risk of cardiovascular disease.

At the moment, however, we are at a **crucial transition point** where we need to take additional measures to capture children's physical behavior as reality-based as possible. In this way, their physical behavior can be supported as feasible in the future. More precise measurements will allow critical decisions to be made based on the most accurate data, which in turn will lead to improved effectiveness and performance of interventions. **Precise measurement devices** are therefore of the utmost importance for physical activity behavior research. Otherwise, we are creating data with an incorrect baseline that will be interpreted by other researchers who may develop inappropriate interventions based on it. Thus, pure methodology must continue to improve while becoming less and less biased by subjective choices in algorithm selection. Thus, more validation studies with short epoch lengths are needed, especially in young children, and the formerly in-house ActiGraph signal preprocessing algorithms, now released by the manufacturer, must be used to compare the data already collected with studies that have used devices from other manufacturers.

We also cannot afford to just collect data and not analyze it. It cannot be the case that large, pooled databases such as ICAD do not include additional studies due to lack of financial and human resources. Here, policymakers must provide further resources to ensure more comprehensive and accurate reporting. The price of failure to do so would be too severe. The data sets exist, they are in the right format, they are collected using the same study design, and they just need to be merged and analyzed. In the case of the ICAD database, this can be done not only at the national level but also at the international level. Therefore, the WHO, the EU and also the governments of the individual countries must collectively provide further resources so that this **global issue** can be studied in more detail. Just talking about the pandemic of inactivity will not help, the existing data must be analyzed in depth, and big data is the appropriate approach for this.

It will also become increasingly important to record the entire period over 24 hours of a day, over a week or more. In addition, the time when the sensors are not worn must be reduced to a minimum. This can be achieved by using, for example, smaller and waterproof devices but also sensors directly attached to the body or even subdermally implanted sensors in the future. This, together with open source and data pooling methods, will allow us to further reduce the large gaps in the accurate acquisition of physical behavior.

The form of **international collaboration** at the international CAPA workshop has to be refocused now and established on a regular basis. There is a need to bring international experts in sports science, sleep research, and sports computer science together with electrical engineers and accelerometer experts, data pooling experts, and Big Data analysts to the same table. This will allow new methods of data acquisition to be refined to the point where bias in data collection is further reduced and we can offer a realistic picture of physical behavior in the future. This is where CAPA can play a central role and become an international meeting place and catalyst in physical behavior research.

Zusammenfassung

Laut Global Health Risks Report der Weltgesundheitsorganisation (WHO) ist etwa ein Drittel aller Todesfälle weltweit auf einige wenige Risikofaktoren zurückzuführen. Unzureichende körperliche Aktivität steht dabei an vierter Stelle der Risikofaktoren. Als körperliche Aktivität einer Person wird dabei jede Bewegung verstanden, die von der Skelettmuskulatur unter Verbrauch von Energie ausgeführt wird. Kurz gesagt bewegen sich viele Menschen einfach allgemein noch zu wenig. Die WHO erfasst die weltweiten Trends der unzureichenden körperlichen Aktivität regelmäßig mit großen gepoolten Datenanalysen. Ein Ergebnis der letzten Analysen ist, dass die globale altersstandardisierte Prävalenz von unzureichender körperlicher Aktivität zuletzt bei fast einem Drittel lag. Der Unterschied zwischen den Geschlechtern war sogar noch größer und für junge Menschen ist die Situation zusätzlich kritischer. Insgesamt war weltweit nur ein Fünftel der Kinder und Jugendlichen in ausreichendem Maße körperlich aktiv. Ein großer Teil der Heranwachsenden hält sich also nicht an die aktuellen Empfehlungen von täglich mindestens 60 Minuten moderater körperliche Aktivität, was sich auf ihren aktuellen und zukünftigen Gesundheitszustand auswirken kann.

Doch wie viel körperliche Aktivität und Bewegung sind genug und wie wird dies überprüft? Eine der ersten Empfehlungen zu dieser Frage wurde 1975 vom American College of Sports Medicine veröffentlicht. Diese Publikation hatte einen großen Einfluss auf die Sportwissenschaft und wurde mehrfach überarbeitet. Im Jahr 2010 veröffentlichte die Weltgesundheitsorganisation die erste internationale Version der Bewegungsrichtlinien. Darauf aufbauend haben viele Länder diese Richtlinien auch als nationale Empfehlungen übernommen, darunter auch Deutschland. Der von der WHO empfohlene Umfang an wöchentlicher oder täglicher körperlicher Aktivität wird als Referenz herangezogen, um festzustellen, welche Personen ein ausreichendes Maß an körperlicher Aktivität erreichen. Infolge dieser weltweit einheitlichen Empfehlungen wurde in vielen Studien per Fragebogen untersucht, inwieweit sich die (Nicht-) Einhaltung dieser Richtlinien auf die Gesundheit auswirkt.

Bevölkerungsbasierte Studien, die die Einhaltung von diesen Empfehlungen in Deutschland untersuchen, sind die DEGS-, KiGGS- und MoMo-Studie. Die „Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland"(KiGGS) erfasst körperliche Aktivität, Gesundheitsparameter sowie soziale und umweltbedingte Determinanten von

Kindern und Jugendlichen in Deutschland. Die **Motorik Modul (MoMo)** Längsschnittstudie ist seit 2003 ein repräsentatives Modul der KiGGS-Studie und gleichzeitig eine Vertiefungsstudie zur Erfassung der körperlichen Aktivität und motorischen Leistungsfähigkeit, die nach der Basisstudie (2003-2006) vom Karlsruher Institut für Technologie in drei Erhebungswellen (Welle 1 2009-2012, Welle 2 2015-2017, Welle 3 2018-2022) durchgeführt wurde. Auch diese Dissertation wurde hauptsächlich im Rahmen der MoMo-Studie durchgeführt, wobei der Schwerpunkt auf der Messung der körperlichen Aktivität mit Beschleunigungsmessern lag.

Bis 2015 wurde zur Erfassung der körperlichen Aktivität im Rahmen der MoMo-Studie ausschließlich der **MoMo-Aktivitätsfragebogen** eingesetzt. Der Fragebogen konzentriert sich auf die Erfassung der zentralen Aspekte Art, Intensität, Häufigkeit und Dauer der typischen bzw. gewohnheitsmäßigen Aktivität. Im Jahr 2015 wurde die Erfassung der körperlichen Aktivität durch Aktivitätsmesser, sogenannte Beschleunigungssensoren, in die MoMo-Studie aufgenommen. Dies hatte den Hintergrund die Datenerhebung um eine kontinuierliche Erfassung der aktuellen Aktivitätsintensität zu erweitern, welche mit Fragebögen nur schwierig zu erfassen ist. Wie man auch an den Verkaufszahlen sieht sind **Aktivitätsmesser** inzwischen ein führender Fitnesstrend. In den Jahren 2017-18 wurden mehr als 100 Millionen Trackinggeräte und beschleunigungsmessende Smartwatches verkauft, 2020 waren es knapp 445 Millionen Einheiten. Beschleunigungssensoren, technisch auch Akzelerometer bezeichnet, haben sich als wichtige Überwachungsinstrumente in der klinischen Forschung und der Gesundheitsförderungen etabliert und werden nun auch häufiger in nationalen Forschungsstudien eingesetzt. ActiGraph-Akzelerometer gehörten zu den ersten Beschleunigungsmessern, die schon früh für die Untersuchung von körperlicher Aktivität und Gesundheit validiert und eingesetzt wurden. Im Jahr 2014 verwendeten mehr als 51 % von 76 Studien mit jeweils mehr als 400 Teilnehmern in 36 Ländern einen ActiGraph-Akzelerometer. Daher entschieden sich auch KiGGS und MoMo für die Nutzung von ActiGraph-Akzelerometern in Welle 2, um die Vergleichbarkeit mit anderen großen epidemiologischen Studien zu gewährleisten.

Heutzutage bieten Akzelerometer die Möglichkeit, valide Daten über die Intensität und den Umfang des körperlichen Verhaltens in Echtzeit über mehrere Tage und Wochen zu erfassen. Um die einzigartigen Aspekte des körperlichen Verhaltens zu beschreiben, können aus diesen **multidimensionalen Daten** eine große Anzahl von Metriken abgeleitet werden. Als Forscher muss man sich bei der Auswahl und Anwendung von Geräten zur Quantifizierung des körperlichen Verhaltens aber unter anderem mit sehr vielen technischen Details der Akzelerometrie auseinandersetzen. Die Auswirkungen dieser technischen Entscheidungen auf die Messgrößen (Energieverbrauch, Aktivitätsintensität, Körperposition und Aktivitätsmuster) können auf verschiedene Weise erfolgen. Das Gerät, die Trageposition (Hüfte, Handgelenk, Oberschenkel) und die Aufzeichnungsparameter (Epochenlänge, Frequenz, Speicherkapazität, Aufzeichnungsfrequenz und Signalfilter) haben einen großen Einfluss auf die gemessene Aktivität. Verschiedene Hintergründe wie Studiendesign (Zweck, wiederholte Messungen) und Dauer (Zeitrahmen, Tragezeit) sowie Datenspeicherung und -auswertung müssen bei der Festlegung der Parameter berücksichtigt werden. Zusätzlich müssen bei der Auswertung mehrere Stellschrauben (Rohdaten, Kontextinformationen, Nicht-Tragezeit, Intensitätsklassifizierung, Compliance) in Abhängigkeit von den Zielgrößen justiert werden. In MoMo ergibt dies beispielsweise ca. 18 Millionen Datenpunkte bei jedem Probanden der einen Akzelerometer für eine Woche trägt. Die Daten aller Probanden gilt es möglichst Zeit effizient zu verarbeiten, aufzubereiten und anschließend zu analysieren.

Die **Hauptziele dieser Dissertation** waren daher, 1) zu diskutieren wie technische Entscheidungen beim Einsatz von Akzelerometern in großen epidemiologischen Studien der Aktivitätsforschung getroffen werden sollten, 2) welche grundlegenden methodischen Aspekte deshalb bei der gerätegestützten Messung der körperlichen Aktivität speziell in der MoMo-Studie angewendet wurden, 3) welche Auswirkungen verschiedene Auswertungsmethoden bei der Vorverarbeitung von Akzelerometerdaten haben, 4) worin sich die Daten der Akzelerometererhebung zu den repräsentativen Daten des MoMo-Aktivitätsfragebogen unterscheiden und 5) welche Ergebnisse sich bei den Akzelerometerdaten im Hinblick auf Unterschiede zwischen Wochentagen und Wochenendtagen ergeben. Diese Fragestellungen wurden in fünf wissenschaftlichen Artikeln untersucht.

In einem **„Überblicksartikel"** spiegeln wir zuerst den Expertenkonsens zum Thema „Beschleunigungsmessung zur Beurteilung des Bewegungsverhaltens"während des 2. internationalen Workshops des Center for the Assessment of Physical Activity (CAPA) wider. Der Artikel verdeutlicht, welche Aspekte bei der Planung und Auswertung von Akzelerometerstudien berücksichtigt werden sollten. Eines der wichtigsten Ergebnisse des Workshops ist, dass man sich gründlich mit bestehenden Validierungsstudien vertraut machen muss und möglichst viele technische Entscheidungen bei der Datenerfassung und -auswertung dokumentiert, um später einen studienübergreifenden Datenvergleich zu ermöglichen. Wir schlagen vor, dass auf der Grundlage des validesten Ansatzes die jeweilige Verhaltensmetrik und die zugehörig passende Methode gewählt wird. Unter diesem Aspekt bestimmt die gewählte Methode die Art des Geräts und den Vorhersagealgorithmus zur Bestimmung der körperlichen Aktivität. Aufgrund der großen Anzahl an möglichen „subjektiven" Entscheidungen vor der eigentlichen Messung verschafft der Artikel einen Überblick über die technischen Details der Beschleunigungsmessung die es zu beachten gilt. Gleichzeitig skizziert er die besten Praktiken bei der Auswahl und Anwendung von Geräten zur Quantifizierung der drei wesentlichen Verhaltenskategorien, die für die Forschungsgemeinschaft von allgemeinem Interesse sind: körperlicher Aktivität, sedentärem Verhalten und Schlaf.

Bislang gab es in Deutschland keine flächendeckenden gerätegestützten Bewegungsdaten für Kinder und Jugendliche. Mit KiGGS und MoMo Welle 2 wurden erstmals bundesweit gerätegestützte Aktivitätsdaten zur körperlichen Aktivität von Kindern und Jugendlichen erhoben. Auf den Empfehlungen des CAPA-Workshops aufbauend haben wir deshalb in unserem **„Studienprotokoll"** die grundlegenden methodischen Aspekte dargelegt, welche wir bei der gerätegestützten Messung der körperlichen Aktivität speziell in der MoMo-Studie angewendet haben. Das Studienprotokoll gibt einen Überblick über die technischen Details und die grundsätzlichen Entscheidungen beim Einsatz von Akzelerometern in groß angelegten epidemiologischen Studien. Dabei sind Einschränkungen durch die vorgegebenen Filter und Auswertungsroutinen besonders zu berücksichtigen.

Um zu klären welche Auswirkungen verschiedene Auswertungsmethoden bei der anschließenden Verarbeitung von Akzelerometerdaten haben, wurde im **„Methoden-Artikel"** der Einfluss von drei speziellen Faktoren untersucht. Wir haben dabei überprüft inwieweit die Faktoren „Nicht-Tragezeit-Algorithmus", „Epochenlänge" und „Intensitäts Schwellenwerte" die Quantifizierung der mit Akzelerometern gemessenen körperlichen Aktivität beeinflussen. Gefunden wurden signifikante Unterschiede bei allen drei Faktoren. Infolgedessen können die daraus resultierenden Unterschiede, je nach Auswahl des Faktors, der geschätzten Werten für sedentäres Verhalten und körperliche Aktivität sehr groß werden. Vor allem das Ausmaß der Diskrepanz zwischen den Epochenlängen hat gezeigt, dass dies bei der Interpretation der Akzelerometerdaten besonders berücksichtigt werden muss. Die Epochenlänge ist damit maßgeblich als Einflussgröße zu betrachten, wenn man Werte körperlicher Aktivität zwischen Studien vergleicht. Wir schlagen außerdem vor, Daten zu poolen und in vereinheitlichter Weise auszuwerten. Neben neuen Validierungsstudien mit kurzen Epochenlängen für jüngere Kinder empfehlen wir auch die Durchführung von Meta-Analysen. Hierbei könnten Daten aus mehreren Studien zur Validierung von Grenzwerten verwendet werden, um einen einvernehmlichen Satz von Schwellenwerten vorzuschlagen, die in verschiedenen Settings und Studien bei Kindern und Jugendlichen verwendet werden können.

Da gerade Kinder ein komplexeres, aber weniger strukturiertes Bewegungsverhalten zeigen als Erwachsene, ist die Bestimmung ihrer vielen spontanen und impulsiven Bewegungen eine besondere Herausforderung für die Ermittlung der körperlichen Aktivität. Weder Fragebögen noch Akzelerometer bieten eine optimale Erfassung aller Facetten der körperlichen Aktivität. Aus diesem Grund wird ein kombinierter multimodaler Ansatz aus Selbstauskünften und gerätegestützten Methoden empfohlen. In einem weiteren Artikel, dem **„Vergleichsartikel"**, haben wir deshalb die mit dem Akzelerometer gemessene körperliche Aktivität und die mit dem MoMo-Fragebogen erfasste körperliche Aktivität verglichen. Anhand der Anzahl der Tage, an denen die Teilnehmenden die WHO-Empfehlung für körperliche Aktivität erreichten, wird in dieser Studie untersucht, inwieweit sich selbstberichtete und gerätebasiert gemessene körperlicher Aktivität unterscheiden. Außerdem wird untersucht ob die mit den beiden Methoden ermittelten Unterschiede in der körperlichen Aktivität zwischen Alters- und Geschlechtergruppen vergleichbar sind. Als Ergebnis fanden wir bei den Akzelerometern, dass nur jeder 25. befragte Person (4%) die WHO-Empfehlung von 60 Minuten täglicher körperlicher Aktivität erreicht. Die selbstberichtete körperliche Aktivität per Fragebogen war etwas höher (9%), aber ebenfalls sehr niedrig. Die Unterschiede zwischen den Methoden sind bei jüngeren Kindern geringer als in älteren Altersgruppen. Je älter die Probanden sind, desto geringer ist der Anteil derjenigen, die die WHO-Empfehlung an jedem Tag einhalten, wobei Mädchen die Empfehlung in allen Altersgruppen seltener einhalten als Jungen. In Deutschland lebende Kinder und Jugendliche erfüllen damit die WHO-Empfehlung für körperliche Aktivität nur in sehr geringem Maße. Während jüngere Kinder in ihrem freien Spiel viel aktiver sind, sollten vor allem Kinder über 10 Jahren und insbesondere Mädchen Ziel von Bewegungsförderungen sein.

Schlussendlich haben wir in einem letzten Artikel einen **„Typischen-Tag"** analysiert. Hierbei haben wir einen durchschnittlichen Tag betrachtet, an dem die Intensität der körperlichen Aktivität mit Hilfe von Akzelerometern gemessen wurde. Ziel war es, besser zu verstehen, wie sich die gerätegestützten Daten zur körperlichen Aktivität bei Kindern und Jugendlichen zwischen Schul- und Wochenendtagen unterscheiden. Wir haben dabei verschiedene Aspekte der Intensität von körperlichem Bewegungsverhalten bestimmt und dabei auch die absoluten und relativen Werte der leichten, moderaten und anstrengenden körperlichen Aktivität, des sedentären Verhaltens sowie die Tragezeit untersucht. Feststellen ließ sich, dass die Tragezeit der Teilnehmenden mit dem Alter ansteigt, vermutlich weil die Wachphase während des Tages zunimmt. Der Prozentsatz der anstrengenden körperlichen Aktivität blieb allerdings über alle Altersgruppen hinweg konstant bei etwa 3%, wobei Mädchen durchweg weniger anstrengenden körperliche Aktivität als Jungen ausübten. Dies führt zu einem absoluten Anstieg der anstrengenden Aktivität im Alter von im Schnitt 5 Minuten. Bei der Tragezeit wurden keine signifikanten Unterschiede zwischen Jungen und Mädchen festgestellt. Interessanterweise unterscheiden sich aber vor allem Freitag als längster Tag und Sonntag als kürzester Tag, wenn man die absoluten Wachzeiten betrachtet gegenüber den anderen Wochentagen. Erstaunlicherweise sind die prozentualen Verteilungen der Intensitäten auch an diesen Tagen annähernd gleich wie während aller restlichen Tage. Deswegen ist eins der Kernergebnisse, dass das Bewegungsverhalten vom Aufwachen bis zum Aufwachen und nicht als fester 24-stündiger Verhaltenszyklus betrachtet werden sollte.

Abschließend bleibt die Tatsache, dass sich Kinder heutzutage zu wenig bewegen. Egal, welche Algorithmen zur Auswertung der Daten verwendet werden, das Ergebnis bleibt im Grunde dasselbe. Dies führt zu einer Vielzahl von Gesundheitsproblemen, wie Übergewicht und einem erhöhten Risiko für Herz-Kreislauf-Erkrankungen.

Im Moment befinden wir uns jedoch an einem **entscheidenden Übergangspunkt**, an dem wir zusätzliche Maßnahmen ergreifen müssen, um das Bewegungsverhalten von Kindern so realitätsnah wie möglich zu erfassen. Auf diese Weise kann ihr körperliches Verhalten in Zukunft so individuell wie nur möglich unterstützt werden. Präzisere Messungen werden es erlauben, kritische Entscheidungen auf der Grundlage der genauesten Daten zu treffen, was wiederum zu einer verbesserten Wirksamkeit und Effektivität von Interventionen führen wird. **Präzise Messgeräte** sind daher von allergrößter Bedeutung für die Erforschung des körperlichen Bewegungsverhaltens. Ansonsten suggerieren wir Daten mit einer falschen Ausgangsbasis, die von anderen Wissenschaftlern interpretiert werden und darauf aufbauend gegebenenfalls unpassende Interventionen entwickeln. Die reine Methodik muss also immer besser werden und gleichzeitig immer weniger durch subjektive Entscheidungen bei der Auswahl der Algorithmen verzerrt werden. Daher sind mehr Validierungsstudien mit kurzen Epochenlängen erforderlich, insbesondere bei kleinen Kindern, und die ehemals firmeninternen ActiGraph-Signalvorverarbeitungsalgorithmen, die jetzt vom Hersteller freigegeben wurden, müssen verwendet werden, um die bereits erhobenen Daten mit Studien, die Geräten anderer Hersteller genutzt haben, zu vergleichen. Wir können es uns außerdem nicht leisten, Daten nur zu sammeln und nicht zu analysieren. Es darf nicht sein, dass große, gepoolte Datenbanken wie die ICAD mangels finanzieller und personeller Ressourcen keine weiteren Studien mit aufnehmen. Hier muss die Politik weitere Mittel bereitstellen, um eine umfassendere und genauere Berichterstattung zu gewährleisten. Der Preis des Versäumnis wäre zu hoch. Die Datensätze sind vorhanden, sie haben das richtige Format, sie werden mit dem gleichen Studiendesign erhoben und müssen nur noch zusammengeführt und ausgewertet werden. Im Falle der ICAD-Datenbank ist dies nicht nur auf nationaler, sondern auch auf internationaler Ebene möglich. Deshalb müssen die WHO, die EU und auch die Regierungen der einzelnen Länder gemeinsam weitere Mittel zur Verfügung stellen, damit dieses **globale Problem** genauer untersucht werden kann. Nur über die Pandemie der Inaktivität zu reden, hilft dabei nicht weiter, die vorhandenen Daten müssen eingehend analysiert werden, und dafür ist Big Data der geeignete Ansatz.
Außerdem wird es immer wichtiger werden, den gesamten Zeitraum über 24 Stunden eines Tages, über eine Woche oder mehr aufzuzeichnen. Zusätzlich gilt es die Zeit, in der die Sensoren nicht getragen werden können, auf ein Minimum zu reduzieren. Dies kann bewerkstelligt werden, indem zukünftig beispielsweise kleinere und wasserdichte Geräte aber auch direkt am Körper angebrachte oder sogar subdermal implantierte Sensoren verwendet werden. Dies wird es uns zusammen mit Open-Source- und Data-Pooling-Methoden ermöglichen, die großen Lücken bei der tatsächlichen Erfassung des Bewegungsverhaltens, weiter zu verringern.
Die Form der **internationalen Zusammenarbeit** beim internationale CAPA-Workshop

gilt es nun neu auszurichten und regelmäßig zu etablieren. Es ist notwendig internationalen Experten aus der Sportwissenschaft, der Schlafforschung und der Sportinformatik zusammen mit Elektrotechnikern und Experten für Beschleunigungssensoren, Experten für Datenpooling und Big-Data-Analysen an einen Tisch zu bringen. Dadurch können neue Methoden der Datenerhebung so weit verfeinert werden, dass die Verzerrung in der Datenerfassung weiter verringert wird und wir zukünftig ein realistisches Bild des Bewegungsverhaltens bieten können. Hier kann die CAPA eine zentrale Rolle spielen und zu einem internationalen Treffpunkt und Katalysator in der körperlichen Verhaltensforschung werden.

Contents

Acknowledgments **iii**

Summary **v**

Zusammenfassung **xi**

List of figures **xix**

List of tables **xxi**

List of Abbreviations **xxiii**

Preface **xxv**
0.1 Published Research xxv
0.2 Motorik-Modul Study (MoMo) xxvi

1 Introduction **1**
1.1 Physical Activity and Health 1
1.2 Documentation of device-based measured PA 8
1.3 Research Problem and relevance 10
1.4 Structure of the thesis 11
References Introduction 15

2 State of the science, best practices and future directions **21**
2.1 Introduction 21
2.2 State of the science 22
2.3 Best practices 24
2.4 Future directions 37
2.5 Conclusion 40
References Consensus Article 43

3 Measurement of Physical Activity with Accelerometry in MoMo **55**
3.1 Abstract 55
3.2 Introduction 56
3.3 Methods 57
3.4 Results 63
3.5 Discussion 64
References Study Protocol 67

4 Processing accelerometer data of children and adolescents in MoMo **73**

4.1 abstract . . . 73
4.2 Background . . . 74
4.3 Methods . . . 76
4.4 Results . . . 78
4.5 Discussion . . . 81
4.6 Conclusion . . . 87
References Methods-Paper . . . 89

5 Comparison of self-reported & device-based, measured physical activity 95
5.1 Abstract . . . 95
5.2 Background . . . 96
5.3 Methods . . . 96
5.4 Results . . . 100
5.5 Discussion . . . 102
5.6 Conclusion . . . 105
References Comparison Article . . . 107

6 Typical day and influence of weekend on physical activity 111
6.1 Abstract . . . 111
6.2 Background . . . 112
6.3 Methods . . . 113
6.4 Results . . . 115
6.5 Discussion . . . 116
6.6 Conclusion . . . 121
References Typical Day Article . . . 125

7 Discussion and Conclusion 129
7.1 Issue of methods . . . 131
7.2 Issue of Development . . . 134
7.3 Issue of Comparability . . . 139
7.4 Final Discussion . . . 146
7.5 Conclusion . . . 151
References Discussion and Conclusion . . . 153

A Appendix 165
A.1 Comparison WHO Guidelines 2010 vs. 2020 . . . 165
A.2 Supplement Material Chapter 4 . . . 167
A.3 Poster Presentations . . . 172

B Publications 175

List of Figures

1 Preface: Logo Motorik Modul Study xxvi

2 Preface: MoMo Samplepoints xxvi

3 Preface: Physical Activity Measurement in MoMo xxvii

1.1 Introduction: Activity Pyramid 2

1.2 Introduction: Scopus “sedentary behavior” 6

1.3 Introduction: 24-hour Physical Behavior Cycle 7

1.4 Introduction: PA Questionnaire Cube 8

1.5 Introduction: Scopus Search “physical activity and acceleromet*” 10

1.6 Introduction: Thesis Structure 13

2.1 Consensus-Article - Figure 1: CAPA Workshop Logo 23

3.1 Study Protocol - Infographic 56

3.2 Study Protocol - Figure 1: ACC Study Design 58

3.3 Study Protocol - Figure 2: Placement of the Device 61

3.4 Study Protocol - Figure 3: Wear time-protocol 62

4.1 Methods-Article - Figure 1: Histogram of WT per hour, for NWT 78

4.2 Methods-Article - Figure 2: Marginal Means of WT for diff. EL and NWT 79

4.3 Methods-Article - Figure 3: Mean Minutes in SED, LPA, MPA & VPA for diff. EL 81

4.4 Methods-Article - Figure 4: Mean Minutes in MVPA for different EL . . 82

4.5 Methods-Article - Figure 5: Cut Point Ranges for 15s EL 85

5.1 Comparison Article - Figure 1: Cross Table Heat Map 100

5.2 Comparison Article - Figure 2: Specific Cross Table Heat Maps 102

5.3 Comparison Article - Figure 3: Differences Plot for Age and Gender . . . 103

6.1 Typical Day Article - Figure 1: MoMo – Activity Pyramid 115
6.2 Typical Day Article - Figure 2: Typical Day 116
6.3 Typical Day Article - Figure 3: Different Intensities across all Weekdays . 117
6.4 Typical Day Article - Figure 4: Different Intensities across all Weekdays and Age 118

7.1 Discussion: Device Placement 135
7.2 Discussion: ActivPal Icon Strip 136
7.3 Discussion: ICAD Logo 139
7.4 Discussion: MO|RE data Logo 139
7.5 Discussion: WHO Guideline Cross Table Heat Map 141
7.6 Discussion: Approaches to assess Physical Behavior 142
7.7 Discussion: 24-hour Physical Behavior Cycle 144
7.8 Discussion: Measuring PA Cubes 147

A.1 Comparison Article - WHO Comparison Heatmaps 166
A.2 Supplement Figure 1: Q-Q Plots for SED 168
A.3 Supplement Figure 2: Q-Q Plots for LPA 169
A.4 Supplement Figure 4: Q-Q Plots for MPA 170
A.5 Supplement Figure 4: Q-Q Plots for VPA 171
A.6 Poster: ISPAH 2018 173
A.7 Poster: dvs Hochschultag 2019 174

List of Tables

3.1 Study Protocol - Table 1: Dropout MoMo 59
3.2 Study Protocol - Table 2: Dropout KiGGS 59
3.3 Study Protocol - Table 3: Accelerometer Criteria 63

4.1 Methods-Article - Table 1: Cut Point Ranges for 1s EL 77
4.2 Methods-Article - Table 2: Valid Days depending on valid hour, epoch length, and NWT 80

5.1 Comparison Article - Table 1: Characteristics of Participants 98
5.2 Comparison Article - Table 2: Accelerometer Criteria 99
5.3 Comparison Article - Table 3: Mean data and 95%-CI 101

6.1 Typical Day Article - Table 1: Participant Characteristics 114
6.2 Typical Day Article - Table 2: Differences of WT on Weekdays and Weekend days 119
6.3 Typical Day Article - Table 3: Overall WT in different Intensities 122
6.4 Typical Day Article - Table 4: Minutes of WT on Weekdays and Weekend days 123

7.1 Discussion - Table 1: Accelerometer Criteria 132

A.1 Supplement - Chapter 4 Table 1 167

List of Abbreviations

MoMo Motorik-Modul Study

KIT Karlsruhe Institute of Technology

KiGGS German Health Interview and Examination Survey for Children and Adolescents

RKI Robert Koch-Institute

DEGS German Health Interview and Examination Survey for Adults

WHO World Health Organization

ICAD International Children's Accelerometry Database

CAPA Center for the Assessment of Physical Activity

NHANES National Health and Nutrition Examination Survey

HELENA Healthy Lifestyle in Europe by Nutrition Adolescence

EHYS European Youth Heart Study

DOI Digital Object Identifier

PA Physical Activity

PB Physical Behavior

MET Metabolic Equivalent

SED Sedentary Behavior

LPA Light Physical Activity

MPA Moderate Physical Activity

VPA Vigorous Physical Activity

MVPA Moderate to Vigorous Physical Activity

EE Energy Expenditure

PAQ Physical Activity Questionnaire

ACC Accelerometer

EMA Ecological Momentary Assessment

AA Ambulatory Assessment

AGD-files Count-based ActiGraph Data files

GT3X ActiGraph Accelerometer models wGT3X-BT and GT3X+

NWT Non-Wear Time

WT Wear Time

EL Epoch Length

VA Vertical Axis

VM Vector Magnitude

BMI Body Mass Index

ISM Isotemporal Substitution Model

CoDA Compositional Data Analysis

ENMO Euclidean Norm Minus One

MAD Mean Amplitude Deviation

GGIR R-package to process multi-day raw accelerometer data

Preface

0.1 Published Research

Parts of this thesis have been published or are under review in peer-reviewed journals. Thus, the following chapters can be read independently from each other:

CHAPTER 2: Burchartz, Alexander; Anedda, Bastian; Auerswald, Tina; Giurgiu, Marco; Hill, Holger; Ketelhut, Sascha; Kolb, Simon; Mall, Christoph; Manz, Kristin; Nigg, Claudio R.; Reichert, Markus; Sprengeler, Ole; Wunsch, Kathrin; Matthews, Charles E. (2020): Assessing physical behavior through accelerometry – State of the science, best practices, and future directions. In: **Psychology of Sport and Exercise** 49, S. 101703. DOI: 10.1016/j. psychsport.2020.101703.

CHAPTER 3: Burchartz, Alexander; Manz, Kristin; Anedda, Bastian; Niessner, Claudia; Oriwol, Doris; Schmidt, Steffen CE; Woll, Alexander (2020): Measurement of Physical Activity and Sedentary Behavior by Accelerometry Among a Nationwide Sample from the KiGGS and MoMo study: Study Protocol. In: **JMIR Research Protocols** 9 (7), e14370. DOI: 10.2196/14370.

CHAPTER 4: Burchartz, Alexander; Kolb, S.; Klos, Leon.; Schmidt, S. CE; Haaren-Mack, B. von; Niessner, C.; & Woll, A. (2023): How specific combinations of epoch length, non-wear time and cut-points influence physical activity - Processing accelerometer data from children and adolescents in the nationwide MoMo study. In: **German Journal of Exercise and Sport Research**.Advance online publication. DOI: 10.1007/s12662-023-00892-9

CHAPTER 5: Burchartz, Alexander; Oriwol, Doris; Kolb, Simon; Schmidt, Steffen CE; Wunsch, Kathrin; Manz, Kristin; Niessner, Claudia; Woll, Alexander (2021): Comparison of self-reported & device-based, measured physical activity among children in Germany. In: **BMC public health** 21 (1). DOI: 10.1186/s12889-021-11114-y.

CHAPTER 6: Burchartz, Alexander; Oriwol, D.; Kolb, S.; Schmidt, S. CE; Haaren-Mack, B. von; Niessner, C.; & Woll, A. (2022). Impact of weekdays versus weekend days on accelerometer measured physical behavior among children and adolescents: Results from the MoMo study. **German Journal of Exercise and Sport Research**, 52(2), 218–227. 10.1007/s12662-022-00811-4

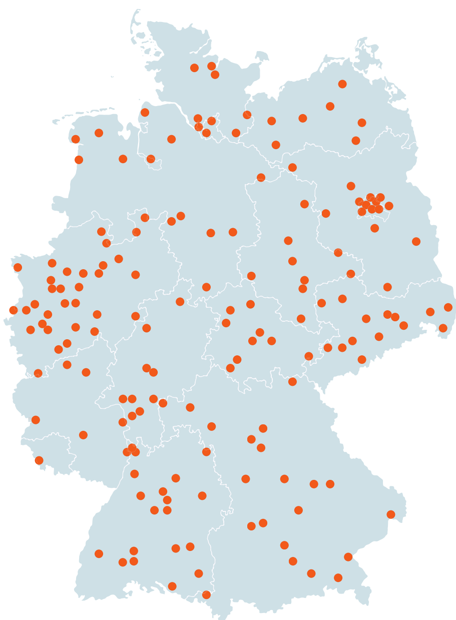
0.2 Motorik-Modul Study (MoMo)

This work was developed within the Motorik-Modul Study (MoMo) (2009-2022): Physical fitness and physical activity as determinants of health development in children and adolescents. The MoMo study logo can be seen in Figure 1. The main objective of MoMo is to track and report physical activity and physical fitness of children and adolescents in a nationwide sample, and significant effort was put into collecting representative data from 167 sample points throughout Germany (see Figure 2).

Figure 1: The logo of the Motorik Modul Study (MoMo)

Motor performance and physical activity are important aspects of healthy development in childhood and adolescence. Until recently, it has not been possible to reliably answer how healthy, active, and motorly fit children and adolescents in Germany are. The Motorik-Modul (MoMo) aims to close this research gap. MoMo is a joint project of the Karlsruhe Institute of Technology (KIT) and the Karlsruhe University of Education (PH Karlsruhe) in cooperation with the German Health Interview and Examination Survey for Children and Adolescents (KiGGS) of the Robert Koch-Institute (RKI). The German Health Interview and Examination Survey for Children and Adolescents (KiGGS) study was funded by the German Federal Ministry of Health, the Ministry of Education and Research, and the Robert Koch Institute. This work was supported by the Federal Ministry of Education and Research (01ER1503) within the long-term research program in public health research. A positive vote of the ethics committee of the Karlsruhe Institute of Technology on September 23, 2014, is available for the study.

Figure 2: Representation of the 167 MoMo sample points all over Germany.

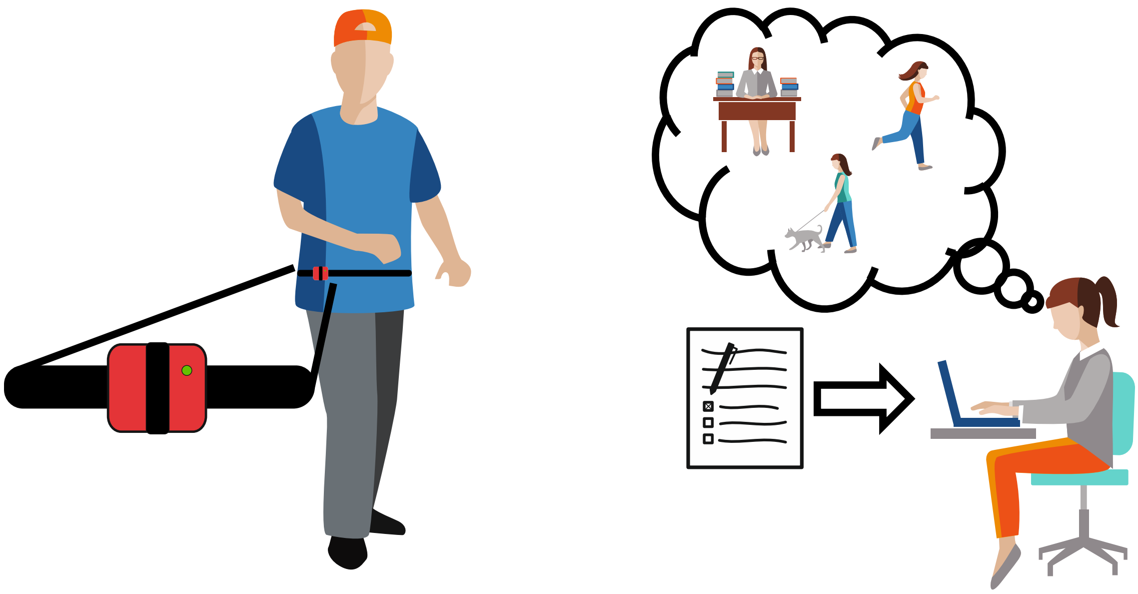


Since participants are selected on a representative nationwide basis, it is possible to make statements about the motor development and activity behavior of children and adolescents throughout Germany. The MoMo study aims to make a long-term contribution to improving the health situation of children and adolescents in Germany. For example, the following questions should be answered: Do active and "fit" children and adolescents remain active, and does physical activity influence chronic diseases? To this end, an analysis of factors that influence physical activity and motor performance will be carried out and their effects on health development in childhood and adolescence will be investigated.

Through cooperation with the RKI, particularly with KiGGS, interactions between motor performance, physical activity, and the health status of children and adolescents can also be derived. The findings of the MoMo study serve as a basis for planning and developing targeted measures related to exercise for prevention and promotion of health.

The focus of this thesis is on the aspect of the MoMo study that deals with measuring physical activity, more precisely, with acquiring physical activity using accelerometers.

Figure 3: Representation of the physical activity measurement in MoMo by accelerometers on the left in contrast to questionnaires on the right.



1. Introduction

1.1 Physical Activity and Health

A quote attributed to Hippocrates, a physician from ancient Greece referred to by many as the "father of medicine," showed an early awareness of the value of physical activity to human health:

> *"If we could give every individual the right amount of nourishment and exercise, not too little and not too much, we would have found the safest way to health."*

In ancient Greece, Hippocrates knew that the right dose of activity leads to better health. Unfortunately, almost two and a half thousand years later, it still seems that humanity has not comprehended this, while in the meantime the human lifestyle has changed dramatically. According to the 2010 World Health Organization (WHO) Global Health Risks Report (World Health Organization, 2010), about one-third of all deaths worldwide can be attributed to a few risk factors. Physical inactivity ranks fourth among the highest risk factors that can be influenced by the individual. Our activity behavior has changed considerably due to technological but also social changes, especially in the last century. The global target for physical activity set for 2025 (a 10% relative reduction in insufficient physical activity) cannot be achieved according to the WHO. Physical activity can mean many things. What exactly does physical activity mean? One of the most well-known definitions is

> *"Physical activity is defined as any bodily movement produced by skeletal muscles that results in energy expenditure"*
>
> Caspersen et al., 1985, p. 126

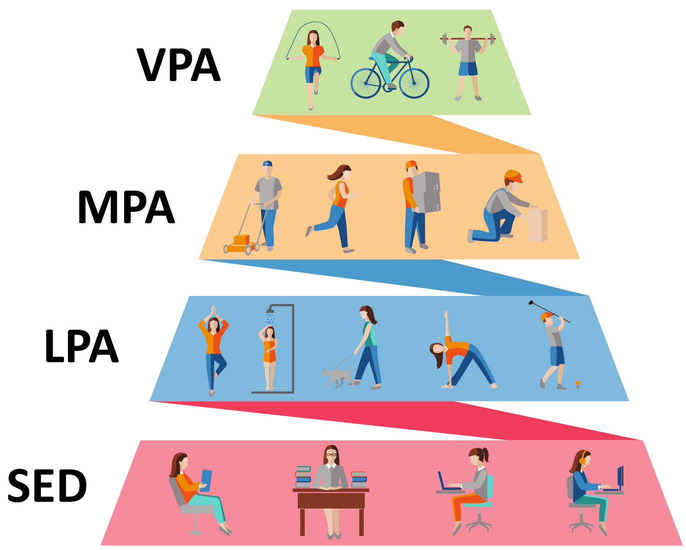
Physical activity (PA) as it is understood in this article includes all types of physical activity such as walking, running, racing, dancing, and playing with light (LPA), moderate

(MPA), and vigorous (VPA) intensities. All areas of daily life are also included (e.g. school, work, transport, leisure, home, or sports). The intensity of physical activity is quantified by the energy expenditure required to perform that activity. The metabolic equivalent of activity is often used as a criterion of energy expenditure, namely, METs. 1 MET describes the measure of oxygen consumption for energy production at rest and is expressed in ml of O_2 consumed per kg of body weight for one minute (ml / kg / min). The multiples reflect the energy expenditure during activity in relation to the resting metabolic rate. On average, a male adult consumes 3.5 ml / kg / min at baseline, and a female consumes approximately 3.15 ml/kg/min. Usually, the male value is used as a reference (that is, 1 MET = 3.5 ml/kg/min). Relative METs can be derived from measuring the baseline energy expenditure of a particular individual. Ainsworth et al. (1993) recorded the associated METs for many activities in their "Compendium of Physical Activities." Usually, physical activity intensities are broken down into different subsets (compare also Figure 1.1:

- LPA (1.5-3 METs),
- MPA (3-6 METs),
- and VPA (>6 METs).

Physical activity and health domains have traditionally been more concerned with the moderate to vigorous (MVPA) domain; however, as stated above, this work uses the term PA for the range from light activity upward (LPA, MPA, and VPA).

Figure 1.1: Illustration of the different activity intensities



"Sedentary behavior (SED)" in contrast to PA, refers to awake behavior characterized by energy expenditure ≤ 1.5 metabolic equivalents (METs) (Tremblay et al., 2017). Typical behaviors with low energy expenditure include sitting, lying, standing, leaning, and reclining. Interestingly, all behaviors occur without locomotion, which is the main difference from PA.

From the perspective of movement, **"Sleep"** is defined by external rest with very little movement during a prolonged period of time, typically in a lying position, usually with the eyes closed. Many vital signs such as pulse, respiratory rate, and blood pressure decrease during sleep in contrast to wakefulness and are characterized by energy expenditure ~ 1 metabolic equivalent (MET).

There is a great deal of confusion about the term **"Physical Inactivity"**. This does not mean sedentary behavior, but rather an insufficient physical activity level to meet present physical activity recommendations (World Health Organization, 2010). An example can better illustrate this difference: One can be physically active enough to reach the target of 60 minutes of moderate to vigorous exercise per day while accumulating a lot of sedentary behavior during the rest of the day.

Just recently the WHO did two pooled data studies (Guthold et al., 2018, 2020) and looked at worldwide trends in insufficient physical activity. The first study (Guthold et al., 2018) included data from 358 population-based surveys in 168 countries with 1.9 million participants to report the prevalence of insufficient physical activity. They used data on physical activity at work, at home, for transport and during leisure time and compared it with WHO physical activity guidelines. Here, in 2016, the global age-standardized prevalence of insufficient physical activity was 27.5%. The gender difference was greater than 8 percentage points (23.4%, for men versus 31.7%, for women). For young people, the situation is even worse. Using cross-sectional studies from 146 countries in a second study (Guthold et al., 2020), they analyzed approximately 1.6 million participants (ages 11-17) and found that a total of 81.0% were not physically active sufficiently worldwide in 2016. This includes 77.6% of boys and 84.7% of girls. Therefore, a large proportion of adolescents do not follow current recommendations for daily physical activity, which can affect their current and future health status.

But how much activity is enough? *"... not too little and not too much ..."* is of limited help.

One of the first guidelines for this question was published by the American College of Sports Medicine in 1975: the Guidelines for Graded Exercise Testing and Exercise Prescription (American College of Sports Medicine, 1975). This publication had a great influence on sports science and was revised several times. However, the greatest influence on modern guidelines was probably the recommendation from the American College of Sports Medicine and the Centers for Disease Control and Prevention, which first recommended exercise for at least 30 minutes in 1995 (Pate et al., 1995). A subsequent nationwide recommendation was issued by the US government in 2008, which also explicitly related health to physical activity (U.S. Department of Health and Human Services, 2008). In 2010, the World Health Organization (WHO) published the first international version of these guidelines (World Health Organization, 2010). Building on this, many countries also adopted these guidelines as national recommendations, including Germany (Rütten & Pfeifer, 2017). The level of weekly or daily physical activity recommended by the WHO is used as a benchmark to determine which individuals achieve a sufficient level of physical activity. As a result of these worldwide standardized recommendations, many studies investigated to what extent (non)adherence to these guidelines has an impact on health.

Population-based studies that examine compliance with guidelines in Germany are the DEGS, KiGGS, and MoMo study. The German National Health Interview and Examination Survey later renamed the German Health Interview and Examination Survey for Adults (Studie zur Gesundheit Erwachsener in Deutschland (DEGS)) had already been

in existence since 1997. The aspects of physical activity measured by DEGS include the frequency, duration and intensity of physical activity, in different domains (leisure time, transportation, work, home), as well as sedentary behavior in different socioeconomic and age groups (young people, adults and older adults) (Scheidt-Nave et al., 2012).
The second study in Germany that captures physical activity, health parameters, as well as social and environmental determinants is the German Health Interview and Examination Survey for Children and Adolescents (KiGGS, Kurth et al., 2008). The Motorik Modul (MoMo) longitudinal study has been a representative module of the KiGGS study since 2003 (Woll et al., 2017). In MoMo, an in-depth study is conducted to assess physical activity and motor performance of children and adolescents living in Germany and was conducted by the Karlsruhe Institute of Technology for three survey waves (Wave 1 2009-2012, Wave 2 2015-2017, Wave 3 2018-2022) after the baseline study (2003-2006). This thesis was also carried out primarily within the framework of the MoMo study with a focus on measuring physical activity with accelerometers. More detailed information on MoMo can be found in Chapter 0.2.

However, until 2013, there were still no recommendations or even national guidelines on physical activity in Germany. After the first results of the first MoMo survey wave were available, Alexander Woll gathered a group of experts around Graf et al. (2013) to publish an expert consensus with suggestions to promote physical activity among children and adolescents in Germany. This group consisted of interdisciplinary scientists and representatives of selected professional societies and associations. They were tasked with considering German age and gender specific characteristics, as well as the characteristics of selected groups concerning special access routes, for example, children from educationally disadvantaged families or with a migration background, and regional differences, and with developing data-based new recommendations explicitly for Germany. The recommendations of Graf et al. (2013) finally included that children and adolescents in Germany should achieve a daily physical activity time of 90 min or more and that daily activities, for example, the active way to school should be explicitly promoted. Furthermore, it was recommended that a daily step count of at least 12,000 steps be achieved in daily life. Subsequently, and in part on the basis of this work, the National Recommendations for Physical Activity and Physical Activity Promotion were published in 2017 following a resolution of the German Bundestag (Rütten & Pfeifer, 2017). Two special characteristics distinguished these recommendations: On the one hand, scientifically based and systematically prepared official recommendations were formulated for Germany for the first time; on the other hand, recommendations for physical activity and promotion of physical activity were systematically linked for the first time, also from an international perspective. For children and adolescents, a daily physical activity time of 90 minutes and more was recommended at moderate to high intensity, 60 minutes of which can be completed through daily activities, such as at least 12,000 steps/day. Additionally, sedentary activities and the use of screen media (TV, computer/tablet, smartphone, etc.) should be kept to a minimum. For elementary school children, the recommendation is a maximum of 60 minutes/day, for adolescents a maximum of 120 minutes/day (Rütten & Pfeifer, 2017).

Based on the experience of studies around the world and developments in the technological field, updated WHO recommendations for physical activity were issued in 2020 (Chaput et al., 2020). These recommendations now include self-reported evidence and modern wearable technology as tools to assess physical activity. Since then, it has been recommended that

> *" children and adolescents should do at least* **an average** *of 60 minutes per day of moderate to vigorous intensity, mostly aerobic physical activity, throughout the week, and vigorous intensity aerobic activities, as well as those that strengthen muscles and bone, should be incorporated at least 3 days a week. "*
>
> Chaput et al., 2020, p. 6

This new recommendation from Chaput et al. (2020), which now looks at the average time spent in physical activity per week, rather than the previous daily time, is very easy to implement, especially for device-based records. On the other hand, it is important to note that now days with activity times longer than 60 min would compensate for those with less activity (Colley et al., 2017). Still, the daily stimulus is very important in children (Dwyer et al., 1983). Taking a look at the exact times spent with MVPA every day will result in fewer days of at least 60 min MVPA when both evaluation methods are compared (Colley et al., 2017). However, the main reason, since the study part with device-based measured physical activity was already planned and started in 2014, the questions in the questionnaire still referred to the 2010 WHO Guidelines (World Health Organization, 2010). In 2020, the youth activity recommendations have changed from a recommendation of at least 60 minutes per day to a recommendation of an average of 60 minutes per day (Chaput et al., 2020). This adaptation will require changes in survey questions and sampling methods for future monitoring.

A second recommendation of the new WHO Guidelines (Chaput et al., 2020) are limitations for sedentary behavior:

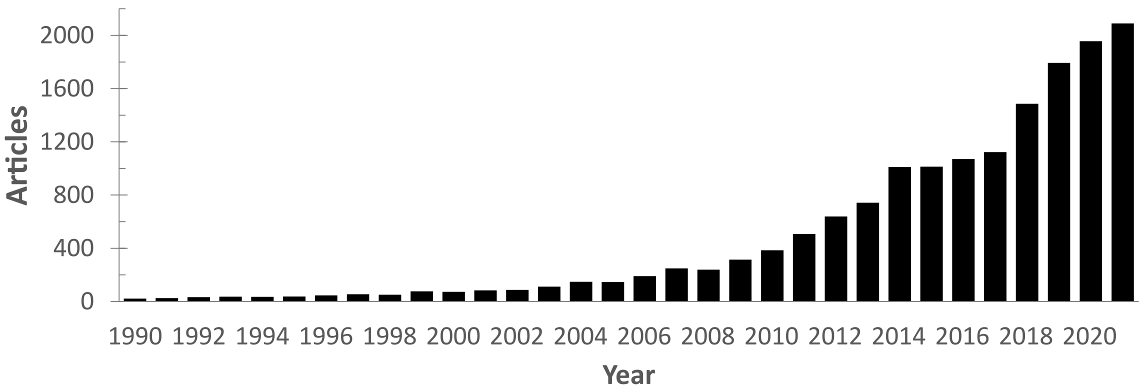
> *"Children and adolescents should limit the amount of time spent being sedentary, particularly the amount of recreational screen time."*
>
> Chaput et al., 2020, p. 7

These are especially important because, as already mentioned, inactivity causes 9% of premature mortality (Lee et al., 2012), but separately also too much sedentary activity leads to another 4% increase in the risk of premature mortality (Rezende et al., 2016). So, on the one hand, you can be active and sedentary at the same time if you perform the recommended amount of physical activity but also linger too long in sedentary behavior; on the other hand, too much inactivity and sedentary behavior add up to an even higher risk of premature mortality.

The literature on physical activity and health contains a multitude of different theoretical models that will only be touched upon here since the topic of the thesis only broadly addresses the effects on health. These models often predict and explain the occurrence of health or disease, as well as fitness or motor performance. The spectrum of models ranges from holistic models that organize the interaction of various constructs and influencing factors to complicated risk factor models for the prediction of various diseases. Against the background of the relevance of physical activity as a factor that influences health, it can be seen that in the decades around the millennium, physical activity was transformed from a rather insignificant factor in risk factor models (e.g. the Salutogenesis Model by Antonovsky (1997)) to a central construct in holistic health models (e.g. the Requirement-Resource Model by Woll (1996), the widespread Health Model by Bouchard et al. (2012) or the Model on the Influence of physical activity on fitness by Buford et al. (2014)).

In recent years, in addition to research on the influence of physical activity on health, a second focus, the study of sedentary time and sedentary behavior as an influencing factor, has grown rapidly (Giurgiu et al., 2020; Owen et al., 2011; Tremblay et al., 2010). In particular, in the last decade, the number of publications on this topic has increased considerably (cf. Figure 1.2). Sleep research has also long been strongly represented as an independent research area.

Figure 1.2: Publications by year with search term 'sedentary behavior', Scopus.com, accessed Nov. 04, 2022.



In simple words, physical activity and sedentary behavior, together with sleep as the third, represent the 24 hours of the day. Sleep represents a special component of sedentary behavior, which is usually characterized by a lying position and a reduction in bodily functions and a state of unconsciousness. Therefore, self-reporting of sleep is difficult. Previously, the effects of PA, SED, and sleep were studied in isolation by most researchers, which is flawed because the time spent on a behavior naturally depends on the composition of the rest of the day (Taylor et al., 2018). However, it is widely known that inadequate sleep is a risk factor for Cardiovascular Disease Risk (Covassin & Singh, 2016). Although still less established than the measurement of PA and SED, accelerometers are also now regularly used to assess sleep. Due to the aforementioned interrelationships between PA, SED, and sleep, these three together form what is known as **"Physical Behavior (PB)"** (Bussmann & van den Berg-Emons, 2013).
There are usually high correlations between behaviors, and an increase in a defined time frame spent on one of the behaviors (e.g. physical activity) leads to a decrease in daily time spent on at least one of the remaining behaviors (e.g. sedentary behavior and / or sleep). This area of "physical activity and health" is currently evolving further into a more comprehensive area of "physical behavior epidemiology." Although sleep is an increasingly important behavioral category, it is only mentioned in extracts, as the author of this thesis article has little practical experience in recording sleep and sleep was not part of the MoMo-Study.

Today, the physical behavior approach goes even a little deeper and focuses not only on the consequences of physical activity on health, but applies a holistic approach to the 24-hour sleep-wake cycle. Since each day is characterized by a combination of the three behaviors: sleep, sedentary behavior, and physical activity, the focus of PB is on mutual influence within this 24h-cycle (Bussmann & van den Berg-Emons, 2013;

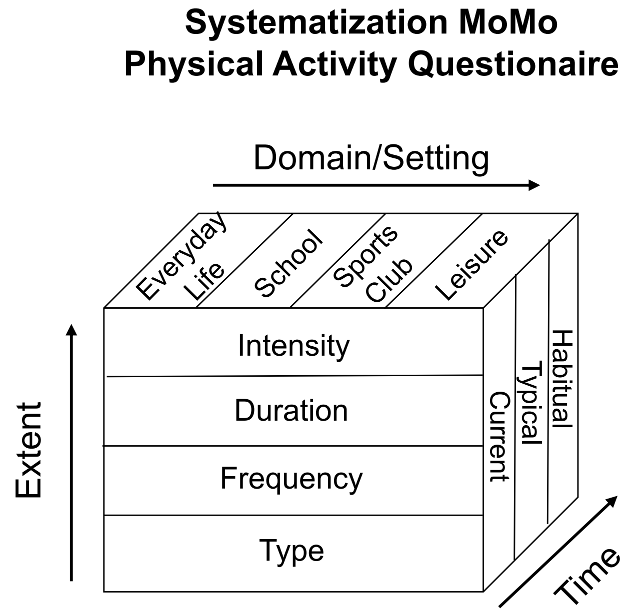
Rosenberger et al., 2019; Tremblay et al., 2017). Tremblay et al., 2017 was one of the first to illustrate this model (compare Figure 1.3) which was modified by Rosenberger et al. (2019) to explore the paradigm of the Interrelationships of all aspects of physical behavior on health effects. In the future, this model can provide the basis for the creation of holistic public health guidelines to combine synergies from different research areas and develop overall interventions. In this thesis, the term physical behavior refers collectively to the multidimensional construct of physical activity, sedentary behavior, and sleep.

Figure 1.3: 24-hour Physical Behavior Cycle (adapted from Tremblay et al., 2017).

Note: The movements that take place in the course of a day are divided into two components: On the outside, different categories are presented based on the posture of the body. Inside, divided by energy expenditure, are the three categories of sleep, physical activity, and sedentary behavior. The proportion of space occupied by each behavior in this figure is not prescriptive of the amount of time that should be spent on those behaviors each day.

The model of the Physical Behavior Cycle with its continuous observation of physical behavior leads to the main topic of this thesis, the device-based recording of physical activity. Until 2015, only the MoMo activity questionnaire was used to record physical activity within the MoMo-Study. This questionnaire is based on the systematization according to Woll (1996) shown in Figure 1.4. The questionnaire focuses on recording the central aspects of the type, intensity, frequency, and duration of the habitual or typical activity. To record these aspects of physical activity as holistically as possible, the MoMo activity questionnaire differentiates in various domains or settings.

Figure 1.4: The facets of physical activity, a modified and translated version of Woll, 1996 for surveying by activity questionnaire



In 2015 the recording of physical activity by accelerometers was added in MoMo, and this approach was extended to include the continuous capture of current activity intensity, which is rather difficult to capture with questionnaires. Some background information and considerations that arise in this context will be covered in more detail in the next two Sections.

1.2 Documentation of device-based measured PA

Hippocrates, as already mentioned, referred to the correct dosage of physical activity. To determine this dose, an exact analysis of the actual condition is required. In ancient Greece, neither stopwatches nor computers could be used for data acquisition. However, due to the increasingly accurate acquisition of data with device-based measured PA in modern times, the question of the correct dosage and the correct and most accurate recording is also now raised.

In this thesis, the term **"device-based measured PA"** is used in contrast to "objective measured PA", because there are so many subjective (or at least relatively arbitrary) decision points in the planning, recording, and analysis of accelerometry data that it is hardly objective. Although device-based PA can also describe other methods in which technical devices record physical activity, here the term mainly refers to the recording of physical activity using accelerometers. Acclerometry is composed of the Latin word "accelero", which means accelerate, and the Greek word "métron", which means measure. Therefore, it is the measure of acceleration. In sports science, accelerometry is now understood synonymously with the collection of physical behavior using **accelerometer (ACC)**. If you look at the term accelerometer in purely technical terms, it is a sensor that can detect accelerations. When PA monitors are based on accelerometers, they use one or more piezoelectric accelerometers, depending on the number of axes. A piezoelectric sensor consists of a piezoelectric element, for example, a bimetal, and a seismic mass. If the seismic mass of the sensor is now moved by motion, a deformation of the piezoelectric element takes place either in the form of a bend (in the case of beam sensors) or direct pressure or tension (in the case of newer sensors with an integrated chip (IC)). This change in electrical polarization results in a charge shift. This shift generates a voltage signal that will be proportional to the registered acceleration. A detailed description of the process can be found in Chen and Bassett (2005). Originally used in elevators and airplanes, accelerometers can now be found in a wide range of devices such as smartphones, activity monitors, pedometers, and game controllers. Since they are very good at detecting motion, there is no question about their usefulness in tracking physical behavior. Certain algorithms can be applied to accelerometer data to determine PA, SED, and sleep. Although the term accelerometer refers to the sensor, in this thesis it is also used synonymously for the entire device (in MoMo, especially for the models of the ActiGraph accelerometers).

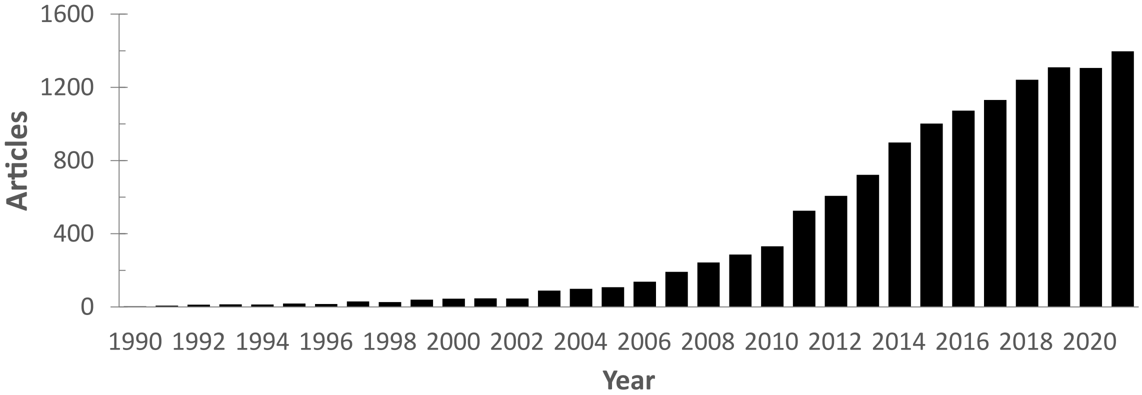
In the 40 years since Montoye's team of exercise physiologists and engineers modified a phono cartridge to measure bodily acceleration (Wong et al., 1981) and predict the energy expenditure of PA (Servais et al., 1984) the device-based measures of PA and related behaviors have emerged as an essential tool for research on PA and health promotion. In fact, the team's prescient observations that

> *"[...]the accelerometer's greatest value may be in the area of categorizing people into various activity-related groups. It could also be used as a device by which people could compare their daily activity to a prescribed level for rehabilitation, weight loss, or personal goals in training."*
>
> Servais et al., 1984

has become fully realized in recent years. PA monitoring devices continue to be the leading fitness trend (Thompson, 2018). More than 100 million tracking devices and accelerometer-enabled smartwatches were sold in 2017–18 (Lamkin, 2018), nearly 445 million devices in 2020 (IDC, 2022). Activity monitors have emerged as important self-monitoring tools in clinical medicine (Lobelo et al., 2018), large-scale health promotion (U.S. Department of Health and Human Services, 2018), and are now commonly used in national surveillance efforts (Burchartz et al., 2020; Colley et al., 2011; Matthews et al., 2008; Troiano et al., 2008). The number of yearly publications using search terms 'exercise or physical activity' and 'acceleromet*' increased from 10 or fewer until 1996 to more than 600 in 2012 and 2013 (Troiano et al., 2014). The same search term yields more than 1400 publications on scopus.com in 2021 (compare Figure 1.5 assessed on November 04, 2022).

ActiGraph accelerometers (Pensacola, FL, USA) were among the first accelerometers to be used early on for the study of PA and health. At the time, the first ActiGraph devices could only measure uniaxial acceleration. In 1998 (at that time Actigraph was just named "Manufacturing Technology Incorporated (MTI)" after being founded under the name "Computer Science Applications Inc. (CSA)"), the European Youth Heart Study (EYHS) was one of the first population-level studies to use accelerometers to quantify PA (Riddoch et al., 2004). ActiGraph released in 2009 the first triaxial GT3X

Figure 1.5: Publications by year with search terms 'exercise or physical activity' and 'accelerometer*,' Scopus.com, accessed Nov. 04, 2022.



model. In 2014, a review by Wijndaele et al. (2015) reported that more than 51% of the 76 studies, each with more than 400 participants, in 36 countries, used an ActiGraph accelerometer. Therefore, in KiGGS and MoMo wave 2, ActiGraph accelerometers (successor models: GT3X+ and wGT3X-BT) were used to enable comparison with other large-scale European studies (Riddoch et al., 2004; Ruiz et al., 2011).

This leads us to the actual core of this thesis. Device-based measurement of physical activity using accelerometers in the MoMo study.

1.3 Research Problem and relevance

A quote from "De Monarchia" from the famous Italian poet and philosopher Dante Alighieri (1265 - 1321), who after all is considered one of the most important philosophers of the European Middle Ages, gives an insight into the importance of each part of the whole:

"Et sicut se habet pars ad totum, sic ordo partialis ad totalem"
translated to
"And as a part stands in relation to the whole, so the order in a part stands to the order in the whole."

Alighieri, 1317

If this statement is transferred to the device-based recording of physical activity, we should not only consider the measured activity as a whole, but also take a close look at each step that is taken to record it.

Today accelerometers offer opportunities for researchers to capture valid data about the intensity and amount of physical behavior in real-time over several days and weeks. From these multidimensional data, a great number of metrics can be derived to capture and describe the unique aspects of physical behavior. As a researcher, you need to go through sometimes excessive technical details of accelerometry when selecting and applying devices to quantify physical behavior. The effects of these decisions on the metrics (energy expenditure, activity intensity, body position, activity patterns) can occur in a variety of ways. The device, carrying position (hip, wrist, thigh), and recording parameters (epoch length (EL), frequency, memory capacity, recording frequency, and filters) have a large influence on the measured activity (Rowlands et al., 2018). Different backgrounds such as study design (purpose, repeated measurements) and duration (time frame, wear time), as well as data storage and evaluation must be taken into account when determining the parameters. Finally, the evaluation must adjust several levers (raw data, context information, non-wear time, intensity classification, compliance) depending on the target variables.

As mentioned above, annual publications using accelerometers in physical behavior research have risen rapidly in recent years. The number of calibration studies and protocols used for this also increased, leading to a very heterogeneous field of outcome variables and results. Due to this large number of possible (subjective) decisions before the actual measurement, one should be thoroughly familiar with existing validation studies when planning a study and document as many technical decisions as possible when recording and evaluating data to enable data comparison across studies.

Therefore, the main objectives of this thesis are as follows:

(1) clarify technical decisions that should be made when using accelerometers in large epidemiological studies of activity,

(2) analyze the basic methodological aspects and important technical decisions in the device-based measurement of physical activity in large scale epidemiological studies, specifically the MoMo study,

(3) investigate the implications of different evaluation methods in the preprocessing of accelerometer data,

(4) compare the data based on the accelerometer recordings in addition to the representative data from the MoMo activity questionnaire,

(5) examine accelerometer data with respect to differences between weekdays and weekend days.

1.4 Structure of the thesis

For this reason, the document is structured as follows (see also Figure 1.6). The Introduction in Chapter 1 presents the background of activity and health research, as well as the development of accelerometry in this field in recent years. Starting with a short historical look at physical activity guidelines, it particularly focuses on Germany and the MoMo-Study. The introduction ends with an outline of the specifics of device-based activity measurement, emphasizing the particular importance of documenting technical decisions, and the scientific research already published. This is followed by the research problem and the main objectives of the thesis, which constitute the body of the document.

Chapter 2 provides the "Consensus Article" (A1 in Fig. 1.6) and reflects the expert consensus of the authors during the 2nd International Workshop for the Center for the Assessment of Physical Activity (CAPA) on the topic of accelerometry for physical behavior assessment. It is in the middle between methodology and data analysis (Figure 1.6), since it clarifies what aspects should be considered in planning and evaluation. This article helps to provide directions and useful references for practitioners and researchers when considering using accelerometers in epidemiological studies.

In Chapter 3, 'the Study Protocol' (A2 in Fig. 1.6), incorporates the problems identified in the Consensus Article and offers researchers the possibility to obtain a specific overview of the decisions for the methods and protocols used to assess device-based physical activity in children and adolescents with accelerometers in the MoMo-Study.

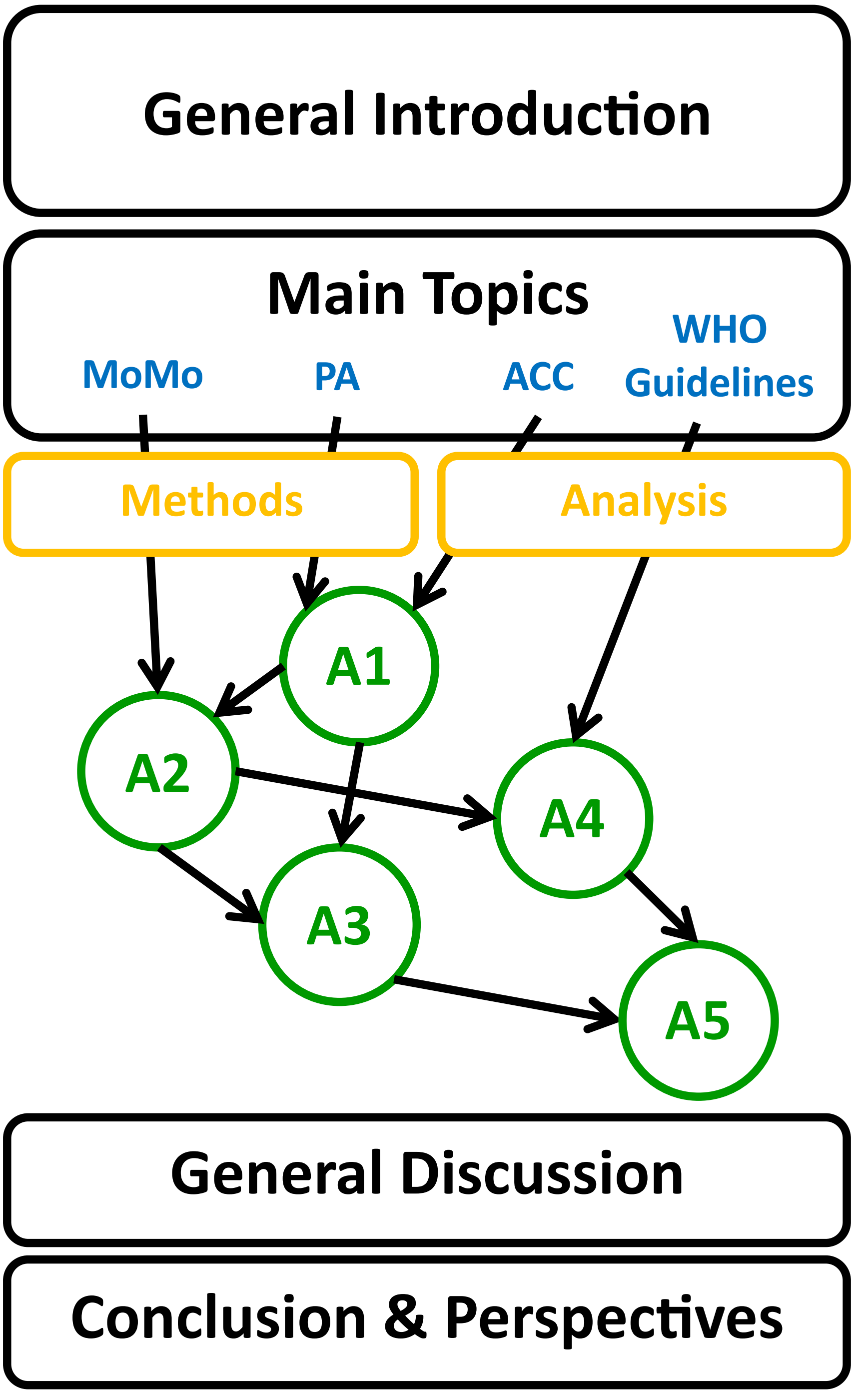
Chapter 4, the "the Methods-Article" (A3 in Fig. 1.6), is the next step after the assessment of physical activity with accelerometers. It explores the implications of choosing various metrics when processing recorded data. Through this examination, it becomes clear why it is so important for reproducibility that all technical decisions in the processing and evaluation of device-based measured physical activity are documented. In this article the different methods are primarily reviewed, but also analytically and critically analyzed, which is why this Chapter is again closer to the transition between methodology and analysis.

There was also strong agreement at the CAPA workshop that neither questionnaires nor accelerometers provide optimal coverage of all aspects of PA and that instead a multimodal, combined approach of self-report and device-based methods is recommended. Therefore, in Chapter 5, "the Comparison Article" (A4 in Fig. 1.6) makes a comparison between device-based measured physical activity and the MoMo physical activity questionnaire. This article clearly falls into the analysis category of the thesis, but also has a slight tendency to the underlying methodology. Based on the number of days on which participants reached the physical activity values given in the WHO guideline, this study examines the difference between self-reported and device-based measured PA and whether the differences in PA between the age and gender groups obtained by two methods are comparable.

Finally, Chapter 6 holds "the Typical Day Article" (A5 in Fig. 1.6) and takes a look on a average day of physical activity intensities measured by accelerometer and compares weekdays and weekend days. Compelling results on individual awake phases on different days of the week pose interesting considerations for future accelerometer studies. For this reason, this Chapter is also located in the outer margin of the analysis section.

This also introduces the transition to Chapter 7, "the Discussion", which contains the general discussion and conclusion of the thesis. It is intended to provide an outlook on what lessons can be learned from the first nationwide physical activity survey using accelerometers in Germany and what future research problems should be addressed next based on this experience. Appendix A contains additional documents and supplements that are referred to or have been created in the context of this work. Appendix B includes a list of all scientific publications by the applicant as author or co-author up to the date of submission.

Figure 1.6: Structure of the Thesis.



Note: The different articels are shown in the green circles, the blue keywords represent main topics, orientation of the articles to the yellow boxes indicates the affiliation whether the article is more methodological in character or has an evaluation focus.
A1: Consensus Article
A2: Study Protocol
A3: Methods-Article
A4: Comparison Article
A5: Typical Day Article

References Introduction

Ainsworth, B. E., Haskell, W. L., Leon, A. S., Jacobs, D. R., Montoye, H. J., Sallis, J. F., & Paffenbarger, R. S. (1993). Compendium of physical activities: classification of energy costs of human physical activities. *Medicine and science in sports and exercise*, *25*(1), 71–80. https://doi.org/10.1249/00005768-199301000-00011

Alighieri, D. (1317). *Monarchia: Edited by Pier Giorgio Ricci. Translated by Prue Shaw (1965)*. Edizione Nazionale. https://dante.princeton.edu/cgi-bin/dante/DispMinorWork.pl?TITLE=Mon&REF=I%20vi%201-4

American College of Sports Medicine. (1975). *Guidelines for graded exercise testing and exercise prescription*. Lea & Febiger.

Antonovsky, A. (1997). *Salutogenese: Zur Entmystifizierung der Gesundheit* (Vol. Band 36). dgvt Verlag.

Bouchard, C., Blair, S. N., & Haskell, W. L. (2012). *Physical Activity and Health*. Human Kinetics. https://books.google.de/books?id=tO96DwAAQBAJ

Buford, T. W., Anton, S. D., Clark, D. J., Higgins, T. J., & Cooke, M. B. (2014). Optimizing the benefits of exercise on physical function in older adults. *PM & R : the journal of injury, function, and rehabilitation*, *6*(6), 528–543. https://doi.org/10.1016/j.pmrj.2013.11.009

Burchartz, A., Manz, K., Anedda, B., Niessner, C., Oriwol, D., Schmidt, S. C. E., & Woll, A. (2020). Measurement of Physical Activity and Sedentary Behavior by Accelerometry Among a Nationwide Sample from the KiGGS and MoMo Study: Study Protocol. *JMIR Research Protocols*, *9*(7), e14370. https://doi.org/10.2196/14370

Bussmann, J. B. J., & van den Berg-Emons, R. J. G. (2013). To total amount of activity..... and beyond: perspectives on measuring physical behavior. *Frontiers in Psychology*, *4*, 463. https://doi.org/10.3389/fpsyg.2013.00463

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Reports*, *100*(2), 126–131.

Chaput, J.-P., Willumsen, J. F., Bull, F. C., Chou, R., Ekelund, U., Firth, J., Jago, R., Ortega, F. B., & Katzmarzyk, P. T. (2020). 2020 WHO guidelines on physical activity and sedentary behaviour for children and adolescents aged 5-17 years: summary of the evidence. *The international journal of behavioral nutrition and physical activity*, *17*(1), 141. https://doi.org/10.1186/s12966-020-01037-z

Chen, K. Y., & Bassett, D. R. (2005). The technology of accelerometry-based activity monitors: current and future. *Medicine and science in sports and exercise*, *37*(11 Suppl), S490–500. https://doi.org/10.1249/01.mss.0000185571.49104.82

Colley, R. C., Carson, V., Garriguet, D., Janssen, I., Roberts, K. C., & Tremblay, M. S. (2017). *Physical activity of Canadian children and youth, 2007 to 2015*. Statistics Canada.

Colley, R. C., Garriguet, D., Janssen, I., Craig, C. L., Clarke, J., & Tremblay, M. S. (2011). Physical activity of Canadian adults: accelerometer results from the 2007 to 2009 Canadian Health Measures Survey. *Health Reports*, *22*(1), 7–14.

Covassin, N., & Singh, P. (2016). Sleep Duration and Cardiovascular Disease Risk: Epidemiologic and Experimental Evidence. *Sleep Medicine Clinics*, *11*(1), 81–89. https://doi.org/10.1016/j.jsmc.2015.10.007

Dwyer, T., Coonan, W. E., Leitch, D. R., Hetzel, B. S., & Baghurst, R. A. (1983). An investigation of the effects of daily physical activity on the health of primary school students in South Australia. *International journal of epidemiology*, *12*(3), 308–313. https://doi.org/10.1093/ije/12.3.308

Giurgiu, M., Niermann, C., Ebner-Priemer, U. W., & Kanning, M. (2020). Accuracy of Sedentary Behavior-Triggered Ecological Momentary Assessment for Collecting Contextual Information: Development and Feasibility Study. *JMIR mHealth and uHealth*, *8*(9), e17852. https://doi.org/10.2196/17852

Graf, C., Beneke, R., Bloch, W., Bucksch, J., Dordel, S., Eiser, S., Ferrari, N., Koch, B., Krug, S., Lawrenz, W., Manz, K., Naul, R., Oberhoffer, R., Quilling, E., Schulz, H., Stemper, T., Stibbe, G., Tokarski, W., Völker, K., & Woll, A. (2013). Vorschläge zur Förderung der körperlichen Aktivität von Kindern und Jugendlichen in Deutschland. *Monatsschrift Kinderheilkunde*, *161*(5), 439–446. https://doi.org/10.1007/s00112-012-2863-6

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2018). Worldwide trends in insufficient physical activity from 2001 to 2016: A pooled analysis of 358 population-based surveys with 1·9 million participants. *The Lancet Global Health*, *6*(10), e1077–e1086. https://doi.org/10.1016/S2214-109X(18)30357-7

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1·6 million participants. *The Lancet Child & Adolescent Health*, *4*(1), 23–35. https://doi.org/10.1016/S2352-4642(19)30323-2

IDC. (2022). Absatz von Wearables weltweit in den Jahren 2014 bis 2021 (in Millionen Stück). Retrieved November 24, 2022, from https://de.statista.com/statistik/daten/studie/515723/umfrage/absatz-von-wearables-weltweit/

Kurth, B.-M., Kamtsiuris, P., Hölling, H., Schlaud, M., Dölle, R., Ellert, U., Kahl, H., Knopf, H., Lange, M., Mensink, G. B. M., Neuhauser, H., Rosario, A. S., Scheidt-Nave, C., Schenk, L., Schlack, R., Stolzenberg, H., Thamm, M., Thierfelder, W., & Wolf, U. (2008). The challenge of comprehensively mapping children's health in a nation-wide health survey: Design of the German KiGGS-Study. *BMC public health*, *8*, 196. https://doi.org/10.1186/1471-2458-8-196

Lamkin, P. (2018). Smartwatch popularity booms with fitness trackers on the slide. Retrieved September 30, 2021, from https://www.forbes.com/sites/paullamkin/2018/02/22/smartwatch-popularity-booms-with-fitness-trackers-on-the-slide

Lee, I.-M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., & Katzmarzyk, P. T. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *The Lancet*, *380*(9838), 219–229. https://doi.org/10.1016/S0140-6736(12)61031-9

Lobelo, F., Rohm Young, D., Sallis, R., Garber, M. D., Billinger, S. A., Duperly, J., Hutber, A., Pate, R. R., Thomas, R. J., Widlansky, M. E., McConnell, M. V., & Joy, E. A. (2018). Routine Assessment and Promotion of Physical Activity in Healthcare Settings: A Scientific Statement From the American Heart Association. *Circulation*, *137*(18), e495–e522. https://doi.org/10.1161/CIR.0000000000000559

Matthews, C. E., Chen, K. Y., Freedson, P. S., Buchowski, M. S., Beech, B. M., Pate, R. R., & Troiano, R. P. (2008). Amount of time spent in sedentary behaviors in the United States, 2003-2004. *American journal of epidemiology*, *167*(7), 875–881. https://doi.org/10.1093/aje/kwm390

Owen, N., Sugiyama, T., Eakin, E. E., Gardiner, P. A., Tremblay, M. S., & Sallis, J. F. (2011). Adults' sedentary behavior determinants and interventions. *American journal of preventive medicine*, *41*(2), 189–196. https://doi.org/10.1016/j.amepre.2011.05.013

Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C., Buchner, D. M., Ettinger, W., Heath, G. W., & King, A. C. (1995). Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *JAMA: The Journal of the American Medical Association*, *273*(5), 402–407. https://doi.org/10.1001/jama.273.5.402

Rezende, L. F. M., Sá, T. H., Mielke, G. I., Viscondi, J. Y. K., Rey-López, J. P., & Garcia, L. M. T. (2016). All-Cause Mortality Attributable to Sitting Time: Analysis of 54 Countries Worldwide. *American journal of preventive medicine*, *51*(2), 253–263. https://doi.org/10.1016/j.amepre.2016.01.022

Riddoch, C. J., Bo Andersen, L., Wedderkopp, N., Harro, M., Klasson-Heggebø, L., Sardinha, L. B., Cooper, A. R., & Ekelund, U. (2004). Physical activity levels and patterns of 9- and 15-yr-old European children. *Medicine and science in sports and exercise*, *36*(1), 86–92. https://doi.org/10.1249/01.MSS.0000106174.43932.92

Rosenberger, M. E., Fulton, J. E., Buman, M. P., Troiano, R. P., Grandner, M. A., Buchner, D. M., & Haskell, W. L. (2019). The 24-Hour Activity Cycle: A New Paradigm for Physical Activity. *Medicine and science in sports and exercise*, *51*(3), 454–464. https://doi.org/10.1249/MSS.0000000000001811

Rowlands, A. V., Mirkes, E. M., Yates, T., Clemes, S., Davies, M. J., Khunti, K., & Edwardson, C. L. (2018). Accelerometer-assessed Physical Activity in Epidemiology: Are Monitors Equivalent? *Medicine and science in sports and exercise*, *50*(2), 257–265. https://doi.org/10.1249/mss.0000000000001435

Ruiz, J. R., Ortega, F. B., Martínez-Gómez, D., Labayen, I., Moreno Aznar, L. A., de Bourdeaudhuij, I., Manios, Y., González-Gross, M., Mauro, B., Molnar, D., Widhalm, K., Marcos, A., Béghin, L., Castillo, M. J., & Sjöström, M. (2011). Objectively measured physical activity and sedentary time in European adolescents: The HELENA study. *American journal of epidemiology*, *174*(2), 173–184. https://doi.org/10.1093/aje/kwr068

Rütten, A., & Pfeifer, K. (Eds.). (2017). *Nationale Empfehlungen für Bewegung und Bewegungsförderung* (Aufl. 1.2.06.17, Vol. Sonderheft 3). Bundeszentrale für gesundheitliche Aufklärung (BZgA).

Scheidt-Nave, C., Kamtsiuris, P., Gößwald, A., Hölling, H., Lange, M., Busch, M. A., Dahm, S., Dölle, R., Ellert, U., Fuchs, J., Hapke, U., Heidemann, C., Knopf, H., Laussmann, D., Mensink, G. B. M., Neuhauser, H., Richter, A., Sass, A.-C., Rosario, A. S., . . . Kurth, B.-M. (2012). German health interview and examination survey for adults (DEGS) - design, objectives and implementation of the first data collection wave. *BMC public health*, *12*, 730. https://doi.org/10.1186/1471-2458-12-730

Servais, S. B., Webster, J. G., & Montoye, H. J. (1984). Estimating human energy expenditure using an accelerometer device. *Journal of Clinical Engineering*, *9*(2), 159–170.

Taylor, R. W., Haszard, J. J., Meredith-Jones, K. A., Galland, B. C., Heath, A.-L. M., Lawrence, J., Gray, A. R., Sayers, R., Hanna, M., & Taylor, B. J. (2018). 24-h movement behaviors from infancy to preschool: cross-sectional and longitudinal relationships with body composition and bone health. *The international journal of behavioral nutrition and physical activity*, *15*(1), 118. https://doi.org/10.1186/s12966-018-0753-6

Thompson, W. R. (2018). Worldwide survey of fitness trends for 2019. *ACSM's Health & Fitness Journal*, *22*(6), 10–17.

Tremblay, M. S., Aubert, S., Barnes, J. D., Saunders, T. J., Carson, V., Latimer-Cheung, A. E., Chastin, S. F. M., Altenburg, T. M., & Chinapaw, M. J. M. (2017). Sedentary Behavior Research Network (SBRN) - Terminology Consensus Project process and outcome. *The international journal of behavioral nutrition and physical activity*, *14*(1), 75. https://doi.org/10.1186/s12966-017-0525-8

Tremblay, M. S., Colley, R. C., Saunders, T. J., Healy, G. N., & Owen, N. (2010). Physiological and health implications of a sedentary lifestyle. *Applied physiology, nutrition, and metabolism = Physiologie appliquee, nutrition et metabolisme*, *35*(6), 725–740. https://doi.org/10.1139/H10-079

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine and science in sports and exercise*, *40*(1), 181–188. https://doi.org/10.1249/mss.0b013e31815a51b3

Troiano, R. P., McClain, J. J., Brychta, R. J., & Chen, K. Y. (2014). Evolution of accelerometer methods for physical activity research. *British journal of sports medicine*, *48*(13), 1019–1023. https://doi.org/10.1136/bjsports-2014-093546

U.S. Department of Health and Human Services. (2008). 2008 Physical Activity Guidelines for Americans. https://health.gov/our-work/nutrition-physical-activity/physical-activity-guidelines/previous-guidelines/2008-physical-activity-guidelines

U.S. Department of Health and Human Services. (2018). Physical Activity Guidelines for Americans, 2nd edition. https://health.gov/paguidelines/second-edition/pdf/Physical_Activity_Guidelines_2nd_edition.pdf

Wijndaele, K., Westgate, K., Stephens, S. K., Blair, S. N., Bull, F. C., Chastin, S. F. M., Dunstan, D. W., Ekelund, U., Esliger, D. W., Freedson, P. S., Granat, M. H., Matthews, C. E., Owen, N., Rowlands, A. V., Sherar, L. B., Tremblay, M. S., Troiano, R. P., Brage, S., & Healy, G. N. (2015). Utilization and Harmonization of Adult Accelerometry Data: Review and Expert Consensus. *Medicine and science in sports and exercise*, *47*(10), 2129–2139. https://doi.org/10.1249/MSS.0000000000000661.

Woll, A. (1996). *Gesundheitsförderung in der Gemeinde : eine empirische Untersuchung zum Zusammenhang von sportlicher Aktivität, Fitneß und Gesundheit bei Personen im mittleren und späteren Erwachsenenalter* (Dissertation). Johann Wolfgang Goethe Universität Frankfurt am Main. Neu-Isenburg, Institut für Sportwissenschaften.

Woll, A., Albrecht, C., & Worth, A. (2017). Motorik-Module (MoMo) – the KiGGS wave 2 module to survey motor performance and physical activity. *Journal of Health Monitoring*, *2*(S3), 66–73. https://doi.org/10.17886/RKI-GBE-2017-110

Wong, T. C., Webster, J. G., Montoye, H. J., & Washburn, R. (1981). Portable accelerometer device for measuring human energy expenditure. *IEEE Transactions on Biomedical Engineering*, (6), 467–471.

World Health Organization. (2010). *Global recommendations on physical activity for health*.

2. State of the science, best practices and future directions

Burchartz, Alexander; Anedda, Bastian; Auerswald, Tina; Giurgiu, Marco; Hill, Holger; Ketelhut, Sascha; Kolb, Simon; Mall, Christoph; Manz, Kristin; Nigg, Claudio R.; Reichert, Markus; Sprengeler, Ole; Wunsch, Kathrin; Matthews, Charles E. (2020): Assessing physical behavior through accelerometry – State of the science, best practices and future directions.

Psychology of Sport and Exercise 49, S. 101703.
DOI: 10.1016/j.psychsport.2020.101703.

2.1 Introduction

Accelerometers offer opportunities for researchers to capture valid data on the intensity and amount of physical behavior (PB) in real time over a period of several days and weeks. From these multidimensional data, a great number of metrics can be derived to capture and describe the unique aspects of PB. The goal of this article is to help the end user of PB monitoring devices (from novice to intermediate experience) wade through the sometimes excessive technical details of accelerometry to outline best practices in selecting and applying devices to quantify three main behavioral categories of common interest to the research community: physical activity (PA), sedentary behavior (SED), and sleep. The effects of these decisions on the metrics (energy expenditure, activity intensity, body position, activity patterns) can occur in a variety of ways. The device, the carrying position (hip, wrist, thigh), and the recording parameters (epoch length (EL), frequency, memory capacity, recording frequency, and filters) have a large influence on the measured activity. Different backgrounds such as study design (purpose, repeated measurements) and duration (time frame, wear time), as well as data storage and evaluation must be taken into account when determining the parameters. Finally, the evaluation must adjust several levers (raw data, context information, non-wear time, intensity classification, compliance) depending on the target variables. Looking into the future, current developments in statistical analysis are discussed, because the research community has not yet reached a consensus on the most promising approach. There are exciting developments ahead of us in the future. Sleep in particular is increasingly being considered an influencing factor for health. Together with the technical developments in sensors that will become incrementally smaller, more accurate, and in the near future will be integrated directly into our clothes or skin, accelerometry is facing exciting times and lots of data to evaluate.

2.2 State of the science

In the 40 years since Montoye's team of exercise physiologists and engineers modified a phono cartridge to measure bodily acceleration (Wong et al., 1981) and predict the energy expenditure of physical activity (PA) (Servais et al., 1984), device-based measures of PA and related behaviors have emerged as an essential tool for research on PA and health promotion. In fact, the team's prescient observations that "the accelerometer's greatest value" may be in the area of categorizing people into various activity-related groups. It could also be used as a device by which people could compare their daily activity with a prescribed level for rehabilitation, weight loss, or personal goals in training" (Servais et al., 1984) has become fully realized in recent years. PA monitoring devices continue to be a leading fitness trend (Thompson, 2018) and more than 100 million tracking devices and accelerometer-enabled smart watches were sold in 2017–18 (Lamkin, 2018). Activity monitors have become important self-monitoring tools in clinical medicine (Lobelo et al., 2018), large-scale health promotion (U.S. Department of Health and Human Services, 2018), and are now commonly used in national surveillance efforts (Burchartz et al., 2020; Colley et al., 2011; Matthews et al., 2008; Troiano et al., 2008). The number of yearly publications by search terms 'exercise or physical activity' and 'acceleromet *' increased from 10 or fewer until 1996 to more than 600 in 2012 and 2013 (Troiano et al., 2014). The same search term yields more than 1200 publications on scopus.com in 2019 (assessed on November 7, 2019).
One of the advantages of accelerometry is that it can collect dense data over a long period of time (days, weeks, and sometimes even months), allowing for a detailed examination of daily behavior. From these multidimensional data, a large number of metrics can be derived to capture and describe the unique aspects of PB. Recent epidemiological studies provide new information on the distinct influence of sedentary behavior (SED), light (LPA), and moderate to vigorous intensity physical activity (MVPA) on health enabled by accelerometry (Diaz et al., 2017; Matthews et al., 2016), and new cohort studies are using accelerometers on a much larger scale and include the assessment of sleep (Doherty et al., 2017; German National Cohort Consortium, 2014). This accelerometer era in PB research might be considered a golden age with great opportunities, but many challenges remain.
Although there are suitable and precise devices available for most applications (e.g., interventions, epidemiology, surveillance), the wide variety of devices and prediction algorithms available for a variety of metrics (e.g., step counts, energy expenditure, intensity classification, posture, sleeping pattern) and limited information from rigorous validation studies make it difficult for the average user to understand what the most appropriate options are for each application. One device may not fit all applications, and users must make informed choices to optimize the results of individual studies. The goal of

this article is to help the end user of PB monitoring devices (novice to intermediate experience) navigate through sometimes excessive technical details of accelerometry to outline best practices and highlight necessary considerations in selecting and applying devices to quantify two major behavioral categories of common interest to the research community: PA and SED. Although sleep is an increasingly important behavioral category, it is only mentioned in extracts, as the authors of the article have little practical experience in recording sleep. Throughout this article, no major distinction is made between accelerometers sold for research purposes (research devices) and those sold to the general population (consumer-devices), it is expected that the decision-making processes apply to each. In many cases, consumer devices are as accurate as research devices and they often provide feedback to participants more easily (Henriksen et al., 2018; Wahl et al., 2017). Nonetheless, consumer-devices rarely provide raw data and mostly use proprietary algorithms, which make the harmonization of various datasets difficult. In addition, a reanalysis of the data with different calculation methods is almost impossible with proprietary algorithms.

Figure 2.1: Logo of the CAPA Workshop held on the 11th and 12th July, 2019 at the Karlsruhe Institute of Technology, Germany

The opinions outlined in this article reflect the accumulated experience of the authors with a wide variety of monitoring devices. The group of authors consists of interdisciplinary researchers in the field of accelerometry and PB that came together at the 2nd International Workshop for the Center for the Assessment of Physical Activity (CAPA) held on the 11th and 12th July, 2019 at the Karlsruhe Institute of Technology, Germany (Fig.2.1). This is not a systematic review, but rather reflects the expert consensus of the authors on the topic of accelerometry for PB assessment.
The following section focuses on best practices and is divided into subsections of:

- Behaviors, highlighting the differences in assessing the different facets of the PB spectrum,
- Metrics, providing insight on what parameters to use for different goals,
- Study Design, showing differences in approaches for a variety of research questions,
- Data collection, processing and (storage/accessibility), describing what still needs to be done after the study has been designed,

- And current developments in statistical analysis, highlighting the need for sophisticated methods to analyze data.

The last section will focus on future directions of the field. Some of these directions are already developing; others will definitely be important topics in the future.
Although there is literature that provides help for users of accelerometer- based measures in surveillance (Pedišić & Bauman, 2015), which gives considerations regarding data collection and processing (Migueles et al., 2017) or focuses on methods in intervention studies (Montoye et al., 2018), the need for directions for new researchers in this field remains. The authors provide these directions and related useful references to help practitioners and researchers when considering using accelerometers.

2.3 Best practices

Currently, accelerometry is the "state of the art" when it comes to device-based measurement of PB. In addition to PA, SED, and sleep are the most common behaviors assessed using triaxial movement acceleration methods within 24-hour measurements.

2.3.1 Study Design

2.3.1.1 Physical activity

"Physical activity is defined as any bodily movement produced by skeletal muscles that results in energy expenditure" (Caspersen et al., 1985, p. 126). When Caspersen and colleagues defined PA in 1985, accelerometers were not yet feasible and widely available for PA assessment. Technological advances have created an increasing interest in the use of accelerometry for the assessment of PA today (Ward et al., 2005; Welk, 2002) using different outcomes (see Section 2.3.2 Metrics for more information). The main advantage of using accelerometry over self-reported (retrospective) PA is that it is not prone to recognition, memory, or social desirability biases of participants, as recall measures are (Adams et al., 2005; Brenner & DeLamater, 2014; Nigg et al., 2012). A detailed discussion of the questionnaire approach can be found in the Physical Activity Questionnaire Article (Nigg & Woll, 2020). A review of the comparison of indirect measures (that is, questionnaires) and direct device-based measures of PA in pediatric populations revealed that 72% of indirect measures overestimated directly measured values (Adamo et al., 2009). However, this trend was not confirmed for adults (Prince et al., 2008), nor for older adult populations (Kowalski et al., 2012). Here, no clear patterns emerged for the mean differences between subjective self-report and device-based measures of PB. Moderating factors of overestimation are less clear and should be examined in future studies. In general, the correlations between device-based and self-report measures are weak throughout the lifespan, showing a discrepancy between both measures. However, the need for device-based monitoring of PA, ideally 24 hours a day, for assessment of real-life activity patterns is often stated.

2.3.1.2 Sedentary behavior

Besides PA, it is also important to objectively measure SED (Lewis et al., 2017), especially since studies have shown that SED has negative effects on health outcomes (Katzmarzyk et al., 2019). According to the internationally accepted definition of SED (Tremblay et al., 2017), it is necessary to capture both characterizations of SED, namely:

body posture (sitting/lying/reclining) and energy expenditure ($\leq$1.5 MET). Moreover, there is an ongoing debate on whether the effects of SED on mental and somatic health are independent of PA or not (Biswas et al., 2015; Ekelund et al., 2016).
Thus, there is a growing need to assess both behaviors over a specified period of time, providing the opportunity to examine the definite effects of both on mental and somatic health. Measurement of SED is also feasible using self-report and device-based measures.

2.3.1.3 Sleep

Sleep forms a special component of SED, which is usually characterized by a lying position and a reduction in bodily functions. Therefore, self-report measures of sleep are difficult, as a sleeping state induces a state of unconsciousness. The effects of PA, SED, and sleep are examined by most researches in isolation (Chaput et al., 2017), what is flawed because time spent in one behavior will naturally depend on the composition of the rest of the day (Taylor et al., 2018). Even if less established than PA and SED measurements, accelerometers can also be used to assess sleep, which is the third component of every day PB. Different sleep quality patterns, such as bed and out-of-bed time (and therefore hours per day spent in bed), nighttime movements, or even clinical sleep problems, can be assessed. But this 24-hour movement recording requires different methods of data analysis than standard multivariate techniques. These methods need to take into account the codependence and proportional nature of compositional data (Chastin et al., 2015; see Section 2.3.5 Current developments in statistical analysis for more information). A study comparing sleep / wake judgments obtained through sleep diary and accelerometer driven data revealed good agreement for nocturnal sleep (Kawada, 2008). Similar findings were obtained from a recent review on the validity and reliability of sleep time questionnaires across the lifespan with moderate to strong correlations between measurements (Nascimento-Ferreira et al., 2016) as well as in a study combining diary and questionnaires with accelerometer data (van Hees et al., 2015). When validated against polysomnography, the state-of-the-art method for detecting sleep, high correlations can be detected, revealing that accelerometry is a good tool for real-life assessment (Jean-Louis et al., 2001).

2.3.2 Metrics

Accelerometry is based on continuous and real-time measurement and recording of movement-induced raw acceleration signals over a specific period of time (EL). Accelerometers register the intensity and duration of single- or multiaxial accelerations and convert these raw data into manufacturer- and model-specific outcome metrics. Before the raw data is converted, it is usually filtered. This should eliminate acceleration frequencies that are not compatible with human movement. Most devices allow the user to choose between different filters when processing the data. As the different filters have a large impact on the outputs, it is important to report information on the filters used. Unfortunately, there are no internationally accepted standards for signal processing (Rowlands, 2007). Due to these differences in raw data processing and filtering, outcome metrics cannot be directly compared across devices (Chen & Bassett, 2005). However, using identical software and algorithms to process outcome measures can help harmonize the data. A recent study by Rowlands (2018) has shown that key PA outcomes derived from data from different devices, which were processed identically thereafter, were largely equivalent.

Since the outcome metrics provided by accelerometers do not have direct physical meaning, they must be translated into a more interpretable unit or measure (Troiano, 2006). For example, different age-specific cut-points can be applied to accelerometer outcome metrics in order to express raw data in terms of time spent in specific behaviors (PA, SED or sleep), in different positions and postures (standing, sitting, lying) and to determine intensities of PA (low, moderate, vigorous; for example, Schaefer et al., 2014). When applying specific cut-off points from calibration studies, it is crucial to follow the same data collection protocol (processing, epoch length, device placement, filter, etc.) that was used in the original study (Migueles et al., 2017). Although common metrics of interest are described in detail in this section, there are new approaches that focus on improving these existing metrics. The finding of metrics for PA volume and intensity derived from raw acceleration has been the focus of recent studies and show great promise (Rowlands, Edwardson, et al., 2018; Rowlands, Mirkes, et al., 2018; Rowlands, Sherar, et al., 2019).

2.3.2.1 Energy expenditure

One key use of accelerometry is to interpret the raw data recorded by the device to estimate energy expenditure (EE). This is one way to make accelerometer data comparable across different types of PA and across a broad range of target groups. Therefore, accelerometer metrics are often transferred into the commonly used metabolic equivalent of a task (MET). The MET values obtained can then be classified as sedentary ($\leq$ 1.5 METs), low (1.5-3 METs), moderate (3 to 6 METs) and vigorous ($>$ 6 METs) PA (Butte et al., 2012)) for adults. However, children are known to have considerably higher basal metabolic rates per unit of body mass than adults. Therefore, the adult MET thresholds do not apply to children (McMurray et al., 2015; Saint-Maurice et al., 2016). For scoring and interpretation of youth PA data, the new Youth Compendium of Physical Activities can be applied (Butte et al., 2012). When interpreting accelerometer data by means of EE one has to consider that multiple different algorithms exist to convert accelerometer outcome metrics into EE outcomes. This may lead to different EE outcomes depending on the algorithm used, so data from different accelerometer models cannot be directly compared if the conversion algorithm is not open to the public. Furthermore, the relationship between a specific activity and EE can vary due to external circumstances (additional loading, changes in altitude, and temperature). Furthermore, depending on the placement, accelerometers are not capable of accurately detecting the intensity of activities that involve the use of the upper extremities and activities with limited hip movement (Prince et al., 2008).
The inclusion of heart rate monitors may have an additive informative value, especially in activities that involve isometric muscular contraction, such as weight bearing exercises and activities such as carrying a load, pushing, and rowing (Jakicic et al., 2004; Kozey-Keadle et al., 2010). A combination of accelerometer analysis and heart rate monitoring may further improve the overall accuracy of EE and the assessment of exercise intensity in free-living situations.

2.3.2.2 Activity intensity

The volume and intensity of PA and SED during a specific time interval (hours, days, and weeks) may be obtained by classifying outcome metrics accumulated in a specific EL (integration of a filtered digitized acceleration signal over a user-specified time interval, see Section 2.3.4 Data collection and processing for more information) with a set of

cut-points. These cut-points function as thresholds for outcome metrics and are used to classify activities as sedentary, light, moderate, vigorous or very vigorous activities (Migueles et al., 2017).
These cut-points are generally validated specifically for a certain model of accelerometer, wear location, age group, and health status of the observed population (Taraldsen et al., 2012). Thereafter, applying cut-points to a specific data set requires following the same data collection and processing criteria that were applied in the original calibration study.

2.3.2.3 Body position and posture

Some accelerometers have the ability to distinguish between different body postures and positions using the accelerometer inclination output. Using both inclination and dynamic acceleration, these devices are able to classify basic posture by distinguishing sedentary activities from upright activities. The ability to identify different body positions and postures is dependent on the placement of the accelerometer on the human body (e.g. hip, wrist, thigh, etc.). However, Carr and Mahar (2012) reported that the correct body position was identified only two-thirds of the time during sedentary activity. This may be attributed to the fact that accelerometry analyzes PB in predefined ELs (Ayabe et al., 2013). Especially small activities such as transitions from a sitting or supine position only last a few seconds and thus may be below the EL chosen and are therefore not resolved. New methods to address this using angle for posture estimation show promise (Vähä-Ypyä et al., 2018).
It is also challenging to monitor the movement behavior of certain study populations, such as toddlers, children, and people with movement disabilities due to the occurrence of 'non-standard' postures, such as kneeling and crawling (Davies et al., 2012). These postures are "non-standard" because the devices categorize body position as sitting / lying, standing, or step. The best way to quantify non-standard postures and the transitions between postures is direct observation or proxy reporting. Additionally, the time spent in a posture does not indicate the effort someone must exert to attain or maintain this posture or movement.

2.3.2.4 Activity patterns

To date, most research has focused on associations between the amount of time spent in SED, LPA, and MVPA and health. However, recent evidence from controlled experimental trials suggests that the pattern of PB may also be related to health outcomes, even when accounting for the total volume of activity (Keadle, Shiroma, et al., 2017). In addition to PB variables such as duration, intensity, and volume, the activity pattern has been suggested as an outcome of PB that may provide additional information beyond reports of activity counts or other outcome metrics (Cavanaugh et al., 2010). Short PA breaks between prolonged periods of SED are especially important to reduce the adverse effects of somatic and mental health effects (Ekelund et al., 2016; Giurgiu et al., 2019). Therefore, determining these intermissions is of relevance, especially in organizational settings such as work and school. In the current literature, the term bout is often used to describe a predefined amount of time without intermission in the same PB. However, with the new emerging possibilities of data analysis, the term 'activity pattern' refers to more than just the summation of bouts.
Another very interesting measurable characteristic is the timing of different PB. Olds et al. (2011) got results associated with late bedtimes and late wake-up times with an unfavorable activity profile and a weight status profile. Matricciani et al. (2019) used GENEActive wrist-worn accelerometers to check the duration, onset, offset, daily variability

and efficiency of sleep. Recent studies used an accelerometer to assess associations between sleep duration, timing, and regularity with measures of adiposity (Zhou et al., 2018) or PA (Xu et al., 2019). The development of population-independent outcome metrics that adequately assess the prevalence of meeting the PB guidelines is another new approach to quantify accelerometer outcomes. Metrics such as the 'magnitude of acceleration' above which a person accumulates the most active 60 (for children) or 30 (adult) minutes are possibly comparable between data sets and a new tool for public health to report on guidelines (Rowlands, Fairclough, et al., 2019).

2.3.2.5 Evaluating the accuracy and precision of outcome metrics

Somewhat surprisingly, it can be difficult to determine the accuracy and precision of the monitoring devices in the setting it is typically used in population-based samples of people living independently of their daily lives, at home, work / school or at leisure. If one is interested in assessing the total volume of PA, estimated as total activity counts or a sum of all bodily acceleration during the monitoring period, then less validation work may be needed. On the other hand, if you want to know the true validity of your estimates of step counts, sedentary time, EE or the duration of activity intensities or sleep, more rigorous validation studies are needed. For example, EE validation can be performed by validating against oxygen intake determined by spiroergometry.
It has become clear that the initial assumption that simple calibration studies using laboratory-based approaches alone would be sufficient to develop accurate prediction algorithms for PB in real life settings was not always tenable. Prediction methods developed using a small set of activities in controlled environments have not always produced valid estimates of target behaviors in real life because they cannot cover the entire spectrum of PB that occurs. Additionally, high-quality validation studies (described in more detail below) are actually relatively rare, especially with the use of free-living data. Many new prediction methods have been created, but few have been rigorously tested. Limited information about which of these prediction methods (e.g. moderate-vigorous intensity cut-points, Lee et al., 2019) are most accurate has led to much confusion in the field. To help clarify this area and identify studies that may provide better validity estimates, Keadle et al. (2019) have recently proposed a framework that outlines specific steps in the monitoring development, calibration, and validation process. Drawing on frameworks used for drug development, they proposed four phases of monitoring development, validation, and application. The initial step (Phase 0) is based on bench tests and refinement of the technical reliability of the monitor. The next steps reflect the calibration of the monitor or the development of the prediction algorithms. Phase I testing includes simpler and more controlled laboratory-based testing of selected activities using fixed start/stop times and the development of the initial prediction method(s). Phase II testing extends the earlier phase and includes implementation of semi-free-living protocols, including transitions between activities, to further develop and refine prediction methods. Critterion measures, such as direct observation and indirect calorimetry, are integral to the calibration monitoring process, and these data are often used to provide initial validity information about new devices or prediction methods. However, since the data used in these "validation" studies are derived from the same study population from which the prediction methods were developed, they can overestimate the actual validity of the method when it is applied in a new study population. Phase III of the development process involves rigorous validation studies of previously developed prediction methods using strong criterion measures (i.e., indirect calorimetry, direct observation, doubly labelled

water) in a study sample different from that used to develop prediction method (i.e., independent sample). Examples of strong Phase III validation studies include that of Toth et al. (2018) who evaluated the validity of step counts from a variety of devices compared to direct video observation; Lyden et al. (2014) who evaluated MVPA estimates using a new machine learning algorithm compared to direct observation; Chomistek et al. (2017) who compared a variety of ActiGraph prediction methods for energy expenditure with doubly labeled water; and Crouter et al. (2013) who compared estimates of MVPA duration using various ActiGraph prediction methods to indirect calorimetry. Implementation of strong criterion measures in independent samples of free-living individuals can provide clear evidence for most accurate precise methods, while reliance on Phase I/II can be useful but leaves much more uncertainty. Unfortunately, Phase III validation studies are relatively rare in most cases and only Phase I or II studies are available to evaluate a method of interest.

The final phase (Phase IV) of the development process involves the application and dissemination of methods that have successfully progressed through previous phases. As data processing and prediction algorithms become more complex, to minimize the requirements of specialized expertise, the development of more user-friendly methods will be required for more effective dissemination and use by the research community.

2.3.2.6 Limitations/perspectives

The accelerometry data should be directly comparable and understandable in physical terms and valid in a wide range of target groups (Taraldsen et al., 2012). Unfortunately, little data is available to provide evidence to determine the most valid variables for different purposes.

A common problem with all acceleration monitors is processing and classifying the data into PB outcome variables. Specifically, different methods for handling the same data can result in significantly different values for the same outcome variables. Thus, although PB measurement with accelerometers may be considered 'objective', there are subjective elements such as setting the EL and cut-points, and consensus guidelines for collecting and processing these device-based measured data are lacking (Heil et al., 2012). Additionally, contextual information related to the setting and type of activity in accelerometry is limited. Therefore, information from other sources (that is, behavior logs) and sensors should be integrated to better understand PB in different contexts and to increase the comparability of measurements by imputing non-wear time (Sprengeler et al., 2017). However, this will increase the participant burden and the evaluation effort of the researchers.

2.3.3 Study Design

2.3.3.1 Choosing a monitor and monitor placement

Choosing the 'best' monitor for a given research, clinical, or intervention application depends on the characteristics of the PB that one wants to measure; the type of study or intervention project at hand; the amount of burden that participants might accept; and the available staff and resources to administer the monitors. These resources include cost of devices, logistics of monitor administration, data storage, and analytic resources available, and increasingly specialized expertise to implement more advanced prediction methods/algorithms.

The first question to ask when deciding on which monitor to use in a given setting and where to locate it in the body is: What aspects of human behavior and metrics of those behaviors should be measured? Different monitoring devices have different strengths and weaknesses to predict different summary metrics. For example, devices worn on the thigh generally provide more accurate and precise measures of body posture than waist or wrist-worn accelerometers, although devices worn at each of these sites may output a summary metric for sedentary time. Additionally, there can be important differences in the accuracy and precision of a given metric, even when they are derived from the same type of monitor. For example, more advanced pattern recognition methods are likely to be more accurate and precise than simple cut-points derived from the same device (Lyden et al., 2014). It is recommended to think about identifying the most valid algorithm to predict the specific metrics of interest, and then select the monitor and monitor placement that can capture data to feed into that algorithm. In other words, identify the summary metrics of primary importance for the study, select the monitoring device, and placement that have adequate validity for the specific study population and study type. These choices also need to be within the resources available for the project and within an acceptable range of burden for the participants.

2.3.3.2 Measurement time frame

The time frame for data collection depends mainly on the research question or the broad measurement objective. In planning a study, researchers must decide how long monitors should be worn each day (e.g., a few hours, waking day, 24 hours), how many days should the data collection period include, and whether seasonal variation in behavior may affect the results (Atkin et al., 2016; Matthews et al., 2001). Choosing a device or prediction method that has higher validity/accuracy should minimize systematic errors in the estimation of PB metrics of primary interest. However, human behavior is inherently variable because humans are not robots that do exactly the same thing every day. This results in a natural variation from day to day in behavior, and our patterns of behavior often change with the seasons and from year to year. Thus, natural variation in behavior must be taken into account when designing measurement protocols and matching the protocol to the broad objectives of the study. Study designs using shorter measurement time frames with the goal of estimating mean values within a study population (versus individual prediction), such as population surveillance and intervention studies, may be more susceptible to the influence of seasonal variation, and the measurement schedule may need to be designed to minimize these effects. Short data collection periods during the day could be useful if the aim of the research is to measure PB in a specific setting or situation such as during physical education in schools. Wake-up time, or time out of bed, is often used for studies of SED and PA, while a 24-hour protocol is also needed if sleeping behavior is part of the research question. Furthermore, wearing the device for 24 h increases the recording time and the compliance with the wear time, therefore, a recording time of 24 h per day is recommended when possible (Migueles et al., 2017; Tudor-Locke et al., 2015). However, from an ethical perspective, it should be carefully reconsidered if a 24-hour measurement period is justified if only daytime data is of interest.

In terms of the number of days of monitoring needed, this choice may depend on the objectives of the study. If the goal is to estimate mean values in a population for surveillance purposes, in theory, only a single day of observation is needed. Migueles et al. (2017) recommend a minimum of four days of valid data (wear time of at least 8–10 h per day), while Trost et al. (2005) claim that PB patterns can be determined with 3–4 days of measurement with over 80% reliability. The general recommendation to capture seven consecutive days of data collection is generally a feasible approach to assess habitual PB patterns in children and adults (Addy et al., 2014; Barreira et al., 2015; Trost et al., 2000). Sampling 7-day periods increases the chance of capturing an adequate number of valid days, which means a compromise between sample size and reliability (Migueles et al., 2017), for meaningful data analysis, and improves the opportunity that weekday and weekend days are part of the data collection period (Addy et al., 2014; Trost et al., 2005). Interestingly, Wolff-Hughes et al. (2016) noted that purposeful sampling of weekend days can lead to biased estimates of population mean values compared to random sampling, raising questions about the common practice of requiring fixed numbers/types of days in ones analyzes. There may be no right or wrong approach and for some purposes including participants with a single day of observation is appropriate, while studies with other goals may need to ensure more days and specific types of day should be included in the analysis. The research community has carefully examined the number of days or observations required to achieve adequate reliability for short-term measures (e.g., ICCs $>$ 0.8). However, less work has been done to understand the variation in behavior from one administration period (e.g., one 7-day period) to the next. In general, studies that examined this type of behavior variation have observed relatively high reliability from one 7-day administration to the next in older women from the United States (Keadle, Sampson, et al., 2017) and middle-aged and older adults in Germany (Jaeschke et al., 2018), indicating that 7-day administration periods reflect relatively long-term average values for PA and SED in the population. Interestingly, behavioral variation in measures, conceptualized as random fluctuations around long-term average behavior (i.e. random measurement error), has a differential impact on statistical results depending on whether the PB variable is used as a dependent or independent variable in models (Hutcheon et al., 2010). When PB variables are used as dependent variables, random error results in no bias in the model-based estimates, but does reduce their precision (i.e., standard errors increase). On the contrary, when PB variables are used as independent variables, random error can introduce bias into model-based association estimates (that is, attenuated beta coefficients) a phenomenon called regression dilution bias (Elliott et al., 1990). This effect should be considered when interpreting analysis including an independent PB variable.

2.3.3.3 Studies with repeated measurements

Study designs, such as interventional or longitudinal studies, that require repeated measurements need some additional considerations in the planning stages to allow the best comparability of the device-based measured PB data collected over time. The monitor administration methods, including wearing position and periods, as well as device settings (e.g., sampling frequency), should remain as consistent as possible within each administration period. Additionally, the wear time determinations and metric prediction methods applied to the raw data should be consistent over time. Furthermore, external factors that could influence the PB, but that are not part of the research question, should be standardized by consistent and purposeful sampling over time. For example, sampling PB with attention to season of the year and/or the days of the week monitored is important to minimize variation due to these factors. Documentation of all procedures and decisions made is recommended, e.g. in the form of a standard operation procedure (SOP).

2.3.3.4 Measurement reactivity

Wearing an accelerometer might influence the behavior of participants, which is known as measurement reactivity (Baumann et al., 2018). The studies showed that the participants modified their PA pattern in the presence of an accelerometer. For example, in an adult sample, a small portion of LPA was reallocated to SED. In another study, Clemes et al. (2008) compared the step counts of the participants who knew they were being tracked with those of the participants who did not know and found a significant increase in the first condition. Devices that display the archived activity or sedentary time might even enhance the effect and motivate participants to change their behavior. However, if the typical activity behavior should be measured, a modification of the normal activity pattern is not desired and could bias the results. Therefore, it is recommended to provide as much information as necessary on the outcome measure (e.g. explaining 'measuring activity' rather than 'measuring steps' or 'measuring movement at the hip') and use devices that do not display results to reduce the reactivity bias of the measurement. Furthermore, longer measurement periods, as well as the exclusion of the first measurement day from data analysis, can counteract bias due to measurement reactivity (Dössegger et al., 2014).

2.3.4 Data collection and processing

It is critical to document all technical decisions for comparison purposes.

2.3.4.1 Data collection

Three decisions must be made before collecting accelerometer data. First, it is a basic requirement to measure and store the raw data of the accelerometer. Unlike previous decades that required the accumulation of data in formats such as counts, the storage capacity of accelerometer devices is not limited, as emphasized thirteen years ago (Rowlands & Eston, 2007). The storage of raw data enables researchers to process the data offline with different algorithms (analyzing body position with static components, time course of the magnitude of activity, time-frequency analyzes) or to reanalyze the data with algorithms newly developed in future. The measurement range (e.g. $\pm$ 8g) and resolution (e.g. 0.01 g, respectively, 0.1 m/s2 per bit) depend on the characteristics of the Micro-Electro-Mechanical Systems (MEMS) sensor selected by the manufacturer of the accelerometer and usually covers the full range of naturally occurring human acceleration values. The sampling frequency (defining the temporal resolution) must be at least twice as high (Nyquist frequency or Nyquist-Shannon sampling theorem Shannon, 1949) as the highest movement frequency component to avoid aliasing effects (this means generating virtual frequency components when the analog accelerometer signal is digitized). Human movement frequencies can reach values of 10 Hz in writing (Teulings & Maarse, 1984) or 10.5 Hz in piano playing (Furuya & Soechting, 2012). Taking into account these two examples, a sampling rate of 21 Hz (2×10.5 Hz) is required to assess the movement frequency. For the assessment of PB, a sampling frequency of 30 Hz normaly meets the Nyquist criterion. If the magnitude of the activity or the movement pattern (e.g. for biomechanical analyses) is in the focus of interest, the sampling rate must be multiple times higher, whereas the accuracy of the measured pattern increases with sampling rate (accuracy – samplerate trade-off; e.g., a sampling rate of five or ten times higher than the movement frequency results in a more or less amount of distortion of the measured signal). Therefore, it depends on the research question how precisely a

movement pattern should be measured, respectively, which grade of distortion is acceptable to increase measuring time (depending on the storage capacity of the device). Kang and Rowe (2015) developed a method for automated and task-specific optimization of sampling rates.
Secondly, as mentioned above, it must be decided how many days or weeks to measure. This decision depends mainly on the study question (for a detailed discussion, see 2.3.3.2 Measurement Timeframe).
The third issue is the epoch setting, i.e., the aggregation level / period length used to analyze the raw acceleration. Aggregation into epochs is necessary to feed raw data into various prediction methods, whether they are equations or simple cut-points. Particularly in children, whose movement behavior pattern is known to be spontaneous and intermittent, an EL between 1 and 5 s or the shortest possible EL is recommended (Banda et al., 2016; Bornstein et al., 2011; Heil et al., 2012; Kettner et al., 2013; Migueles et al., 2017; Sigmund et al., 2014). On the contrary, long ELs such as 60 s are known to underestimate MVPA in children (Guinhouya et al., 2013). Here it must be noted that with longer ELs, short vigorous PA intervals are more often detected as moderate PA, and short intervals of moderate PA are at the same time detected as LPA. The accelerometer studies chose to use an EL of 60 s due to the smaller storage capacities in the early years. With better storage capacities today, smaller ELs are no longer a problem and are highly recommended. The optimal EL in terms of health outcomes is yet unknown.

2.3.4.2 Context information

Using diaries or questionnaires assessing the PB context is essential for specific research questions. For example, the simultaneous application of self-reports and device-based measurements enables one to reliably quantify the volume/intensity of PB within a specific time frame, such as work or school hours, and the timing of organized sports (Helmerhorst et al., 2012; Reilly et al., 2008; Sprengeler et al., 2017; Westerterp, 2009). If schedules are available, this information can be collected at the group level in a less time-consuming manner, e.g. school classes (Sprengeler et al., 2019). If they are not available, the use of wear time protocols or log-books needs to be considered (Burchartz et al., 2020). Combining self-reports and device-based measures can allow for a comprehensive understanding of PB (Sallis & Saelens, 2000). The accelerometer can only capture accelerations to create an activity profile of the frequency, intensity, and duration of the activity when worn, so detailed information about the type, quality, and context of the activity is needed for a complete understanding of the assessed PB.

2.3.4.3 Non-wear time

Since most accelerometers are not waterproof (with a few exceptions, the latest sensors from GENEActiv and Axivity), essential amounts of aquatic PA will not be measured. In this case or when the participants forget to wear the accelerometer, the non-wear time must be identified (via self-report or algorithms). In general, algorithms for detecting non-wear times consist of time intervals with successive values of the lowest valid acceleration with or without a tolerance of several minutes in which small accelerations are allowed, with optional windows of zero acceleration before and after this tolerance. The algorithm of Barouni et al. (2020) also analyzed the respiration pattern to differentiate the wear time from the non-wear time. Toftager et al. (2013) recommend that the different algorithms be used only for the specific subgroups in which they have been validated.

There are different algorithms for different age groups and more studies are needed to test the precision of each algorithm for these different age groups (Migueles et al., 2017). The non-wear time data need to be excluded from the data analysis, otherwise they will likely be categorized as SED. This introduces a bias as non-wear time is normally associated with some kind of activity with an intensity higher than SED. As mentioned above, recording 24 h movement requires different data analysis methods since during sleep movement is even lower than during SED (Chastin et al., 2015). When devices are not worn during sports activities, the imput of MVPA for the duration of sports activities should be considered with caution. It is highly unlikely that the entire duration will be spent doing MVPA. It is not recommended to enter the total time spent during sports activities as MVPA as stated in a diary. An imputation of 50% of the reported duration is a good solution if the average PA value is to be analyzed from an epidemiological point of view, particularly in children (Fairclough & Stratton, 2006; Hollis et al., 2016; Sprengeler et al., 2019).

2.3.4.4 Compliance

Trained study staff should hand out the devices and motivate participants to wear the device at all times. It is best that the participants are introduced to how to wear the device correctly. The important points of wearing the accelerometer (placement, wearing times, data protection, and return of the device) should be summarized in an information sheet to be distributed to participants. In particular, concerns can be dispelled by pointing out that an activity sensor (accelerometer) is present and used to detect PB (as opposed to camera, GPS, WLAN, Bluetooth, or similar). In some countries, it is possible to obtain an ethics statement to collect GPS data – even for children, which provide rich data sets on PB.
In some sports, the organizer / organization prohibits the use of electronic devices to record and evaluate activities in real time. Additional information sheets with further information about the study for teachers, trainers, coaches, parents, or employers, depending on the setting in which the study takes place, should also be provided. This information may prevent devices from needing to be removed at sporting events or during work/school, which would result in data loss. When returning the device by mail, an addressed and stamped envelope should be provided to participants to make it as easy as possible for them to send it back. A telephone follow-up protocol can be implemented if the devices are not returned (Burchartz et al., 2020). The already recommended 24-hour recording time can also increase compliance. On the one hand, it provides a complete picture of activity during waking hours, sleep patterns, and inactive phases. On the other hand, the compliance is better as participants do not have to pay attention to taking off the device at certain times (e.g. sleep).

2.3.5 Current developments in statistical analysis

2.3.5.1 Acknowledgement of the intrinsic co-dependency between human behaviors during a finite amount of time

Measurement of human behavior such as PA, SED, and sleep with accelerometer devices results in data sets representing observations of a finite amount of time, e.g., 8 h, 12-h, 24-h, or 7 days. Within the observed time frame, the amount of an individual's behavior summed up to 100%, and an increase in one behavior ultimately leads to a decrease in at least one of the other behaviors. These circumstances have important implications for the interpretation of the results of the statistics used and the interpretation of the results in relation to health benefits. Standard procedures of multivariate statistical analyzes (e.g., linear regression, ANOVA) assume that the time spent on one behavior is independent of the time spent on any other behavior and that the amounts are potentially infinite. However, accelerometry always deals with finite amounts of time, and applying, for example, a standard linear regression technique may in some cases lead to multicollinearity issues. Over decades, research on human behavior in relation to health has most often not taken advantage of the interrelationships between time-linked behaviors by examining the specific behavior of interest in isolation, rather than investigating the dependencies that do exist (Mekary et al., 2009). Recently, several research groups have worked on different strategies to recognize the codependency between behaviors and apply more sophisticated statistical approaches, such as the Isotemporal Substitution Model (ISM), Compositional Data Analysis (CoDA), Multivariate Pattern Analysis or Bayesian Dirichlet distributions (Aadland et al., 2018; Chastin et al., 2015; Mekary & Ding, 2019; van der Merwe, 2018). In October 2019, an international workshop was held in Granada organized by Francisco Ortega with the focus on statistical methods to analyze accelerometer-measured PA. As a result of this workshop, there is a detailed consensus article on analytical approaches to evaluate associations with accelerometer-determined PB in epidemiological studies (Migueles et al., 2021). This article provides an expert description and a discussion of the statistical approaches currently available for use in epidemiological studies, as well as highlighting their strengths and limitations.

2.3.5.2 Isotemporal substitution model/compositional isotemporal substitution model

The ISM was originally developed by Mekary et al. (2009) to analyze PA and SED data. The general aim of the ISM is to estimate the effect of replacing one specific behavior with another behavior for the exact same amount of time. With a setup of different models, the association with a specific health outcome variable (e.g., BMI, biomarkers) for a predicted replacement of one behavior with another can be estimated. The ISM approach is able to control for all relevant behavior- related variables and the observed total time. It is important to note that, in most cases, absent multiple measures of behavior over time, the association only describes the predicted influence of changing behavior rather than the effect of an actual change.

2.3.5.3 Compositional Data Analysis

The CoDA approach was developed by Chastin et al. (2015) to analyze PA and SED data. CoDA accounts for the constraint structure of the data, namely the finite amount of, time of e.g. 24-h. Such data can only exist in a specific constraint geometric space, the so-called simplex, according to the Aitchison geometry (Aitchison, 1982). However, standard statistical techniques apply Euclidean geometry and assume potentially infinite data. To analyze data of compositional nature, the data first has to be expressed as isometric log ratios (ilr). The compositions of the ilr coordinates can then be analyzed in Euclidean geometry with standard statistical techniques. CoDA can be applied with the aim of using the composition of human behaviors as an independent variable or to predict the change in specific health-related outcome variables (e.g., BMI, biomarkers) in relation to behavior change. Dumuid, Stanford, Pedišić, et al. (2018) have recently developed and tested a compositional ISM in an empirical study.

2.3.5.4 Multivariate Pattern Analysis

Multivariate pattern analysis is widely used in pharmaceutical and metabolomics studies and has recently been adapted by Aadland et al. (2018) to analyze accelerometry-derived PA and SED data in relation to cardiometabolic health. Multivariate pattern analysis accounts for the multicollinearity of PA and SED data and provides a solution to analyze a more detailed spectrum of PA intensities in comparison to the established but rather broad overall summary measures (e.g., SED, LPA and MVPA). Furthermore, Aadland, Andersen, et al. (2019) showed that the explained variance of metabolic health was ten times when applying the full spectrum of PA intensities from the three axes compared to the traditional summary of MVPA derived by counts per minute from the vertical axis alone.

2.3.5.5 Bayesian Dirichlet distribution

For several research questions, it may be of interest that a composition of human behaviors such as PA and SED is the dependent variable. A Bayesian approach using a Dirichlet distribution is suggested (van der Merwe, 2018), which accounts for the compositional nature and allows one or more independent variables in a regression model. As an empirical example, how player positions in team sports affect the amounts of standing, walking, and running, measured via accelerometry, during the game, which can be of great importance in sport science and practical coaching situations.

2.3.5.6 Practical suggestion

The research community has not yet reached a consensus on the most promising approach, and both the original ISM and the compositional ISM have tended to show broadly similar results (Dumuid, Stanford, Martín-Fernández, et al., 2018). Mekary and Ding (2019) argue in their comment article on a compositional ISM study by Biddle et al. (2018) that their original ISM has several advantages compared to compositional ISM, such as a straightforward interpretation and intuitive articulation of the results with respect to PA guidelines. The original ISM uses absolute values of the time spent in a specific behavior instead of relative values that are used in the compositional ISM by Biddle et al. (2018) and Dumuid, Stanford, Pedišić, et al. (2018) and Fairclough et al. (2017). Aadland, Kvalheim, et al. (2019a) compared multiple linear regression, CoDA, and multivariate pattern analysis in an empirical study. Substantial differences in associations between PA intensities and cardiometabolic health were identified, and the authors argue that multivariate pattern analysis should be considered in future studies. Each approach has its strengths, limitations, and practical relevance. Therefore, researchers must carefully inspect the approach that fits best with their research objective and data. Helpful guidelines for the analytical process are available: Mekary et al. (2009) provide detailed information regarding the applied substitutional model and other related models in the appendix; Chastin et al. (2015) provide detailed information regarding the compositional nature of sleep, SED, and PA, and Dumuid, Stanford, Martín-Fernández, et al. (2018) provide a description and sample R code in the Supplementary Material accompanied to the respective article. Aadland, Kvalheim, et al. (2019b) provide a tutorial that guides the reader on how to conduct multivariate pattern analysis with respect to PA and health. van der Merwe (2018) also provides a description, sample R code, and an example data set to adapt his Bayesian approach to handle compositional data in regression modelling. It is important to note that the application of these statistical approaches is rather complex in comparison with standard procedures of multivariate statistical analyzes (e.g. linear regression, ANOVA). Therefore, researchers are encouraged to seek the advice of appropriate experts.

2.4 Future directions

2.4.1 Algorithms for intensity detection and validation studies

Traditionally, the intensity of PA obtained from accelerometer data is determined by the accumulated number (the outcome metric, sometimes referred to as 'counts') of threshold exceedances (so-called cut-off points) per time unit (EL). The raw signals of the accelerations measured by the accelerometer are processed and evaluated by various methods. Various validation studies for different age groups and devices determine cut-off points for SED, light, moderate, and vigorous intensities (Migueles et al., 2017; Schaefer et al., 2014). This has proven to be successful, as the intensities for different age groups and different target groups differ greatly from each other. However, because of the large number of devices available, the number of intensity algorithms is also very high. So far, there are no uniform international standards that specify how validation studies should be conducted for the different target groups, so that the results calculated afterwards would be comparable. Similarly, an unanswered question using ELs is whether and how measuring in smaller epochs affects estimates of PA minutes and meeting PA guidelines.
However, we recommend not using device-based outcome metrics anymore because it is not always known or reported how they are computed. Many commercially available devices keep these methods for count calculation proprietary. Newer device manufacturers (e.g. activpal, movisens, GENEActiv), mostly with a scientific background, have therefore increasingly opted for an open-science approach, making comparability between devices possible. Future intensity calculation algorithms should be based on raw data from accelerometers and be open to the public. This open science approach would allow studies to compare their data with raw data from other studies by applying the different algorithms to their own raw data set.
This is why pattern recognition algorithms are emerging in the field right now. Due to the amount of raw data available for evaluation, different algorithms, for example, from the field of speech recognition, were tested in the last years to be used to find PB patterns in accelerometer data. Different studies and researchers like Farrahi et al. (2019) and Smith et al. (2019) or the group around Stewart Trost (Ahmadi et al., 2019; Tong et al., 2019) are currently working on the use of supervised learning algorithms, as well as deep learning and convolutional neural networks, to predict activity patterns and EE from body-worn accelerometers.

2.4.2 Sleep detection

As sleep is related to many health outcomes, future studies should also use the opportunity of accelerometers to assess different sleep outcomes, such as duration of sleep, sleeping habits, including sleep movements, health issues related to sleep such as apnea or naps during the day. Differentiation of sleep and SED can be difficult at times (e.g., the difference between lying on the couch watching television and sleeping on the couch in front of the TV). An important point to consider is the positioning of the accelerometer. Positioning the sensor on the thigh makes it possible to detect different body positions very well. Smith et al. (2019) validated different body placements of the accelerometer and found that the hip may be superior for sleep timing and quantity metrics, whereas the wrist may be superior for sleep quality metrics. In the future, more studies like the one by Barouni et al. (2020) are needed to differentiate between wake, sleep and non-wear periods. Studies that confirm the best placement for accelerometers during PA, SED and sleep like Leppänen et al. (2019) are needed as well.

2.4.3 Timing of PA

2.4.3.1 Sleep/active cycles

Moreover, the inclusion of sleep pattern assessment can provide deeper insights into sleep / wake patterns of participants. After achieving more accurate algorithms for sleep and daytime nap detection (potentially by including other sensors, or measuring heartrate variability), gaining deeper insights into inter-individual daily routines and making it possible to more precisely determine the relationship between active and sedentary versus sleep times should be the focus of future research.

2.4.3.2 Activity throughout the day

Moreover, accelerometer use makes it possible to examine different activities in terms of duration and intensity over the whole day (if the device is worn throughout the day). Using these data, researchers can make specific statements about when, for how long, and at what intensity a person is active. Here, daily patterns can be examined, that is, if someone is walking to work every day at the same time or if someone has any activity routines. This also allows investigation of the within-day transfer or compensation effects of PA or SED.

2.4.3.3 Non-wear time protocols via ambulatory assessment

The manual input of non-wear time protocols into PA data is very time-consuming and distorts the collected device-based data set with subjective evaluations of activity. A more sustainable approach is to use ambulatory assessment in combination with 24-hour recording.
Triggered e-diaries can ask subjects about activity right after certain events have been detected (no device worn, periods of high activity, or SED). This means that non-wear times and especially the reasons for that can also be recorded relative precisely. This allows at least the recording of PA while the device was not worn, and the participant can receive feedback on how much activity was not recorded, as well as the activity type and context. Due to continuous technical progress and the constantly decreasing size of the sensors, simultaneous use of several sensors will be conceivable in the future. This could combine the advantages of different carrying positions and also improve the detection of certain PB (Reichert et al., 2020).

2.4.3.4 Smart patches and clothes

To encounter compliance problems and the difficulties in comparing data intraindividually due to different wear times, body-mounted sensors, smart patches, or smart clothes are promising attempts. This technology also has the potential to be used in automated activity profiling systems that produce a continuous record of activity patterns over extended periods of time (Preece et al., 2009). However, a 24-hour assessment faces multiple challenges: First, the adherence of the participants must be ensured. People are more likely to wear a monitoring device if it does not interfere with their daily habits and activities (Evenson et al., 2015). A hip-worn device, for example, is not feasible for tight clothes; a wrist-worn device is not suitable for a craftsperson wearing gloves, and either device is not suitable for a swimmer if it is not waterproof. A promising approach to face these problems is the establishment of more (validated) user-friendly equipment, like (waterproof) smart patches (Schneller et al., 2017). A microchip collecting triaxial acceleration data (and possibly more) mounted in a small patch that can easily be adhered at any body location. The first commercial products (e.g., biobeat patch or Moio.care smart patch) are already on the market. At present, these products are still associated with high costs for individual smart patches. In addition, these devices can often be used once or a limited number of times and must then be disposed of, which produces a lot of waste and is not as sustainable as reusable devices, such as typical accelerometers. Reusable devices that attach to the body with bio patches are preferable to one-use devices. Due to technological development, sensors and device sizes are getting smaller and are therefore easier to use. A promising approach into this direction has already been made by leading sports manufacturers by inserting such microchips into sportswear (e.g., shirts like the Hexokin Smart Shirt, or shoes like the Digitsole Smartshoe or the Skiin smart underwear) which can be used more than once. But even smart clothes still have a very limited lifespan, moisture during washing and the strain on cables and sensors caused by movement limit the length of time they can be used. A potential future direction is the development of subdermal accelerometer computer chip implants. The potential in terms of data collection is great; however, this raises major ethical issues as subdermal microchips cannot easily be removed. This would limit human rights with respect to privacy and make them 'transparent humans' with no chance of escaping permanent (possible) observation through data gained within these chips.

2.4.4 Combining accelerometry with other PA assessment methods

To obtain more comprehensive pictures of PA, SED, and sleep behavior, it is recommended to integrate accelerometry with ecological momentary assessment (EMA). EMA is a reliable instrument for gaining Big Data, allowing one to make differentiated assumptions about people's daily lives due to high-resolution data points which can be obtained by:

a) collect data from large numbers of subjects using, for example, mobile phones and by

b) assessing a large number of different measures from subjects (e.g., GPS, heart rate, heart rate variability, electrodermal activity, context, etc.; Hesse et al., 2015; Hidalgo-Mazzei et al., 2016).

Both approaches can help to understand real life settings in different ways and are valuable approaches to various science topics (Berger et al., 2017). A detailed discussion of this approach can be found in the Ambulatory Assessment Article (Reichert et al., 2020).

2.5 Conclusion

Accelerometry is the state of the art when it comes to device-based measurement of PB. The advantage of accelerometry is that it can collect dense data over a long period of time, allowing a detailed examination of daily behavior. From these multidimensional data, a great number of metrics can be derived to capture and describe the unique aspects of PB.

Besides PA, SED and sleep are the most commonly evaluated behaviors. Various carrying positions and sensors are available for different areas of application. The complex and dense data resulting from device-based measured PB as well as the various options with respect to devices, data collection, and data analysis, can also be a challenge for researchers. Furthermore, the different approaches used in studies can lead to limited comparability and reproducibility of the study results. The numerous considerations mentioned lead to the conclusion that:

- ≫ A recording time of 24-h per day is recommended for at least seven days (Migueles et al., 2017; Tudor-Locke et al., 2015);
- ≫ Consider existing validation studies when planning one's own studies and to document as many technical decisions as possible when recording and evaluating data to allow data comparison across studies;
- ≫ There is a critical need for better validation studies (Phase III studies in the Keadle et al. (2019) framework). They are needed to clarify questions about the accuracy of various prediction methods;
- ≫ Determine format and sampling frequency of the acceleration data (a sampling frequency of 30 Hz normally meets the Nyquist criterion) the recording time (a 24 h recording of at least one complete week using the shortest possible EL (1s) is recommended) which can be converted to longer ELs if needed;
- ≫ In addition to the device-based measurement, assess the type of activity performed during non-wear time and reasons for non-wear of the devices for a complete understanding of the PB assessed;
- ≫ There are algorithms that can determine non-wearing periods, but EMA methods are endorsed to capture contextual information and activities during non-wearing periods as mentioned in Reichert et al. (2020).

Device-based measured PA is often assumed accurate and to reflect actual PB. We propose that researchers choose their method based on the most valid approach for their given behavioral metric. From that perspective, the method chosen will dictate the type of device and the prediction algorithm. However, the values from, e.g., accelerometers are still estimates, and in the absence of satisfactory agreement with ground-truth gold standard measures of free-living PA should not be interpreted as 'actual' PA levels. The research community has not yet reached a consensus on the most promising approach in statistical analyses of device-based measured data, besides the inherent multicollinearity within data based on human behavior during a finite amount of time should be carefully considered. Each approach has its strengths, limitations, and practical relevance. Therefore, researchers must carefully inspect the approach that fits best with their research

objective and data.
In exercise psychology, accelerometry is a valuable tool for experimental and correlational research, as well as for developing individualized training programs. In addition to gaining information such as duration, length, and height and triaxial acceleration within a passed training session, accelerometer data are also used in post-match analyses in team sports. For (self-) observational purposes, accelerometer-based information can be helpful in two main areas focused by exercise psychology: in motivation of people to engage in any kind of exercise, or athletes; and in barrier management. Here, a combination of accelerometry with other features (like EMA, diaries) is expected to be the means of choice. In addition, individual feedback methods can be used to improve the enjoyment of exercise or training.
The best practices section of this article also provides valuable information for exercise psychologists and points to further literature to reach a fundamental understanding of the use of accelerometers in exercise psychology. It should be used as a starting point for exercise psychologists considering the use of accelerometers. The section of future directions shows opportunities for further research, and especially ambulatory assessment shows great promise in the field of exercise psychology. The goal of this report is to help the end user of PB monitoring devices wade through sometimes excessive technical details of accelerometry to outline best practices in selecting and applying devices to quantify three major behavioral categories of common interest to the research community: PA, SED, and sleep. There are still many challenges, but we also have exciting developments, e.g., technical developments in sensors which will be even smaller and more accurate ahead of us in the future.

Authors' contributions

Alexander Burchartz was responsible for the overall conception and design of this manuscript. Charles Matthews, Bastian Anedda & Alexander Burchartz were responsible for the State of the Science section. The Best Practice Chapter was divided and written by the following: Behaviors by Kathrin Wunsch & Marco Giurgiu; Metrics by Sascha Ketelhut & Alexander Burchartz the process of writing the manuscript as well as the statistical analysis and interpretation of data; Study Desing by Charles Matthews, Kristin Manz & Tina Auerswald; Data Collection and Processing by Ole Sprengeler, Simon Kolb, Holger Hill & Alexander Burchartz; Current developments in statistical analysis by Christoph Mall & Alexander Burchartz. The Future Directions Chapter was written by Alexander Burchartz, Kathrin Wunsch & Claudio Nigg. The Conclusion was written by Alexander Burchartz & Claudio Nigg. Charles Matthews, Markus Reichert, Kathrin Wunsch, Claudio Nigg & Alexander Burchartz were responsible for the critical revision. All authors read and approved the final manuscript.

References Consensus Article

Aadland, E., Andersen, L. B., Anderssen, S. A., & Resaland, G. K. (2018). A comparison of 10 accelerometer non-wear time criteria and logbooks in children. *BMC public health*, *18*(1), 323. https://doi.org/10.1186/s12889-018-5212-4

Aadland, E., Andersen, L. B., Resaland, G. K., & Kvalheim, O. M. (2019). Interpretation of Multivariate Association Patterns between Multicollinear Physical Activity Accelerometry Data and Cardiometabolic Health in Children-A Tutorial. *Metabolites*, *9*(7). https://doi.org/10.3390/metabo9070129.

Aadland, E., Kvalheim, O. M., Anderssen, S. A., Resaland, G. K., & Andersen, L. B. (2019a). Multicollinear physical activity accelerometry data and associations to cardiometabolic health: challenges, pitfalls, and potential solutions. *The international journal of behavioral nutrition and physical activity*, *16*(1), 74. https://doi.org/10.1186/s12966-019-0836-z

Aadland, E., Kvalheim, O. M., Anderssen, S. A., Resaland, G. K., & Andersen, L. B. (2019b). The Triaxial Physical Activity Signature Associated with Metabolic Health in Children. *Medicine and science in sports and exercise*, *51*(10), 2173–2179. https://doi.org/10.1249/mss.0000000000002021.

Adamo, K. B., Prince, S. A., Tricco, A. C., Connor-Gorber, S., & Tremblay, M. S. (2009). A comparison of indirect versus direct measures for assessing physical activity in the pediatric population: A systematic review. *International journal of pediatric obesity : IJPO : an official journal of the International Association for the Study of Obesity*, *4*(1), 2–27. https://doi.org/10.1080/17477160802315010

Adams, S. A., Matthews, C. E., Ebbeling, C. B., Moore, C. G., Cunningham, J. E., Fulton, J., & Hebert, J. R. (2005). The effect of social desirability and social approval on self-reports of physical activity. *American journal of epidemiology*, *161*(4), 389–398. https://doi.org/10.1093/aje/kwi054

Addy, C. L., Trilk, J. L., Dowda, M., Byun, W., & Pate, R. R. (2014). Assessing preschool children's physical activity: how many days of accelerometry measurement. *Pediatric Exercise Science*, *26*(1), 103–109. https://doi.org/10.1123/pes.2013-0021

Ahmadi, M. N., Brookes, D., Chowdhury, A., Pavey, T., & Trost, S. G. (2019). Free-living Evaluation of Laboratory-based Activity Classifiers in Preschoolers. *Medicine and science in sports and exercise*. https://doi.org/10.1249/MSS.0000000000002221

Aitchison, J. (1982). The Statistical Analysis of Compositional Data. *Journal of the Royal Statistical Society: Series B (Methodological)*, *44*(2), 139–160. https://doi.org/10.1111/j.2517-6161.1982.tb01195.x

Atkin, A. J., Sharp, S. J., Harrison, F. L., Brage, S., & van Sluijs, E. M. F. (2016). Seasonal Variation in Children's Physical Activity and Sedentary Time. *Medicine & Science in Sports & Exercise*, *48*(3), 449–456. https://doi.org/10.1249/MSS.0000000000000786

Ayabe, M., Kumahara, H., Morimura, K., & Tanaka, H. (2013). Epoch length and the physical activity bout analysis: an accelerometry research issue. *BMC research notes*, *6*, 20. https://doi.org/10.1186/1756-0500-6-20

Banda, J. A., Haydel, K. F., Davila, T., Desai, M., Bryson, S., Haskell, W. L., Matheson, D., & Robinson, T. N. (2016). Effects of Varying Epoch Lengths, Wear Time Algorithms, and Activity Cut-Points on Estimates of Child Sedentary Behavior and Physical Activity from Accelerometer Data. *PLOS ONE*, *11*(3), e0150534. https://doi.org/10.1371/journal.pone.0150534

Barouni, A., Ottenbacher, J., Schneider, J., Feige, B., Riemann, D., Herlan, A., El Hardouz, D., & McLennan, D. (2020). Ambulatory sleep scoring using accelerometers-distinguishing between nonwear and sleep/wake states. *PeerJ*, *8*, e8284. https://doi.org/10.7717/peerj.8284

Barreira, T. V., Schuna, J. M., Tudor-Locke, C. E., Chaput, J.-P., Church, T. S., Fogelholm, M., Hu, G., Kuriyan, R., Kurpad, A., Lambert, E. V., Maher, C. A., Maia, J., Matsudo, V., Olds, T., Onywera, V., Sarmiento, O. L., Standage, M., Tremblay, M. S., Zhao, P., & Katzmarzyk, P. T. (2015). Reliability of accelerometer-determined physical activity and sedentary behavior in school-aged children: a 12-country study. *International journal of obesity supplements*, *5*(Suppl 2), S29–35. https://doi.org/10.1038/ijosup.2015.16

Baumann, S., Groß, S., Voigt, L., Ullrich, A., Weymar, F., Schwaneberg, T., Dörr, M., Meyer, C., John, U., & Ulbricht, S. (2018). Pitfalls in accelerometer-based measurement of physical activity: The presence of reactivity in an adult population. *Scandinavian journal of medicine & science in sports*, *28*(3), 1056–1063. https://doi.org/10.1111/sms.12977.

Berger, I., Obeid, J., Timmons, B. W., & DeMatteo, C. (2017). Exploring Accelerometer Versus Self-Report Sleep Assessment in Youth With Concussion. *Global pediatric health*, *4*, 2333794X17745973. https://doi.org/10.1177/2333794X17745973

Biddle, G. J. H., Edwardson, C. L., Henson, J., Davies, M. J., Khunti, K., Rowlands, A. V., & Yates, T. (2018). Associations of Physical Behaviours and Behavioural Reallocations with Markers of Metabolic Health: A Compositional Data Analysis. *International journal of environmental research and public health*, *15*(10). https://doi.org/10.3390/ijerph15102280.

Biswas, A., Oh, P. I., Faulkner, G. E., Bajaj, R. R., Silver, M. A., Mitchell, M. S., & Alter, D. A. (2015). Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. *Annals of internal medicine*, *162*(2), 123–132. https://doi.org/10.7326/M14-1651

Bornstein, D. B., Beets, M. W., Byun, W., & McIver, K. L. (2011). Accelerometer-derived physical activity levels of preschoolers: a meta-analysis. *Journal of Science and Medicine in Sport*, *14*(6), 504–511. https://doi.org/10.1016/j.jsams.2011.05.007

Brenner, P. S., & DeLamater, J. D. (2014). Social Desirability Bias in Self-reports of Physical Activity: Is an Exercise Identity the Culprit? *Social Indicators Research*, *117*(2), 489–504. https://doi.org/10.1007/s11205-013-0359-y

Burchartz, A., Manz, K., Anedda, B., Niessner, C., Oriwol, D., Schmidt, S. C. E., & Woll, A. (2020). Measurement of Physical Activity and Sedentary Behavior by Accelerometry Among a Nationwide Sample from the KiGGS and MoMo Study: Study Protocol. *JMIR Research Protocols*, *9*(7), e14370. https://doi.org/10.2196/14370

Butte, N. F., Ekelund, U., & Westerterp, K. R. (2012). Assessing physical activity using wearable monitors: measures of physical activity. *Medicine and science in sports and exercise*, *44*(1 Suppl 1), S5–12. https://doi.org/10.1249/mss.0b013e3182399c0e.

Carr, L. J., & Mahar, M. T. (2012). Accuracy of intensity and inclinometer output of three activity monitors for identification of sedentary behavior and light-intensity activity. *Journal of obesity*, *2012*, 460271. https://doi.org/10.1155/2012/460271.

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Reports*, *100*(2), 126–131.

Cavanaugh, J. T., Kochi, N., & Stergiou, N. (2010). Nonlinear analysis of ambulatory activity patterns in community-dwelling older adults. *The journals of gerontology. Series A, Biological sciences and medical sciences*, *65*(2), 197–203. https://doi.org/10.1093/gerona/glp144.

Chaput, J.-P., Saunders, T. J., & Carson, V. (2017). Interactions between sleep, movement and other non-movement behaviours in the pathogenesis of childhood obesity. *Obesity reviews : an official journal of the International Association for the Study of Obesity*, *18 Suppl 1*, 7–14. https://doi.org/10.1111/obr.12508

Chastin, S. F. M., Palarea-Albaladejo, J., Dontje, M. L., & Skelton, D. A. (2015). Combined Effects of Time Spent in Physical Activity, Sedentary Behaviors and Sleep on Obesity and Cardio-Metabolic Health Markers: A Novel Compositional Data Analysis Approach. *PLoS One*, *10*(10), e0139984. https://doi.org/10.1371/journal.pone.0139984.

Chen, K. Y., & Bassett, D. R. (2005). The technology of accelerometry-based activity monitors: current and future. *Medicine and science in sports and exercise*, *37*(11 Suppl), S490–500. https://doi.org/10.1249/01.mss.0000185571.49104.82

Chomistek, A. K., Yuan, C., Matthews, C. E., Troiano, R. P., Bowles, H. R., Rood, J., Barnett, J. B., Willett, W. C., Rimm, E. B., & Bassett, D. R. (2017). Physical Activity Assessment with the ActiGraph GT3X and Doubly Labeled Water. *Medicine and science in sports and exercise*, *49*(9), 1935–1944. https://doi.org/10.1249/MSS.0000000000001299

Clemes, S. A., Matchett, N., & Wane, S. L. (2008). Reactivity: an issue for short-term pedometer studies? *British journal of sports medicine*, *42*(1), 68–70. https://doi.org/10.1136/bjsm.2007.038521

Colley, R. C., Garriguet, D., Janssen, I., Craig, C. L., Clarke, J., & Tremblay, M. S. (2011). Physical activity of Canadian adults: accelerometer results from the 2007 to 2009 Canadian Health Measures Survey. *Health Reports*, *22*(1), 7–14.

Crouter, S. E., DellaValle, D. M., Haas, J. D., Frongillo, E. A., & Bassett, D. R. (2013). Validity of ActiGraph 2-Regression Model, Matthews Cut-Points, and NHANES Cut-Points for Assessing Free-Living Physical Activity. *Journal of Physical Activity and Health*, *10*(4), 504–514. https://doi.org/10.1123/jpah.10.4.504

Davies, G., Reilly, J. J., & Paton, J. Y. (2012). Objective measurement of posture and posture transitions in the pre-school child. *Physiological measurement*, *33*(11), 1913–1921. https://doi.org/10.1088/0967-3334/33/11/1913.

Diaz, K. M., Howard, V. J., Hutto, B., Colabianchi, N., Vena, J. E., Safford, M. M., Blair, S. N., & Hooker, S. P. (2017). Patterns of Sedentary Behavior and Mortality in U.S. Middle-Aged and Older Adults: A National Cohort Study. *Annals of internal medicine*, *167*(7), 465–475. https://doi.org/10.7326/M17-0212

Doherty, A., Jackson, D., Hammerla, N., Plötz, T., Olivier, P., Granat, M. H., White, T., van Hees, V. T., Trenell, M. I., Owen, C. G., Preece, S. J., Gillions, R., Sheard, S., Peakman, T., Brage, S., & Wareham, N. J. (2017). Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. *PLoS One*, *12*(2), e0169649. https://doi.org/10.1371/journal.pone.0169649

Dössegger, A., Ruch, N., Jimmy, G., Braun-Fahrländer, C., Mäder, U., Hänggi, J. M., Hofmann, H., Puder, J. J., Kriemler, S., & Bringolf-Isler, B. (2014). Reactivity to accelerometer measurement of children and adolescents. *Medicine and science in sports and exercise*, *46*(6), 1140–1146. https://doi.org/10.1249/MSS.0000000000000215

Dumuid, D., Stanford, T. E., Martín-Fernández, J.-A., Pedišić, Ž., Maher, C. A., Lewis, L. K., Hron, K., Katzmarzyk, P. T., Chaput, J.-P., Fogelholm, M., Hu, G., Lambert, E. V., Maia, J., Sarmiento, O. L., Standage, M., Barreira, T. V., Broyles, S. T., Tudor-Locke, C. E., Tremblay, M. S., & Olds, T. (2018). Compositional data analysis for physical activity, sedentary time and sleep research. *Statistical methods in medical research*, *27*(12), 3726–3738. https://doi.org/10.1177/0962280217710835.

Dumuid, D., Stanford, T. E., Pedišić, Ž., Maher, C. A., Lewis, L. K., Martín-Fernández, J.-A., Katzmarzyk, P. T., Chaput, J.-P., Fogelholm, M., Standage, M., Tremblay, M. S., & Olds, T. (2018). Adiposity and the isotemporal substitution of physical activity, sedentary time and sleep among school-aged children: a compositional data analysis approach. *BMC public health*, *18*(1), 311. https://doi.org/10.1186/s12889-018-5207-1.

Ekelund, U., Steene-Johannessen, J., Brown, W. J., Fagerland, M. W., Owen, N., Powell, K. E., Bauman, A., & Lee, I.-M. (2016). Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. *The Lancet*, *388*(10051), 1302–1310. https://doi.org/10.1016/S0140-6736(16)30370-1

Elliott, P., Coats, A., Radaelli, A., & Clark, S. J. (1990). Regression dilution bias. *The Lancet*, *335*(8699), 1230–1231. https://doi.org/10.1016/0140-6736(90)92762-7

Evenson, K. R., Sotres-Alvarez, D., Deng, Y. U., Marshall, S. J., Isasi, C. R., Esliger, D. W., & Davis, S. (2015). Accelerometer adherence and performance in a cohort study of US Hispanic adults. *Medicine and science in sports and exercise*, *47*(4), 725–734. https://doi.org/10.1249/mss.0000000000000478.

Fairclough, S. J., Dumuid, D., Taylor, S., Curry, W., McGrane, B., Stratton, G., Maher, C. A., & Olds, T. (2017). Fitness, fatness and the reallocation of time between children's daily movement behaviours: an analysis of compositional data. *The international journal of behavioral nutrition and physical activity*, *14*(1), 64. https://doi.org/10.1186/s12966-017-0521-z.

Fairclough, S. J., & Stratton, G. (2006). A Review of Physical Activity Levels during Elementary School Physical Education. *Journal of Teaching in Physical Education*, *25*(2), 240–258. https://doi.org/10.1123/jtpe.25.2.240

Farrahi, V., Niemelä, M., Kangas, M., Korpelainen, R., & Jämsä, T. (2019). Calibration and validation of accelerometer-based activity monitors: A systematic review of machine-learning approaches. *Gait & posture*, *68*, 285–299. https://doi.org/10.1016/j.gaitpost.2018.12.003

Furuya, S., & Soechting, J. F. (2012). Speed invariance of independent control of finger movements in pianists. *Journal of neurophysiology*, *108*(7), 2060–2068. https://doi.org/10.1152/jn.00378.2012

German National Cohort Consortium. (2014). The German National Cohort: aims, study design and organization. *European journal of epidemiology*, *29*(5), 371–382. https://doi.org/10.1007/s10654-014-9890-7

Giurgiu, M., Koch, E. D., Plotnikoff, R. C., Ebner-Priemer, U. W., & Reichert, M. (2019). Breaking Up Sedentary Behavior Optimally to Enhance Mood. *Medicine and science in sports and exercise*. https://doi.org/10.1249/MSS.0000000000002132

Guinhouya, B. C., Samouda, H., & de Beaufort, C. (2013). Level of physical activity among children and adolescents in Europe: a review of physical activity assessed objectively by accelerometry. *Public health*, *127*(4), 301–311. https://doi.org/10.1016/j.puhe.2013.01.020.

Heil, D. P., Brage, S., & Rothney, M. P. (2012). Modeling physical activity outcomes from wearable monitors. *Medicine and science in sports and exercise*, *44*(1 Suppl 1), S50–60. https://doi.org/10.1249/mss.0b013e3182399dcc.

Helmerhorst, H. J. F., Brage, S., Warren, J. M., Besson, H., & Ekelund, U. (2012). A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. *The international journal of behavioral nutrition and physical activity*, *9*, 103. https://doi.org/10.1186/1479-5868-9-103.

Henriksen, A., Haugen Mikalsen, M., Woldaregay, A. Z., Muzny, M., Hartvigsen, G., Hopstock, L. A., & Grimsgaard, S. (2018). Using Fitness Trackers and Smartwatches to Measure Physical Activity in Research: Analysis of Consumer Wrist-Worn Wearables. *Journal of medical Internet research*, *20*(3), e110. https://doi.org/10.2196/jmir.9157

Hesse, B. W., Moser, R. P., & Riley, W. T. (2015). From Big Data to Knowledge in the Social Sciences. *The Annals of the American Academy of Political and Social Science*, *659*(1), 16–32. https://doi.org/10.1177/0002716215570007

Hidalgo-Mazzei, D., Murru, A., Reinares, M., Vieta, E., & Colom, F. (2016). Big Data in mental health: a challenging fragmented future. *World psychiatry : official journal of the World Psychiatric Association (WPA)*, *15*(2), 186–187. https://doi.org/10.1002/wps.20307

Hollis, J. L., Williams, A. J., Sutherland, R., Campbell, E., Nathan, N., Wolfenden, L., Morgan, P. J., Lubans, D. R., & Wiggers, J. (2016). A systematic review and meta-analysis of moderate-to-vigorous physical activity levels in elementary school physical education lessons. *Preventive medicine*, *86*, 34–54. https://doi.org/10.1016/j.ypmed.2015.11.018.

Hutcheon, J. A., Chiolero, A., & Hanley, J. A. (2010). Random measurement error and regression dilution bias. *BMJ (Clinical research ed.)*, *340*, c2289. https://doi.org/10.1136/bmj.c2289.

Jaeschke, L., Steinbrecher, A., Jeran, S., Konigorski, S., & Pischon, T. (2018). Variability and reliability study of overall physical activity and activity intensity levels using 24 h-accelerometry-assessed data. *BMC public health*, *18*(1), 530. https://doi.org/10.1186/s12889-018-5415-8.

Jakicic, J. M., Marcus, M., Gallagher, K. I., Randall, C., Thomas, E., Goss, F. L., & Robertson, R. J. (2004). Evaluation of the SenseWear Pro Armband to assess energy expenditure during exercise. *Medicine and science in sports and exercise*, *36*(5), 897–904. https://doi.org/10.1249/01.mss.0000126805.32659.43.

Jean-Louis, G., Kripke, D. F., Cole, R. J., Assmus, J. D., & Langer, R. D. (2001). Sleep detection with an accelerometer actigraph: comparisons with polysomnography. *Physiology & Behavior*, *72*(1-2), 21–28. https://doi.org/10.1016/s0031-9384(00)00355-3

Kang, M., & Rowe, D. A. (2015). Issues and Challenges in Sedentary Behavior Measurement. *Measurement in Physical Education and Exercise Science*, *19*(3), 105–115. https://doi.org/10.1080/1091367X.2015.1055566

Katzmarzyk, P. T., Powell, K. E., Jakicic, J. M., Troiano, R. P., Piercy, K., & Tennant, B. (2019). Sedentary Behavior and Health: Update from the 2018 Physical Activity Guidelines Advisory Committee. *Medicine and science in sports and exercise*, *51*(6), 1227–1241. https://doi.org/10.1249/MSS.0000000000001935

Kawada, T. (2008). Agreement rates for sleep/wake judgments obtained via accelerometer and sleep diary: a comparison. *Behavior Research Methods*, *40*(4), 1026–1029. https://doi.org/10.3758/brm.40.4.1026.

Keadle, S. K., Lyden, K., Strath, S. J., Staudenmayer, J. W., & Freedson, P. S. (2019). Perspectives for Progress - A Framework to Evaluate Devices that Assess Physical Behavior. *Exercise and Sport Sciences Reviews*, 1. https://doi.org/10.1249/jes.0000000000000206

Keadle, S. K., Sampson, J. N., Li, H., Lyden, K., Matthews, C. E., & Carroll, R. J. (2017). An Evaluation of Accelerometer-derived Metrics to Assess Daily Behavioral Patterns. *Medicine and science in sports and exercise*, *49*(1), 54–63. https://doi.org/10.1249/mss.0000000000001073.

Keadle, S. K., Shiroma, E. J., Kamada, M., Matthews, C. E., Harris, T. B., & Lee, I.-M. (2017). Reproducibility of Accelerometer-Assessed Physical Activity and Sedentary Time. *American journal of preventive medicine*, *52*(4), 541–548. https://doi.org/10.1016/j.amepre.2016.11.010.

Kettner, S., Kobel, S., Fischbach, N., Drenowatz, C., Dreyhaupt, J., Wirt, T., Koch, B., & Steinacker, J. M. (2013). Objectively determined physical activity levels of primary school children in south-west Germany. *BMC public health*, *13*, 895. https://doi.org/10.1186/1471-2458-13-895.

Kowalski, K., Rhodes, R., Naylor, P.-J., Tuokko, H., & MacDonald, S. (2012). Direct and indirect measurement of physical activity in older adults: a systematic review of the literature. *The international journal of behavioral nutrition and physical activity*, *9*, 148. https://doi.org/10.1186/1479-5868-9-148.

Kozey-Keadle, S., Lyden, K., Howe, C. A., Staudenmayer, J. W., & Freedson, P. S. (2010). Accelerometer output and MET values of common physical activities. *Medicine and science in sports and exercise*, *42*(9), 1776–1784. https://doi.org/10.1249/mss.0b013e3181d479f2.

Lamkin, P. (2018). Smartwatch popularity booms with fitness trackers on the slide. Retrieved September 30, 2021, from https://www.forbes.com/sites/paullamkin/2018/02/22/smartwatch-popularity-booms-with-fitness-trackers-on-the-slide

Lee, I.-M., Shiroma, E. J., Kamada, M., Bassett, D. R., Matthews, C. E., & Buring, J. E. (2019). Association of Step Volume and Intensity With All-Cause Mortality in Older Women. *JAMA internal medicine*. https://doi.org/10.1001/jamainternmed.2019.0899.

Leppänen, M. H., Migueles, J. H., Cadenas-Sanchez, C., Henriksson, P., Mora-Gonzalez, J., Henriksson, H., Labayen, I., Löf, M., Esteban-Cornejo, I., & Ortega, F. B. (2019). Hip and wrist accelerometers showed consistent associations with fitness and fatness in children aged 8-12 years. *Acta paediatrica (Oslo, Norway : 1992)*. https://doi.org/10.1111/apa.15043

Lewis, B. A., Napolitano, M. A., Buman, M. P., Williams, D. M., & Nigg, C. R. (2017). Future directions in physical activity intervention research: expanding our focus to sedentary behaviors, technology, and dissemination. *Journal of behavioral medicine*, *40*(1), 112–126. https://doi.org/10.1007/s10865-016-9797-8

Lobelo, F., Rohm Young, D., Sallis, R., Garber, M. D., Billinger, S. A., Duperly, J., Hutber, A., Pate, R. R., Thomas, R. J., Widlansky, M. E., McConnell, M. V., & Joy, E. A. (2018). Routine Assessment and Promotion of Physical Activity in Healthcare Settings: A Scientific Statement From the American Heart Association. *Circulation*, *137*(18), e495–e522. https://doi.org/10.1161/CIR.0000000000000559

Lyden, K., Keadle, S. K., Staudenmayer, J. W., & Freedson, P. S. (2014). A method to estimate free-living active and sedentary behavior from an accelerometer. *Medicine and science in sports and exercise*, *46*(2), 386–397. https://doi.org/10.1249/mss.0b013e3182a42a2d.

Matricciani, L., Fraysse, F., Grobler, A. C., Muller, J., Wake, M., & Olds, T. (2019). Sleep: population epidemiology and concordance in Australian children aged 11-12 years and their parents. *BMJ open*, *9*(Suppl 3), 127–135. https://doi.org/10.1136/bmjopen-2017-020895

Matthews, C. E., Chen, K. Y., Freedson, P. S., Buchowski, M. S., Beech, B. M., Pate, R. R., & Troiano, R. P. (2008). Amount of time spent in sedentary behaviors in the United States, 2003-2004. *American journal of epidemiology*, *167*(7), 875–881. https://doi.org/10.1093/aje/kwm390

Matthews, C. E., Freedson, P. S., Hebert, J. R., Stanek, E. J., Merriam, P. A., Rosal, M. C., Ebbeling, C. B., & Ockene, I. S. (2001). Seasonal variation in household, occupational, and leisure time physical activity: longitudinal analyses from the seasonal variation of blood cholesterol study. *American journal of epidemiology*, *153*(2), 172–183. https://doi.org/10.1093/aje/153.2.172

Matthews, C. E., Keadle, S. K., Troiano, R. P., Kahle, L., Koster, A., Brychta, R. J., van Domelen, D., Caserotti, P., Chen, K. Y., Harris, T. B., & Berrigan, D. (2016). Accelerometer-measured dose-response for physical activity, sedentary time, and mortality in US adults. *The American journal of clinical nutrition*, *104*(5), 1424–1432. https://doi.org/10.3945/ajcn.116.135129

McMurray, R. G., Butte, N. F., Crouter, S. E., Trost, S. G., Pfeiffer, K. A., Bassett, D. R., Puyau, M. R., Berrigan, D., Watson, K. B., & Fulton, J. E. (2015). Exploring Metrics to Express Energy Expenditure of Physical Activity in Youth. *PLoS One*, *10*(6), e0130869. https://doi.org/10.1371/journal.pone.0130869

Mekary, R. A., & Ding, E. L. (2019). Isotemporal Substitution as the Gold Standard Model for Physical Activity Epidemiology: Why It Is the Most Appropriate for Activity Time Research. *International journal of environmental research and public health*, *16*(5). https://doi.org/10.3390/ijerph16050797.

Mekary, R. A., Willett, W. C., Hu, F. B., & Ding, E. L. (2009). Isotemporal substitution paradigm for physical activity epidemiology and weight change. *American journal of epidemiology*, *170*(4), 519–527. https://doi.org/10.1093/aje/kwp163.

Migueles, J. H., Aadland, E., Andersen, L. B., Brønd, J. C., Chastin, S. F. M., Hansen, B. H., Konstabel, K., Kvalheim, O. M., McGregor, D. E., Rowlands, A. V., Sabia, S., van Hees, V. T., Walmsley, R., Ortega, F. B., & External Review Group. (2021). GRANADA consensus on analytical approaches to assess associations with accelerometer-determined physical behaviours (physical activity, sedentary behaviour and sleep) in epidemiological studies. *British journal of sports medicine*. https://doi.org/10.1136/bjsports-2020-103604

Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Delisle Nyström, C., Mora-Gonzalez, J., Löf, M., Labayen, I., Ruiz, J. R., & Ortega, F. B. (2017). Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports medicine (Auckland, N.Z.)*, *47*(9), 1821–1845. https://doi.org/10.1007/s40279-017-0716-0

Montoye, A. H. K., Moore, R. W., Bowles, H. R., Korycinski, R., & Pfeiffer, K. A. (2018). Reporting accelerometer methods in physical activity intervention studies: a systematic review and recommendations for authors. *British journal of sports medicine*, *52*(23), 1507–1516. https://doi.org/10.1136/bjsports-2015-095947

Nascimento-Ferreira, M. V., Collese, T. S., de Moraes, A. C. F., Rendo-Urteaga, T., Moreno Aznar, L. A., & Carvalho, H. B. (2016). Validity and reliability of sleep time questionnaires in children and adolescents: A systematic review and meta-analysis. *Sleep medicine reviews*, *30*, 85–96. https://doi.org/10.1016/j.smrv.2015.11.006.

Nigg, C. R., Jordan, P. J., & Atkins, A. (2012). Behavioral measurement in exercise psychology. In G. Tenenbaum, R. C. Eklund, & A. Kamata (Eds.), *Measurement in sport and exercise psychology* (pp. 455–464). Human Kinetics.

Nigg, C. R., & Woll, A. (2020). Best practices and future research directions: Consensus from the 2 International Workshop of the Center for the Assessment of Physical Activity (CAPA). *Psychology of Sport and Exercise*, *50*, 101734. https://doi.org/10.1016/j.psychsport.2020.101734

Olds, T., Maher, C. A., & Matricciani, L. (2011). Sleep duration or bedtime? Exploring the relationship between sleep habits and weight status and activity patterns. *Sleep*, *34*(10), 1299–1307. https://doi.org/10.5665/SLEEP.1266

Pedišić, Ž., & Bauman, A. (2015). Accelerometer-based measures in physical activity surveillance: current practices and issues. *British journal of sports medicine*, *49*(4), 219–223. https://doi.org/10.1136/bjsports-2013-093407

Preece, S. J., Goulermas, J. Y., Kenney, L. P. J., Howard, D., Meijer, K., & Crompton, R. (2009). Activity identification using body-mounted sensors–a review of classification techniques. *Physiological measurement*, *30*(4), R1–33. https://doi.org/10.1088/0967-3334/30/4/r01.

Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Connor-Gorber, S., & Tremblay, M. S. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: A systematic review. *The international journal of behavioral nutrition and physical activity*, *5*, 56. https://doi.org/10.1186/1479-5868-5-56

Reichert, M., Giurgiu, M., Koch, E. D., Wieland, L. M., Lautenbach, S., Neubauer, A. B., von Haaren-Mack, B., Schilling, R., Timm, I., Notthoff, N., Marzi, I., Hill, H., Brüssler, S., Eckert, T., Fiedler, J., Burchartz, A., Anedda, B., Wunsch, K., Gerber, M., . . . Liao, Y. (2020). Ambulatory Assessment for Physical Activity Research: State of the Science, Best Practices and Future Directions. *Psychology of Sport and Exercise*, *50*. https://doi.org/10.1016/j.psychsport.2020.101742

Reilly, J. J., Penpraze, V., Hislop, J., Davies, G., Grant, S., & Paton, J. Y. (2008). Objective measurement of physical activity and sedentary behaviour: review with new data. *Archives of disease in childhood*, *93*(7), 614–619. https://doi.org/10.1136/adc.2007.133272.

Rowlands, A. V. (2018). Moving Forward With Accelerometer-Assessed Physical Activity: Two Strategies to Ensure Meaningful, Interpretable, and Comparable Measures. *Pediatric Exercise Science*, *30*(4), 450–456. https://doi.org/10.1123/pes.2018-0201

Rowlands, A. V., Edwardson, C. L., Davies, M. J., Khunti, K., Harrington, D. M., & Yates, T. (2018). Beyond Cut Points: Accelerometer Metrics that Capture the Physical Activity Profile. *Medicine and science in sports and exercise*, *50*(6), 1323–1332. https://doi.org/10.1249/MSS.0000000000001561

Rowlands, A. V., Fairclough, S. J., Yates, T., Edwardson, C. L., Davies, M. J., Munir, F., Khunti, K., & Stiles, V. H. (2019). Activity Intensity, Volume, and Norms: Utility and Interpretation of Accelerometer Metrics. *Medicine and science in sports and exercise*, *51*(11), 2410–2422. https://doi.org/10.1249/MSS.0000000000002047

Rowlands, A. V., Mirkes, E. M., Yates, T., Clemes, S., Davies, M. J., Khunti, K., & Edwardson, C. L. (2018). Accelerometer-assessed Physical Activity in Epidemiology: Are Monitors Equivalent? *Medicine and science in sports and exercise*, *50*(2), 257–265. https://doi.org/10.1249/mss.0000000000001435

Rowlands, A. V., Sherar, L. B., Fairclough, S. J., Yates, T., Edwardson, C. L., Harrington, D. M., Davies, M. J., Munir, F., Khunti, K., & Stiles, V. H. (2019). A data-driven, meaningful, easy to interpret, standardised accelerometer outcome variable for global surveillance. *Journal of Science and Medicine in Sport*, *22*(10), 1132–1138. https://doi.org/10.1016/j.jsams.2019.06.016

Rowlands, A. V. (2007). Accelerometer Assessment of Physical Activity in Children: An Update. *Pediatric Exercise Science*, *19*(3), 252–266. https://doi.org/10.1123/pes.19.3.252

Rowlands, A. V., & Eston, R. G. (2007). The Measurement and Interpretation of Children's Physical Activity. *Journal of sports science & medicine*, *6*(3), 270–276.

Saint-Maurice, P. F., Kim, Y., Welk, G. J., & Gaesser, G. A. (2016). Kids are not little adults: what MET threshold captures sedentary behavior in children? *European journal of applied physiology*, *116*(1), 29–38. https://doi.org/10.1007/s00421-015-3238-1

Sallis, J. F., & Saelens, B. E. (2000). Assessment of physical activity by self-report: status, limitations, and future directions. *Research quarterly for exercise and sport*, *71 Suppl 2*, 1–14. https://doi.org/10.1080/02701367.2000.11082780

Schaefer, C. A., Nigg, C. R., Hill, J. O., Brink, L. A., & Browning, R. C. (2014). Establishing and evaluating wrist cutpoints for the GENEActiv accelerometer in youth. *Medicine and science in sports and exercise*, *46*(4), 826–833. https://doi.org/10.1249/MSS.0000000000000150

Schneller, M. B., Bentsen, P., Nielsen, G., Brønd, J. C., Ried-Larsen, M., Mygind, E., & Schipperijn, J. (2017). Measuring Children's Physical Activity: Compliance Using Skin-Taped Accelerometers. *Medicine and science in sports and exercise*, *49*(6), 1261–1269. https://doi.org/10.1249/mss.0000000000001222

Servais, S. B., Webster, J. G., & Montoye, H. J. (1984). Estimating human energy expenditure using an accelerometer device. *Journal of Clinical Engineering*, *9*(2), 159–170.

Shannon, C. E. (1949). Communication in the Presence of Noise. *Proceedings of the IRE*, *37*(1), 10–21. https://doi.org/10.1109/JRPROC.1949.232969

Sigmund, E., Sigmundová, D., Hamrik, Z., & Madarásová Gecková, A. (2014). Does participation in physical education reduce sedentary behaviour in school and throughout the day among normal-weight and overweight-to-obese Czech children aged 9-11 years? *International journal of environmental research and public health*, *11*(1), 1076–1093. https://doi.org/10.3390/ijerph110101076.

Smith, C., Galland, B. C., Taylor, R. W., & Meredith-Jones, K. A. (2019). ActiGraph GT3X+ and Actical Wrist and Hip Worn Accelerometers for Sleep and Wake Indices in Young Children Using an Automated Algorithm: Validation With Polysomnography. *Frontiers in psychiatry*, *10*, 958. https://doi.org/10.3389/fpsyt.2019.00958

Sprengeler, O., Buck, C., Hebestreit, A., Wirsik, N., & Ahrens, W. (2019). Sports Contribute to Total Moderate to Vigorous Physical Activity in School Children. *Medicine and science in sports and exercise*, *51*(8), 1653–1661. https://doi.org/10.1249/mss.0000000000001948.

Sprengeler, O., Wirsik, N., Hebestreit, A., Herrmann, D., & Ahrens, W. (2017). Domain-Specific Self-Reported and Objectively Measured Physical Activity in Children. *International journal of environmental research and public health*, *14*(3). https://doi.org/10.3390/ijerph14030242

Taraldsen, K., Chastin, S. F. M., Riphagen, I. I., Vereijken, B., & Helbostad, J. L. (2012). Physical activity monitoring by use of accelerometer-based body-worn sensors in older adults: a systematic literature review of current knowledge and applications. *Maturitas*, *71*(1), 13–19. https://doi.org/10.1016/j.maturitas.2011.11.003.

Taylor, R. W., Haszard, J. J., Meredith-Jones, K. A., Galland, B. C., Heath, A.-L. M., Lawrence, J., Gray, A. R., Sayers, R., Hanna, M., & Taylor, B. J. (2018). 24-h movement behaviors from infancy to preschool: cross-sectional and longitudinal relationships with body composition and bone health. *The international journal of behavioral nutrition and physical activity*, *15*(1), 118. https://doi.org/10.1186/s12966-018-0753-6

Teulings, H. L., & Maarse, F. J. (1984). Digital recording and processing of handwriting movements. *Human Movement Science*, *3*, 193–217.

Thompson, W. R. (2018). Worldwide survey of fitness trends for 2019. *ACSM's Health & Fitness Journal*, *22*(6), 10–17.

Toftager, M., Kristensen, P. L., Oliver, M., Duncan, S., Christiansen, L. B., Boyle, E., Brønd, J. C., & Troelsen, J. (2013). Accelerometer data reduction in adolescents: effects on sample retention and bias. *International Journal of Behavioral Nutrition and Physical Activity*, *10*, 140. https://doi.org/10.1186/1479-5868-10-140

Tong, C., Zhang, J., Chowdhury, A., & Trost, S. G. (2019). An Interactive Visualization Tool for Sensor-based Physical Activity Data Analysis. In Unknown (Ed.), *Proceedings of the Australasian Computer Science Week Multiconference* (pp. 1–4). ACM. https://doi.org/10.1145/3290688.3290734

Toth, L. P., Park, S., Springer, C. M., Feyerabend, M. D., Steeves, J. A., & Bassett, D. R. (2018). Video-Recorded Validation of Wearable Step Counters under Free-living Conditions. *Medicine and science in sports and exercise*, *50*(6), 1315–1322. https://doi.org/10.1249/MSS.0000000000001569

Tremblay, M. S., Aubert, S., Barnes, J. D., Saunders, T. J., Carson, V., Latimer-Cheung, A. E., Chastin, S. F. M., Altenburg, T. M., & Chinapaw, M. J. M. (2017). Sedentary Behavior Research Network (SBRN) - Terminology Consensus Project process and outcome. *The international journal of behavioral nutrition and physical activity*, *14*(1), 75. https://doi.org/10.1186/s12966-017-0525-8

Troiano, R. P. (2006). Translating accelerometer counts into energy expenditure: advancing the quest. *Journal of applied physiology (Bethesda, Md. : 1985)*, *100*(4), 1107–1108. https://doi.org/10.1152/japplphysiol.01577.2005.

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine and science in sports and exercise*, *40*(1), 181–188. https://doi.org/10.1249/mss.0b013e31815a51b3

Troiano, R. P., McClain, J. J., Brychta, R. J., & Chen, K. Y. (2014). Evolution of accelerometer methods for physical activity research. *British journal of sports medicine*, *48*(13), 1019–1023. https://doi.org/10.1136/bjsports-2014-093546

Trost, S. G., McIver, K. L., & Pate, R. R. (2005). Conducting accelerometer-based activity assessments in field-based research. *Medicine and science in sports and exercise*, *37*(11 Suppl), S531–43. https://doi.org/10.1249/01.mss.0000185657.86065.98

Trost, S. G., Pate, R. R., Freedson, P. S., Sallis, J. F., & Taylor, W. C. (2000). Using objective physical activity measures with youth: how many days of monitoring are needed? *Medicine and science in sports and exercise*, *32*(2), 426–431. https://doi.org/10.1097/00005768-200002000-00025.

Tudor-Locke, C. E., Barreira, T. V., Schuna, J. M., Mire, E. F., Chaput, J.-P., Fogelholm, M., Hu, G., Kuriyan, R., Kurpad, A., Lambert, E. V., Maher, C. A., Maia, J., Matsudo, V., Olds, T., Onywera, V., Sarmiento, O. L., Standage, M., Tremblay, M. S., Zhao, P., . . . Katzmarzyk, P. T. (2015). Improving wear time compliance with a 24-hour waist-worn accelerometer protocol in the International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE). *The international journal of behavioral nutrition and physical activity*, *12*, 11. https://doi.org/10.1186/s12966-015-0172-x.

U.S. Department of Health and Human Services. (2018). Physical Activity Guidelines for Americans, 2nd edition. https://health.gov/paguidelines/second-edition/pdf/Physical_Activity_Guidelines_2nd_edition.pdf

Vähä-Ypyä, H., Husu, P., Suni, J., Vasankari, T., & Sievänen, H. (2018). Reliable recognition of lying, sitting, and standing with a hip-worn accelerometer. *Scandinavian journal of medicine & science in sports*, *28*(3), 1092–1102. https://doi.org/10.1111/sms.13017

van der Merwe, S. (2018). A method for Bayesian regression modelling of composition data. http://arxiv.org/pdf/1801.02954v1

van Hees, V. T., Sabia, S., Anderson, K. N., Denton, S. J., Oliver, J., Catt, M., Abell, J. G., Kivimäki, M., Trenell, M. I., & Singh-Manoux, A. (2015). A Novel, Open Access Method to Assess Sleep Duration Using a Wrist-Worn Accelerometer. *PLoS One*, *10*(11), e0142533. https://doi.org/10.1371/journal.pone.0142533.

Wahl, Y., Düking, P., Droszez, A., Wahl, P., & Mester, J. (2017). Criterion-Validity of Commercially Available Physical Activity Tracker to Estimate Step Count, Covered Distance and Energy Expenditure during Sports Conditions. *Frontiers in physiology*, *8*, 725. https://doi.org/10.3389/fphys.2017.00725

Ward, D. S., Evenson, K. R., Vaughn, A., Rodgers, A. B., & Troiano, R. P. (2005). Accelerometer use in physical activity: best practices and research recommendations. *Medicine and science in sports and exercise*, *37*(11 Suppl), S582–8. https://doi.org/10.1249/01.mss.0000185292.71933.91.

Welk, G. J. (Ed.). (2002). *Physical activity assessments for health-related research*. Human Kinetics.

Westerterp, K. R. (2009). Assessment of physical activity: a critical appraisal. *European journal of applied physiology*, *105*(6), 823–828. https://doi.org/10.1007/s00421-009-1000-2.

Wolff-Hughes, D. L., McClain, J. J., Dodd, K. W., Berrigan, D., & Troiano, R. P. (2016). Number of accelerometer monitoring days needed for stable group-level estimates of activity. *Physiological measurement*, *37*(9), 1447–1455. https://doi.org/10.1088/0967-3334/37/9/1447.

Wong, T. C., Webster, J. G., Montoye, H. J., & Washburn, R. (1981). Portable accelerometer device for measuring human energy expenditure. *IEEE Transactions on Biomedical Engineering*, (6), 467–471.

Xu, F., Adams, S. K., Cohen, S. A., Earp, J. E., & Greaney, M. L. (2019). Relationship between Physical Activity, Screen Time, and Sleep Quantity and Quality in US Adolescents Aged 16−19. *International journal of environmental research and public health*, *16*(9). https://doi.org/10.3390/ijerph16091524

Zhou, M., Lalani, C., Banda, J. A., & Robinson, T. N. (2018). Sleep duration, timing, variability and measures of adiposity among 8- to 12-year-old children with obesity. *Obesity science & practice*, *4*(6), 535–544. https://doi.org/10.1002/osp4.303

3. Measurement of Physical Activity with Accelerometry in MoMo

Slightly modified version of the 1^{st} published article

Burchartz, Alexander; Manz, Kristin; Anedda, Bastian; Niessner, Claudia; Oriwol, Doris; Schmidt, Steffen CE; Woll, Alexander (2020): Measurement of Physical Activity and Sedentary Behavior by Accelerometry Among a Nationwide Sample from the KiGGS and MoMo Study: Study Protocol

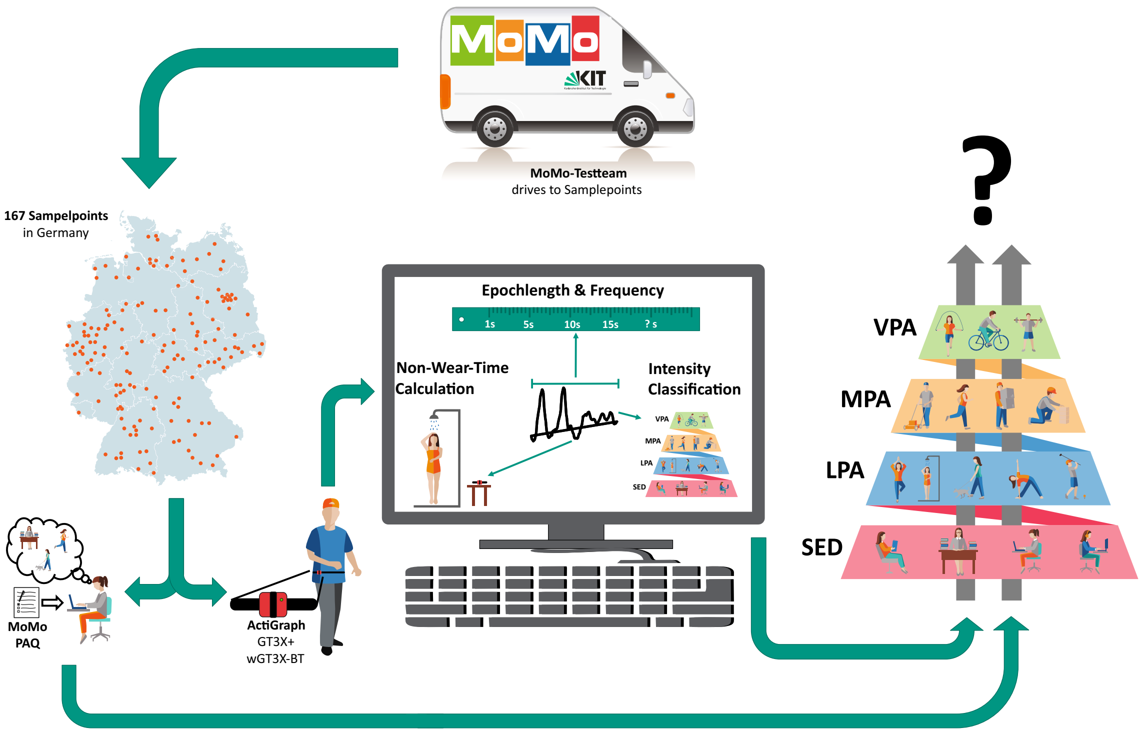
JMIR Research Protocols 9 (7), e14370.
DOI: 10.2196/14370.

3.1 Abstract

Background: Currently, no nationwide objective physical activity data for children and adolescents living in Germany exists. The German Health Interview and Examination Survey for Children and Adolescents (KiGGS) and the Motorik-Modul Study (MoMo) are a national cohort study that has incorporated accelerometers in its most recent data collection wave (wave 2, since 2014). This wave 2 marks the first nationwide collection of objective data on physical activity of children and adolescents living in Germany.
Objective: The purpose of this protocol is to describe the methods used in the KiGGS and MoMo study to capture the intensity, frequency, and duration of physical activity with accelerometers.
Methods: Participants (N=11,003, aged 6 to 31 years) were instructed to wear an ActiGraph GT3X+ or wGT3X-BT accelerometer laterally on the right hip. Accelerometers were worn on consecutive days during waking hours, including at least 4 valid weekdays and 1 weekend day (wear time >8 hours) in the evaluation. A non-wear time protocol was also implemented.
Results: Data collection was completed by October 2017. Data harmonization took place in 2018. The first accelerometer results from this wave were published in 2019, and detailed analyzes are ready to be submitted in 2020.

Conclusions: This study protocol provides an overview of the technical details and basic choices when using accelerometers in large-scale epidemiological studies. At the same time, the restrictions imposed by the specified filters and the evaluation routines must be taken into account.

Figure 3.1: Study Protocol Infographic



3.2 Introduction

The health benefits of regular physical activity are well documented in the public health literature. However, assessing physical activity in children remains difficult because the energy expenditure of a small active person can be as high as that of a large inactive person (Cain et al., 2013; Ekelund et al., 2004; Gabrys et al., 2015; Rowlands, 2007). Because children show more complex but less structured movement behaviors than adults (Gabrys et al., 2015; Graf et al., 2013), capturing their many spontaneous and impulsive movements is a great challenge for physical activity assessment (Müller et al., 2011). Currently, questionnaires are still the most widely used subjective method to assess physical activity. One of the greatest advantages of questionnaires is their versatility. In addition to recording the duration, frequency, and intensity of physical activity, questionnaire methods can also be used to collect information about the type of physical activity, which has only recently become possible with accelerometers. Furthermore, in the context of large-scale epidemiological or health science studies, questionnaires are the only feasible alternative for practical and financial reasons (Jekauc et al., 2013). In contrast, many empirical studies have already shown that the level of physical activity subjectively assessed by questionnaires is often overestimated (Jekauc et al., 2013; Slootmaker et al., 2009; Wijndaele et al., 2015). In particular, irregular and unstructured activities in everyday life are difficult to retrieve from memory correctly. In recent years, accelerometers have been used more frequently in large-scale studies

(Riddoch et al., 2004; Ruiz et al., 2011; Troiano et al., 2014; Wanner et al., 2017) because they have become more feasible, more accurate, and much more affordable.

Although accelerometers are used more frequently, there is no consensus on the use of accelerometers for the assessment of physical activity in nationwide studies in adolescents or in children (Ettienne et al., 2016; Migueles et al., 2017; Schaefer et al., 2014). Due to the rapid development in this field and the extremely large amounts of data collected, many current studies do not accurately document the use of accelerometers in detail (e.g. technical details of settings and evaluation) (Migueles et al., 2017). This complicates the replication and comparison of these studies because there are only a few representative studies worldwide (Riddoch et al., 2004; Tudor-Locke et al., 2003). Until 2014, no nationwide study had been performed in Germany in which physical activity was measured with accelerometers.

The objective of the Motorik-Modul study (MoMo), as part of the German Health Interview and Examination Survey for Children and Adolescents (KiGGS), was to establish regular monitoring of physical fitness and physical activity of children and adolescents living in Germany and to gain insight into their determinants and consequences for health outcomes. The MoMo study was established in 2003 and is based on a cohort-sequence design; four measurement waves (baseline and waves 1, 2, and 3) were planned from 2003 to 2021. Up to 2014 (baseline and wave 1), physical activity was assessed solely using a validated physical activity questionnaire (PAQ) (Jekauc et al., 2013). In KiGGS and MoMo wave 2, physical activity was also assessed by accelerometers.

The purpose of this study protocol is to explain the challenges faced when using accelerometers in the MoMo and KiGGS studies as an example of a large-scale epidemiological study and to detail the methods and protocols used to capture physical activity in children and adolescents with accelerometers in Germany.

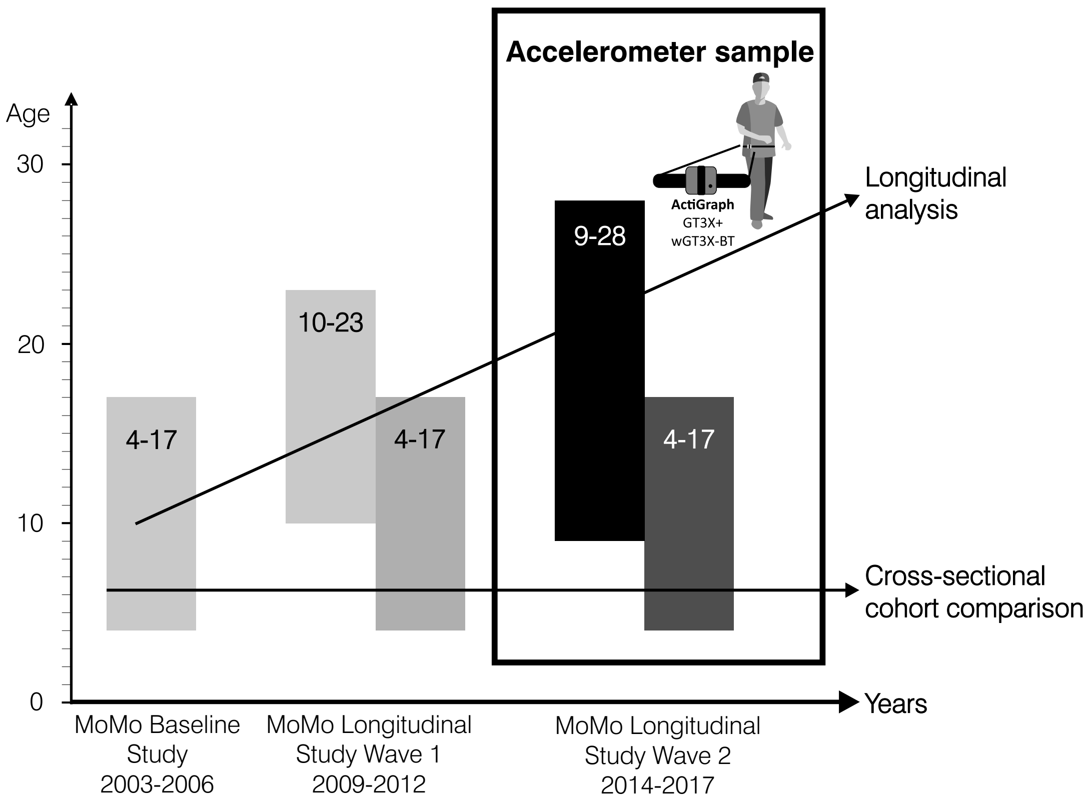
3.3 Methods

3.3.1 Study Design

KiGGS is part of the German health monitoring system established by the Robert Koch Institute. KiGGS research topics include physical health, mental health, health-related behavior, health care, prevention, and social and environmental determinants. The study design and the sampling procedure are described in detail elsewhere (Kurth et al., 2008). The core KiGGS survey is supplemented by the MoMo study, an in-depth study to assess the physical activity and motor performance of children and adolescents living in Germany that is being conducted by the Karlsruhe Institute of Technology. MoMo is being carried out with a subsample of KiGGS participants, as described in (Wagner et al., 2014; Woll et al., 2017). The KiGGS team established temporary study centers at 167 sample points throughout Germany. Participants aged 0 to 17 years were randomly selected from 167 registration offices and invited for interviews, physical examinations, and laboratory tests. In the study centers, KiGGS participants were asked if they were

willing to participate in the MoMo study. If they consented, the interviews and physical examinations for the MoMo study took place approximately six to eight weeks later.
To date, three assessments have been conducted in KiGGS and MoMo: baseline (2003-2006; sample sizes: KiGGS n=17,641; ages: 0 to 17 years; MoMo, n=4528; ages: 4 to 17 years), the first follow-up (wave 1) between 2009 and 2012 (sample sizes: KiGGS, n=12,368; ages: 0 to 17 years; MoMo, n=5106; ages: 4 to 23 years), and the second follow-up (wave 2) between 2014 and 2017 (sample sizes: KiGGS, n=15,023; ages: 0 to 17 years; MoMo, n=5689; ages: 4 to 30 years). Wave 3 of the MoMo Study is currently underway (2018-2020; compare (Mauz et al., 2017; Wagner et al., 2014; Woll et al., 2017; Worth et al., 2015)). For the follow-up surveys (waves 1, 2, and 3), participants in the baseline survey were reinvited (longitudinal subjects). In addition, for cross-sectional analysis, a new sample of children aged 0 to 6 years was drawn in KiGGS wave 1, and in wave 2, a new sample of participants aged 0 to 17 years was drawn. For detailed sample sizes, see (Mauz et al., 2017; Woll et al., 2017). In KiGGS and MoMo wave 2 (2014-2017), accelerometry was used for the first time in this study to measure physical activity.

Figure 3.2: MoMo study design with the accelerometer sample in wave 2. MoMo: Motorik-Modul.



3.3.2 Accelerometer Sample

In KiGGS wave 2, all longitudinal participants aged $\geq$ 10 years (n=6465) were included in the accelerometer sample. In MoMo wave 2, all cross-sectional participants (n=4538) in the MoMo wave 2 sample who did not receive an accelerometer in KiGGS wave 2 were asked to wear one (Figure 3.2). Thus, a total of 11,003 participants were asked to wear an accelerometer. Participants who had impairments that prevented them from wearing an accelerometer were excluded. Participants who dropped out (did not agree to wear an accelerometer or experienced technical problems) in the MoMo and KiGGS studies are listed in Table 3.1 and Table 3.2, respectively.

Table 3.1: Details of the MoMo study participants asked to wear an accelerometer (N=4538), n (%).

Participation	Participants
Agreed to wear an accelerometer	2105 (46.4)
Dropped out because they did not agree to wear an accelerometer	2433 (53.6)
Downloaded accelerometer records	1974 (43.5)
Dropped out due to technical problems or did not wear the accelerometer	131 (2.9)

Table 3.2: Details of the KiGGS study participants asked to wear an accelerometer (N=6465), n (%).

Participation	Participants
Agreed to wear an accelerometer	5040 (78.0)
Dropped out because they did not agree to wear an accelerometer	1425 (22.0))
Downloaded accelerometer records	4750 (73.5)
Dropped out due to technical problems or did not wear the accelerometer	426 (6.6)

3.3.3 Types of Accelerometers

In KiGGS and MoMo wave 2, ActiGraph accelerometers (models: GT3X+ and wGT3X-BT) were used to enable comparison with other large-scale European studies (Riddoch et al., 2004; Ruiz et al., 2011). Both the heart rate monitor and the Bluetooth wireless interface of each accelerometer were deactivated during testing. The accelerometers were equipped with a triaxial acceleration sensor (range: $\pm$6g, sensitivity: 3mg, axes: horizontal right-left (x), vertical (y), and horizontal front-back (z)); they could record acceleration data at rates ranging from 30-100 Hertz and store them in epoch lengths of 1-240 seconds (compare with ActiGraph, 2013). The settings are described in Section 3.3.5 'Initializing the Devices'. The physical dimensions of the devices were $4.6\times3.3\times1\times5$ centimeters, their weight was 19 grams, and they used a rechargeable lithium polymer battery.

3.3.4 Assessment Period and Registration Protocol

MoMo and KiGGS accelerometer data sets were considered valid with a minimum wear time of 8 hours of recordings on 4 weekdays and 1 weekend day. These scoring policies are also consistent with the inclusion requirements in the International Children's Accelerometry Database (ICAD) (Sherar et al., 2011). Furthermore, the literature suggests that 4 days is a reasonable measurement time, thus reducing the burden on participants and making it easier for researchers to collect sufficient data for the formation of recommendations related to general health guidelines (Colley et al., 2011; Mâsse et al., 2005; Matthews et al., 2002; Toftager et al., 2013). To archive the highest possible number of valid data sets, the accelerometer should be worn for 7 days (8 days in the MoMo

study) following the day of examination in the study center. The assessment period of at least seven days ensured the inclusion of weekdays and weekend days. This inclusion is recommended due to differences in physical activity during the week and on weekends (Donaldson et al., 2016; Matthews et al., 2002; Trost et al., 2005). An additional analysis was planned of data sets with 10 or more hours of recording on each of the 5-7 days. This analysis was included because recent studies (Aadland & Ylvisaker, 2015; Donaldson et al., 2016; Herrmann et al., 2013; Warren et al., 2010) propose the use of longer accelerometer wear times in hours and days to provide better estimates of daily activity.
The accelerometer should only be removed at bedtime, during activities that risk damaging the device (eg martial arts), or when the participant is exposed to water (eg swimming and showering).

3.3.5 Initializing the Devices: Epoch Length, Sampling Frequency, and Filter

Each ActiGraph activity monitor was initialized using a standardized procedure prior to being given to the participant. The monitors used the latest firmware (v1.9.2. for wGT3X-BT and v3.2.1 for GT3X+), a unique output filename, and a sampling frequency of 30 Hz. In research with adults, the accelerometer signal was processed in epoch lengths of 10-60 seconds. Due to the sporadic activity of children, an epoch length between 1 and 5 seconds or the shortest possible epoch length is recommended (Bornstein et al., 2011; Edwardson & Gorely, 2010; Migueles et al., 2017; Vanhelst et al., 2012). The ActiGraph models used in the KiGGS and MoMo study store the collected raw data. Furthermore, the data are downloaded in epoch lengths of 1 second, reducing memory space and enabling faster data processing afterward. The devices can be used with two different filters when processing the data: a normal filter and a low-frequency extension filter (the implementation and algorithms of the filters are not open to the public). The normal filter is known to detect accelerations in the range of 0.25-2.5 Hz (Tryon & Williams, 1996). To capture slower movements, the low-frequency extension filter establishes an unknown lower threshold (ActiGraph, 2013). During vigorous physical activity, the human body produces hip accelerations of up to 3.4 Hz (Cavagna & Franzetti, 1986; Cavagna et al., 1991). Even higher frequencies were documented in the wrist when performing physical activity (Fairclough et al., 2016; Hildebrand et al., 2014). Considering these limitations, the normal filter was configured to recognize as many accelerations as possible.
In KiGGS, the device was set up to start measurement at 12:00 AM the day after the examination and stop measurement at midnight after 7 days of recording. A pilot study prior to data collection in MoMo revealed that many participants were confused by the standard "no flashing" mode of the ActiGraph device during recording. Therefore, a flash mode was activated during the recording, in which the device showed a green flashing light-emitting diode (LED). In MoMo, the device was programmed to start at 12:00 a.m. the day the participants underwent their motor performance tests to avoid the confusion of "no flashing" noted above. The measurement stopped at midnight after 8 days of recording. In the MoMo study, the recordings of the first day were not considered for data analysis because participants received the devices at different times throughout the day, depending on the initial examination time. Furthermore, the first day served as an adaptation period for the participants.

3.3.6 Placement of the Device

In KiGGS and MoMo, the device was placed laterally on top of the right anterior superior iliac spine with the closure on top, and then secured with an elastic belt (see Figure 3.3). Compared to wrist attachment of the device, hip monitor placement provides better acceleration detection due to the limited frequency range of the normal ActiGraph filter. The higher movement frequencies at the wrist would be outside the range of that filter. Additionally, most of the cut-off points for intensity estimation are validated with the device placed on the hip (Hänggi et al., 2013; Hildebrand et al., 2014; Migueles et al., 2017; Romanzini et al., 2014), and it is the most common carrying position for accelerometers (Anastasopoulou et al., 2012). More importantly, studies show a more accurate classification of intensities when the device is placed on the hip than on the wrist (Crouter et al., 2015; Hänggi et al., 2013; Hildebrand et al., 2014; Migueles et al., 2017; Stec & Rawson, 2012).

Figure 3.3: Participant wearing an accelerometer device on top of the right anterior superior iliac spine.

3.3.7 Non-wear Time Protocol (Logbook)

The accelerometer can only capture accelerations when worn; therefore, detailed information on the type of activity and the reasons for not wearing the devices is needed for a complete understanding of the evaluated physical activity. Therefore, participants were asked to complete a non-wear time protocol (see Figure 3.4). The combination of self-reports and device-based measures allows a better understanding of physical activity behavior (Sallis & Saelens, 2000). When both self-reported and device-based measured physical activity assessments are available, it is also possible to cross-validate the non-wear time calculated by the algorithms and the self-reported non-wear time. With information about the reasons for not wearing the device, statistics can be created for activities that were not captured, and adjustment factors can be calculated.

Figure 3.4: Sample MoMo wear time-protocol (translated).

Subject-ID: __ __ __ __ __ __ **Aktivity Wear-time Protocol**

	Got up in the morning	Accelerometer applied (in the morning)	Gone to bed	Accelerometer put down (in the evening)	Not worn		Activity during unsupported time	Special Events
		(e.g. 8.30 am)	(e.g. 21.30 pm)	(e.g. 22.00 pm)	from	to		
Example Day 1 Date: 20.01.2014	7:00 AM	7:15 AM	10:00 PM	10:15 PM	6:00 PM	7:00 PM	Injury risk at judo	
					7:20 PM	7:45 PM	Showering	

3.3.8 Transfer and Return of the Devices

Trained study assistants at the study centers distributed the devices. The participants chose the appropriate belt size and were shown how to properly wear the device. Important aspects of accelerometer use (placement, wear times, and device return) were summarized in an information sheet provided to participants. At the end of the measurement period, the device, belt, and protocol were returned by mail. Therefore, an addressed and stamped envelope was provided to the participants. A follow-up protocol was implemented by telephone if the devices were not returned after 2 weeks.

3.3.9 Data Download and Preparation

After receiving the devices, the data was downloaded as gt3x files using ActiLife Version 6.13.3 software (ActiGraph). The MoMo team marked all data sets with less than 4+1 days of wear time as invalid. Additionally, ActiGraphData (AGD) files with data for all three axes and an epoch length of 1 s were created for data analysis. These data sets (1 second epoch length) can be converted into data sets with other epoch lengths, allowing comparison with studies using different data analysis protocols (Choi et al., 2011; Evenson et al., 2008; Freedson et al., 2011; Hänggi et al., 2013; Romanzini et al., 2014; Sasaki et al., 2011; Troiano et al., 2008). This is a faster way to process the data than to convert raw data files (*.gt3x) to data files (*.agd) with different epoch lengths. Additionally, the time saved during the calculation is enormous when analyzing large data sets of several thousand participants. The resulting data sets would be the same. The gt3x files were stored separately to allow a more in-depth analysis of the raw data after planned evaluations (e.g. with the GGIR software package (van Hees et al., 2019)).

3.3.10 Planned Data Processing

To ensure comparability with studies already included in the ICAD (Sherar et al., 2011), we decided to perform the analysis using the ActiGraph count system first. Before data analysis, data regarding non-wear time must be preprocessed. Therefore, the wear time values from the non-wear time protocols were compared with the calculated values of different non-wear time algorithms. In this study, the algorithms invented by Choi et al. (2011) and Troiano et al. (2008) were considered to determine wear time. The Choi algorithm using a 90-minute window ($\pm$30 minutes) for capturing non-wear time was found to be the most practical because there is no need to record for 24 hours and the other non-wear time algorithms with a 60 minute window found too many incorrectly classified non-wear times (Jaeschke et al., 2017). Furthermore, the non-wear time identified by Choi is independent of the used epoch length (Anedda et al., 2018). Different cut-points were calibrated for the classification of intensity of physical activity for different epoch lengths. The cut-off points used the most frequently were based on 1-second (Hänggi et al., 2013; Sasaki et al., 2011), on 15-second (Evenson et al., 2008; Romanzini et al., 2014), or on 60-second (Freedson et al., 2011; Freedson et al., 1998) epochs. In our analysis, different cut-off points for intensity classification will be used for different age groups because the age range of the study sample includes children, adolescents, and adults (6-27 years). The cut points from Evenson et al. (2008) for participants aged 6 to 8 years, Hänggi et al. (2013) for participants aged 9 to 11 years, Romanzini et al. (2014) for participants aged 12 to 19 years, and Sasaki et al. (2011) for adult participants are currently being considered. A summary of the accelerometer data processing criteria (suggested by Migueles et al. (2017)) can be found in Table 3.

Table 3.3: Accelerometer criteria

Accelerometer data processing criterion	Definition in this study
Placement of the device	Laterally on top of the right anterior superior iliac spine
Sampling frequency	30 hertz
Filter	Normal ActiGraph GT3X filter
Epoch lengths	1 second with possibility to convert to 5, 10, 15, 30, and 60 seconds
Non-wear time definition	Choi et al. (2011): 90-minute time window for consecutive zero/nonzero counts; allowance of 2-minute intervals of nonzero counts with an up/downstream 30-minute consecutive zero counts window
Valid days/valid weeks	8 hours of recording on at least four weekdays and one additional weekend day
Population age range	6-27 years (children, adolescents, and young adults)
Sedentary and physical activity intensity classification and cut point algorithms	6-10 years: Evenson et al. (2008) 11-17 years: Romanzini et al. (2014) Young adults: Sasaki et al. (2011)[a]

Note: List of accelerometer data processing criteria in KiGGS and MoMo (suggested by Migueles et al. (2017)). [a]To be determined; definition listed is under consideration.

3.4 Results

Data collection was completed in October 2017, and data harmonization was performed in 2018. The first accelerometer results from this wave were published in 2019 (Burchartz et al., 2019; Woll et al., 2019). Detailed analyzes are ready to be submitted in 2020.

First, data analysis should focus on gender and age differences in daily physical activity levels, as well as compliance with the physical activity recommendations of the World Health Organization (WHO). Furthermore, we plan to perform an in-depth analysis of the associations between physical activity and different health-related parameters (eg, obesity) and socioeconomic parameters (eg, education) by considering various data mining methods. This includes investigation of cross-links and trends between questionnaires and device-based collected activity data (eg, physical activity differences between groups in both data sets).

3.5 Discussion

3.5.1 Summary

Currently, there are different concepts of collecting and processing accelerometer data for the assessment of physical activity among children and adolescents (Migueles et al., 2017). Many studies do not provide detailed descriptions of their data collection and data handling processes. This complicates the replication of and comparability between studies. Therefore, the purpose of this study protocol was to explain the challenges faced when using accelerometers in the MoMo and KiGGS studies as an example of a large-scale epidemiological study and to detail the methods and protocols used to capture physical activity in children and adolescents with accelerometers in Germany.

3.5.2 Strengths

This study protocol provides an extensive list of considerations for measuring physical activity and sedentary behavior by accelerometry in a large sample. These include technical details of the device used and the reasoning behind the device choice, the reasoning behind the a priori data collection proceedings (assessment period and registration protocol, device initialization, device placement, non-wear time protocol), and the data processing methods. Furthermore, thoughts on feasibility issues (transfer and return of the devices, data download and preparation) are provided. Researchers planning similar studies are given all the information needed for replication. This enables comparability with other large European studies such as the European Youth Heart Study (Riddoch et al., 2004) and the Healthy Lifestyle in Europe by Nutrition in Adolescence (HELENA) Study (Ruiz et al., 2011; Vanhelst et al., 2012; Vicente-Rodriguez et al., 2007) due to similarities in methodology. A multimodal approach to the use of self-reports and accelerometers is recommended (Jekauc et al., 2014) and can combine the advantages of the different methods (eg precision and breadth of detection) because no single procedure provides optimal detection in all situations. The MoMo-PAQ was developed to measure habitual physical activity in general, while current physical activity was measured using accelerometers for 1 week. Combining these two methods of assessing physical activity provides the opportunity to present a more comprehensive picture of the actual participant's physical activity and can provide a basis for planning health-enhancing physical activity programs for specific target groups.

3.5.3 Challenges

Although ActiGraph accelerometers are used in many studies to record physical activity, there are technical issues associated with these devices; therefore, their limitations must be considered. The use of the normal ActiGraph filter removes signals with a frequency greater than 2.5 Hz. However, while performing vigorous physical activity, the human body produces accelerations at the hip up to a frequency of 3.4 Hz (Cavagna & Franzetti, 1986; Cavagna et al., 1991). Due to this limitation, activities with higher movement frequencies (that is, in the vigorous activity spectrum) may not be correctly assessed. In the context of MoMo and KiGGS, this will not be an issue, because all activities in this frequency range will be classified as vigorous and more detailed investigations are currently not planned. However, evaluating physical activity based on raw data is recommended for unbiased data processing that conforms to the open science approach. This requires more complex and advanced algorithms and evaluation methods. The first applied analysis will resort to a comparable evaluation with counts; however, future discussions on this topic are needed (Brønd & Arvidsson, 2016; Hill et al., 2015), and complex data analysis methods must be adapted. Although irregular and unstructured daily activities are recorded more accurately using accelerometers than questionnaires, there are still improvements to be made. Devices can only measure physical activity when they are worn.

Therefore, physical activity that occurs during the non-wear time is not included in the data sets. This creates a need for methods that include additional information from these non-wear times, such as a non-wear time protocol that adjusts for the missing physical activity.

Taking into account the wide range of ages in the sample, different cut-points and epoch lengths are suitable for the data in this study; however, calibration studies for such a broad sample do not exist. This leads to data accuracy issues when only one calibration study was used or to comparability issues within the sample when multiple calibration studies were used for different subsets of the sample.

3.5.4 Implications and Perspectives

For future waves of data collection, the non-wear time protocol should be improved. The frequency of reasons for non-wear must be analyzed so that the wearing instructions can be refined. This could potentially increase the wear time of devices. Furthermore, the non-wear time protocol should assess the intensities of physical activity more precisely during non-wear times. This would lead to a more complete assessment of all physical activity that occurs, and more detailed feedback could be given to participants. Future studies should examine the accuracy of different algorithms for detecting non-wear times for different age groups (Migueles et al., 2017). The impact of different thresholds for the classification of intensity of physical activity and the choice of the right epoch length for the target population based on age will be of interest as long as proprietary counts are used. Therefore, it is recommended to analyze multiple implemented cut-point algorithms and identify the one that best fits the sample at hand. Both methods of assessing physical activity should be compared between different target groups. Furthermore, the adherence to the recommendations of physical activity of the WHO should be examined.

3.5.5 Conclusion

This study protocol will help researchers obtain an overview of the decisions about the methods and protocols used to assess device-based physical activity in children and adolescents with accelerometers in Germany.

Authors' contributions

AB was responsible for the general conception and design of this manuscript. CN & DO supported the process of writing the manuscript. AB, CN, DO, SS, KM were responsible for the acquisition of the data. KM, BA, CN, DO, SS were responsible for the critical revision. CN & AW contributed to the design of the study. All authors read and approved the final manuscript.

References Study Protocol

Aadland, E., & Ylvisaker, E. (2015). Reliability of objectively measured sedentary time and physical activity in adults. *PLoS One*, *10*(7), e0133296. https://doi.org/10.1371/journal.pone.0133296

ActiGraph. (2013). GT3X+ and wGT3X+ Device Manual. http://actigraphcorp.com/wp-content/uploads/2015/11/GT3X-wGT3X-Device-Manual-110315.pdf

Anastasopoulou, P., Shammas, L., & Hey, S. (2012). Assessment of human gait speed and energy expenditure using a single triaxial accelerometer. *2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. https://doi.org/10.1109/BSN.2012.7

Anedda, B., Albrecht, C., Burchartz, A., Oriwol, D., Schmidt, S. C. E., & Woll, A. (2018). Processing accelerometer data of children and adolescents: Influence of wear time algorithm, epoch length and cut-off points on age specific subgroups. In M. H. Murphy, C. Boreham, G. de Vito, & E. Tsolakidis (Eds.), *European College of Sport Science* (p. 129). http://wp1191596.server-he.de/DATA/CONGRESSES/DUBLIN_2018/Documents/DUBLIN_BOA_WEB.pdf

Bornstein, D. B., Beets, M. W., Byun, W., & McIver, K. L. (2011). Accelerometer-derived physical activity levels of preschoolers: a meta-analysis. *Journal of Science and Medicine in Sport*, *14*(6), 504–511. https://doi.org/10.1016/j.jsams.2011.05.007

Brønd, J. C., & Arvidsson, D. (2016). Sampling frequency affects the processing of Actigraph raw acceleration data to activity counts. *Journal of applied physiology (Bethesda, Md. : 1985)*, *120*(3), 362–369. https://doi.org/10.1152/japplphysiol.00628.2015

Burchartz, A., Anedda, B., Oriwol, D., Kolb, S., Schmidt, S. C. E., & Woll, A. (2019). Erreichen Kinder und Jugendliche in Deutschland die WHO-Richtlinien für körperliche Aktivität: - erhoben durch Akzelerometer und Fragebogen. In A. Arampatzis, S. Braun, K. Schmitt, & B. Wolfarth (Eds.), *Sport im öffentlichen Raum* (p. 324). Feldhaus.

Cain, K. L., Sallis, J. F., Conway, T. L., van Dyck, D., & Calhoon, L. (2013). Using Accelerometers in Youth Physical Activity Studies: A Review of Methods. *Journal of Physical Activity and Health*, *10*(3), 437–450. https://doi.org/10.1123/jpah.10.3.437

Cavagna, G. A., & Franzetti, P. (1986). The determinants of the step frequency in walking in humans. *The Journal of physiology*, *373*(1), 235–242. https://physoc.onlinelibrary.wiley.com/doi/pdf/10.1113/jphysiol.1986.sp016044

Cavagna, G. A., Willems, P. A., Franzetti, P., & Detrembleur, C. (1991). The two power limits conditioning step frequency in human running. *The Journal of physiology*, *437*(1), 95–108. https://physoc.onlinelibrary.wiley.com/doi/pdf/10.1113/jphysiol.1991.sp018586

Choi, L., Liu, Z., Matthews, C. E., & Buchowski, M. S. (2011). Validation of accelerometer wear and nonwear time classification algorithm. *Medicine and science in sports and exercise*, *43*(2), 357–364. https://doi.org/10.1249/MSS.0b013e3181ed61a3

Colley, R. C., Garriguet, D., Janssen, I., Craig, C. L., Clarke, J., & Tremblay, M. S. (2011). Physical activity of Canadian adults: accelerometer results from the 2007 to 2009 Canadian Health Measures Survey. *Health Reports*, *22*(1), 7–14.

Crouter, S. E., Flynn, J. I., & Bassett Jr, D. R. (2015). Estimating physical activity in youth using a wrist accelerometer. *Medicine and science in sports and exercise*, *47*(5), 944. https://doi.org/10.1249/MSS.0000000000000502

Donaldson, S. C., Montoye, A. H. K., Tuttle, M. S., & Kaminsky, L. A. (2016). Variability of Objectively Measured Sedentary Behavior. *Medicine and science in sports and exercise*, *48*(4), 755–761. https://doi.org/10.1249/MSS.0000000000000828

Edwardson, C. L., & Gorely, T. (2010). Epoch length and its effect on physical activity intensity. *Medicine and science in sports and exercise*, *42*(5), 928–934. https://doi.org/10.1249/MSS.0b013e3181c301f5

Ekelund, U., Yngve, A., Brage, S., Westerterp, K. R., & Sjöström, M. (2004). Body movement and physical activity energy expenditure in children and adolescents: How to adjust for differences in body size and age. *The American journal of clinical nutrition*, *79*(5), 851–856. https://doi.org/10.1093/ajcn/79.5.851

Ettienne, R., Nigg, C. R., Li, F., Su, Y., McGlone, K., Luick, B., Tachibana, A., Carran, C., Mercado, J., & Novotny, R. (2016). Validation of the Actical Accelerometer in Multiethnic Preschoolers: The Children's Healthy Living (CHL) Program. *Hawai'i journal of medicine & public health : a journal of Asia Pacific Medicine & Public Health*, *75*(4), 95–100.

Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S., & McMurray, R. G. (2008). Calibration of two objective measures of physical activity for children. *Journal of sports sciences*, *26*(14), 1557–1565. https://doi.org/10.1080/02640410802334196

Fairclough, S. J., Noonan, R., Rowlands, A. V., van Hees, V. T., Knowles, Z., & Boddy, L. M. (2016). Wear compliance and activity in children wearing wrist and hip mounted accelerometers. *Medicine & Science in Sports & Exercise*, *48*(2), 245–253. https://doi.org/10.1249/MSS.0000000000000771

Freedson, P. S., Lyden, K., Kozey-Keadle, S., & Staudenmayer, J. W. (2011). Evaluation of artificial neural network algorithms for predicting METs and activity type from accelerometer data: Validation on an independent sample. *Journal of Applied Physiology*, *111*(6), 1804–1812. https://doi.org/10.1152/japplphysiol.00309.2011

Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and science in sports and exercise*, *30*(5), 777–781. https://doi.org/10.1097/00005768-199805000-00021

Gabrys, L., Thiel, C., Tallner, A., Wilms, B., Müller, C., Kahlert, D., Jekauc, D., Frick, F., Schulz, H., Sprengeler, O., Hey, S., Kobel, S., & Vogt, L. (2015). Akzelerometrie zur Erfassung körperlicher Aktivität. *Sportwissenschaft*, *45*(1), 1–9. https://doi.org/10.1007/s12662-014-0349-5

Graf, C., Beneke, R., Bloch, W., Bucksch, J., Dordel, S., Eiser, S., Ferrari, N., Koch, B., Krug, S., Lawrenz, W., Manz, K., Naul, R., Oberhoffer, R., Quilling, E., Schulz, H., Stemper, T., Stibbe, G., Tokarski, W., Völker, K., & Woll, A. (2013). Vorschläge zur Förderung der körperlichen Aktivität von Kindern und

Jugendlichen in Deutschland. *Monatsschrift Kinderheilkunde*, *161*(5), 439–446. https://doi.org/10.1007/s00112-012-2863-6

Hänggi, J. M., Phillips, L. R. S., & Rowlands, A. V. (2013). Validation of the GT3X ActiGraph in children and comparison with the GT1M ActiGraph. *Journal of Science and Medicine in Sport*, *16*(1), 40–44. https://doi.org/10.1016/j.jsams.2012.05.012

Herrmann, S. D., Barreira, T. V., Kang, M., & Ainsworth, B. E. (2013). How many hours are enough? Accelerometer wear time may provide bias in daily activity estimates. *Journal of Physical Activity and Health*, *10*(5), 742–749. https://doi.org/10.1123/jpah.10.5.742

Hildebrand, M., van Hees, V. T., Hansen, B. H., & Ekelund, U. (2014). Age group comparability of raw accelerometer output from wrist-and hip-worn monitors. *Medicine & Science in Sports & Exercise*, *46*(9), 1816–1824. https://doi.org/10.1249/MSS.0000000000000289

Hill, H., Ottenbacher, J., & Ebner-Priemer, U. W. (Eds.). (2015). *What counts count... A comparison of count and movement acceleration intensity data in the assessment of physical activity*.

Jaeschke, L., Luzak, A., Steinbrecher, A., Jeran, S., Ferland, M., Linkohr, B., Schulz, H., & Pischon, T. (2017). 24 h-accelerometry in epidemiological studies: Automated detection of non-wear time in comparison to diary information. *Scientific Reports*, *7*(1), 2227. https://doi.org/10.1038/s41598-017-01092-w

Jekauc, D., Reimers, A. K., & Woll, A. (2014). Methoden der Aktivitätsmessung bei Kindern und Jugendlichen. *B&G Bewegungstherapie und Gesundheitssport*, *30*(02), 79–82. https://doi.org/10.1055/s-0033-1361578

Jekauc, D., Wagner, M. O., Kahlert, D., & Woll, A. (2013). Reliabilität und Validität des MoMo-Aktivitätsfragebogens für Jugendliche (MoMo-AFB). *Diagnostica*, *59*(2), 100–111. https://doi.org/10.1026/0012-1924/a000083

Kurth, B.-M., Kamtsiuris, P., Hölling, H., Schlaud, M., Dölle, R., Ellert, U., Kahl, H., Knopf, H., Lange, M., Mensink, G. B. M., Neuhauser, H., Rosario, A. S., Scheidt-Nave, C., Schenk, L., Schlack, R., Stolzenberg, H., Thamm, M., Thierfelder, W., & Wolf, U. (2008). The challenge of comprehensively mapping children's health in a nation-wide health survey: Design of the German KiGGS-Study. *BMC public health*, *8*, 196. https://doi.org/10.1186/1471-2458-8-196

Mâsse, L. C., Fuemmeler, B. F., Anderson, C. B., Matthews, C. E., Trost, S. G., Catellier, D. J., & Treuth, M. S. (2005). Accelerometer data reduction: a comparison of four reduction algorithms on select outcome variables. *Medicine and science in sports and exercise*, *37*(11 Suppl), S544–54. https://doi.org/10.1249/01.mss.0000185674.09066.8a

Matthews, C. E., Ainsworth, B. E., Thompson, R. W., & Bassett, D. R. (2002). Sources of variance in daily physical activity levels as measured by an accelerometer. *Medicine and science in sports and exercise*, *34*(8), 1376–1381. https://doi.org/10.1097/00005768-200208000-00021

Mauz, E., Gößwald, A., Kamtsiuris, P., Hoffmann, R., Lange, M., von Schenck, U., Allen, J., Butschalowsky, H., Frank, L., Hölling, H., et al. (2017). Neue Daten für Taten. Die Datenerhebung zur KiGGS Welle 2 ist beendet. *Journal of Health Monitoring*, *S3*. https://doi.org/10.17886/RKI-GBE-2017-099

Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Delisle Nyström, C., Mora-Gonzalez, J., Löf, M., Labayen, I., Ruiz, J. R., & Ortega, F. B. (2017). Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other

Outcomes: A Systematic Review and Practical Considerations. *Sports medicine (Auckland, N.Z.)*, *47*(9), 1821–1845. https://doi.org/10.1007/s40279-017-0716-0

Müller, C., Winter, C., Mogwitz, M.-S., & Rosenbaum, D. (2011). Validierung von neun Bewegungssensoren bei Kindern und jungen Erwachsenen. *Sportwissenschaft*, *41*(1), 8–15. https://doi.org/10.1007/s12662-011-0167-y

Riddoch, C. J., Bo Andersen, L., Wedderkopp, N., Harro, M., Klasson-Heggebø, L., Sardinha, L. B., Cooper, A. R., & Ekelund, U. (2004). Physical activity levels and patterns of 9- and 15-yr-old European children. *Medicine and science in sports and exercise*, *36*(1), 86–92. https://doi.org/10.1249/01.MSS.0000106174.43932.92

Romanzini, M., Petroski, E. L., Ohara, D., Dourado, A. C., & Reichert, F. F. (2014). Calibration of ActiGraph GT3X, Actical and RT3 accelerometers in adolescents. *European journal of sport science*, *14*(1), 91–99. https://doi.org/10.1080/17461391.2012.732614

Rowlands, A. V. (2007). Accelerometer Assessment of Physical Activity in Children: An Update. *Pediatric Exercise Science*, *19*(3), 252–266. https://doi.org/10.1123/pes.19.3.252

Ruiz, J. R., Ortega, F. B., Martínez-Gómez, D., Labayen, I., Moreno Aznar, L. A., de Bourdeaudhuij, I., Manios, Y., González-Gross, M., Mauro, B., Molnar, D., Widhalm, K., Marcos, A., Béghin, L., Castillo, M. J., & Sjöström, M. (2011). Objectively measured physical activity and sedentary time in European adolescents: The HELENA study. *American journal of epidemiology*, *174*(2), 173–184. https://doi.org/10.1093/aje/kwr068

Sallis, J. F., & Saelens, B. E. (2000). Assessment of physical activity by self-report: status, limitations, and future directions. *Research quarterly for exercise and sport*, *71 Suppl 2*, 1–14. https://doi.org/10.1080/02701367.2000.11082780

Sasaki, J. E., John, D., & Freedson, P. S. (2011). Validation and comparison of ActiGraph activity monitors. *Journal of Science and Medicine in Sport*, *14*(5), 411–416. https://doi.org/10.1016/j.jsams.2011.04.003

Schaefer, C. A., Nigg, C. R., Hill, J. O., Brink, L. A., & Browning, R. C. (2014). Establishing and evaluating wrist cutpoints for the GENEActiv accelerometer in youth. *Medicine and science in sports and exercise*, *46*(4), 826–833. https://doi.org/10.1249/MSS.0000000000000150

Sherar, L. B., Griew, P., Esliger, D. W., Cooper, A. R., Ekelund, U., Judge, K., & Riddoch, C. J. (2011). International children's accelerometry database (ICAD): Design and methods. *BMC public health*, *11*, 485. https://doi.org/10.1186/1471-2458-11-485

Slootmaker, S. M., Schuit, A. J., Chinapaw, M. J., Seidell, J. C., & van Mechelen, W. (2009). Disagreement in physical activity assessed by accelerometer and self-report in subgroups of age, gender, education and weight status. *The international journal of behavioral nutrition and physical activity*, *6*, 17. https://doi.org/10.1186/1479-5868-6-17

Stec, M. J., & Rawson, E. S. (2012). Estimation of resistance exercise energy expenditure using triaxial accelerometry. *The Journal of Strength & Conditioning Research*, *26*(5), 1413–1422. https://doi.org/10.1519/JSC.0b013e318248d7b4

Toftager, M., Kristensen, P. L., Oliver, M., Duncan, S., Christiansen, L. B., Boyle, E., Brønd, J. C., & Troelsen, J. (2013). Accelerometer data reduction in adolescents: effects on sample retention and bias. *International Journal of Behavioral Nutrition and Physical Activity*, *10*, 140. https://doi.org/10.1186/1479-5868-10-140

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine and science in sports and exercise*, *40*(1), 181–188. https://doi.org/10.1249/mss.0b013e31815a51b3

Troiano, R. P., McClain, J. J., Brychta, R. J., & Chen, K. Y. (2014). Evolution of accelerometer methods for physical activity research. *British journal of sports medicine*, *48*(13), 1019–1023. https://doi.org/10.1136/bjsports-2014-093546

Trost, S. G., McIver, K. L., & Pate, R. R. (2005). Conducting accelerometer-based activity assessments in field-based research. *Medicine and science in sports and exercise*, *37*(11 Suppl), S531–43. https://doi.org/10.1249/01.mss.0000185657.86065.98

Tryon, W. W., & Williams, R. (1996). Fully proportional actigraphy: A new instrument. *Behavior Research Methods*, *28*(3), 392–403. https://doi.org/10.3758/BF03200519

Tudor-Locke, C. E., Ainsworth, B. E., Adair, L. S., & Popkin, B. M. (2003). Physical activity in Filipino youth: The Cebu Longitudinal Health and Nutrition Survey. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity*, *27*(2), 181–190. https://doi.org/10.1038/sj.ijo.802207

van Hees, V. T., Fang, Z., Mirkes, E. M., Heywood, J., Zhao, J. H., Joan, C. P., Sabia, S., & Migueles, J. H. (2019). GGIR. https://doi.org/10.5281/ZENODO.1051064

Vanhelst, J., Béghin, L., Duhamel, A., Bergman, P., Sjöström, M., & Gottrand, F. (2012). Comparison of uniaxial and triaxial accelerometry in the assessment of physical activity among adolescents under free-living conditions: The HELENA study. *BMC medical research methodology*, *12*(1), 26. https://doi.org/10.1186/1471-2288-12-26

Vicente-Rodriguez, G., Libersa, C., Mesana, M. I., Béghin, L., Iliescu, C., Moreno Aznar, L. A., Dallongeville, J., & Gottrand, F. (2007). Healthy lifestyle by nutrition in adolescence (HELENA). A new EU funded project. *Therapie*, *62*(3), 259–270. https://doi.org/10.2515/therapie:2007050

Wagner, M. O., Bös, K., Jekauc, D., Karger, C., Mewes, N., Oberger, J., Reimers, A. K., Schlenker, L., Worth, A., & Woll, A. (2014). Cohort profile: The Motorik-Modul Longitudinal Study: physical fitness and physical activity as determinants of health development in German children and adolescents. *International journal of epidemiology*, *43*(5), 1410–1416. https://doi.org/10.1093/ije/dyt098

Wanner, M., Richard, A., Martin, B., Faeh, D., & Rohrmann, S. (2017). Associations between self-reported and objectively measured physical activity, sedentary behavior and overweight/obesity in NHANES 2003-2006. *International journal of obesity (2005)*, *41*(1), 186–193. https://doi.org/10.1038/ijo.2016.168

Warren, J. M., Ekelund, U., Besson, H., Mezzani, A., Geladas, N., & Vanhees, L. (2010). Assessment of physical activity–a review of methodologies with reference to epidemiological research: A report of the exercise physiology section of the European Association of Cardiovascular Prevention and Rehabilitation. *European Journal of Cardiovascular Prevention & Rehabilitation*, *17*(2), 127–139. https://doi.org/10.1097/HJR.0b013e32832ed875

Wijndaele, K., Westgate, K., Stephens, S. K., Blair, S. N., Bull, F. C., Chastin, S. F. M., Dunstan, D. W., Ekelund, U., Esliger, D. W., Freedson, P. S., Granat, M. H., Matthews, C. E., Owen, N., Rowlands, A. V., Sherar, L. B., Tremblay, M. S., Troiano, R. P., Brage, S., & Healy, G. N. (2015). Utilization and Harmonization of

Adult Accelerometry Data: Review and Expert Consensus. *Medicine and science in sports and exercise*, *47*(10), 2129–2139. https://doi.org/10.1249/MSS.0000000000000661.

Woll, A., Albrecht, C., & Worth, A. (2017). Motorik-Module (MoMo) – the KiGGS wave 2 module to survey motor performance and physical activity. *Journal of Health Monitoring*, *2*(S3), 66–73. https://doi.org/10.17886/RKI-GBE-2017-110

Woll, A., Oriwol, D., Anedda, B., Burchartz, A., Hanssen-Doose, A., Kopp, M., Niessner, C., Schmidt, S. C. E., Bös, K., & Worth, A. (2019). Körperliche Aktivität, motorische Leistungsfähigkeit und Gesundheit in Deutschland: Ergebnisse aus der Motorik-Modul-Längsschnittstudie (MoMo). https://doi.org/10.5445/IR/1000095369

Worth, A., Woll, A., Albrecht, C., Karger, C., Mewes, N., Oberger, J., Schlenker, L., Schmidt, S. C. E., Wagner, M. O., & Bös, K. (Eds.). (2015). MoMo-Längsschnittstudie "Physical Fitness and Physical Activity as Determinants of Health Development in Children and Adolescents": Testmanual zu den motorischen Tests und den anthropometrischen Messungen (KIT Scientific Reports ; 7700). https://doi.org/10.5445/ksp/1000047434

4. Processing accelerometer data of children and adolescents in MoMo

Slightly modified version of the 5^{th} published article

Burchartz, Alexander; Kolb, S.; Klos, Leon.; Schmidt, S. C.E.; Haaren-Mack, B. von; Niessner, C.; & Woll, A. (2023): How specific combinations of epoch length, non-wear time and cut-points influence physical activity - Processing accelerometer data from children and adolescents in the nationwide MoMo study.

German Journal of Exercise and Sport Research, Advance online publication.
DOI: 10.1007/s12662-023-00892-9

4.1 abstract

This study assesses three factors that influence the quantification of children's and adolescents' physical activity (PA) using accelerometers: Selection of 1) non-wear algorithm, 2) epoch length and 3) cut-points. 1,525 participants from MoMo Wave 3 (2018-2022), aged 6 to 17 years, wore ActiGraph GT3X accelerometers during waking hours. Acceleration counts were reintegrated into lengths of 1, 5, 15, 30, and 60 s epochs. Two non-wear time algorithms and two sets of cut-points were applied to each epoch length. Significant differences ($p < .01$) were found in both the comparison of the non-wear time algorithms and the comparison of the cut-points when the different epoch lengths were considered. This may result in large differences in estimated sedentary behavior and PA values. We propose to pool the data by merging and combining multiple accelerometer datasets from different studies and evaluate them in a harmonized way in the future. In addition to the need for future validation studies using short epoch lengths for young children, we also propose to conduct meta-analyses. This allows using data from multiple studies to validate cut-off points and to propose a consensual set of cut-off points that can be used in different settings and projects. The high discrepancy between results when comparing different epoch lengths has to be considered when interpreting

accelerometer data and is regarded a confounding variable when comparing levels of PA between studies.

Key Words: guidelines, physical behavior, data analysis, youth, vertical axis

4.2 Background

When measuring physical activity (PA) with devices, especially accelerometers, the objective is to detect it by measuring the acceleration occurring at a specific point on the human body (Burchartz, Anedda, et al., 2020). To report physical behavior (PB) in public health settings, epidemiological studies use times spent at certain levels of activity intensity throughout the day and the week as result. Usually, these intensities are divided into sedentary behavior (SED), light (LPA), moderate (MPA), and vigorous (VPA) intensity PA, although there are few more complex evaluation possibilities (Pfeiffer et al., 2022). To convert the captured accelerations into these intensity levels, the collected data needs to be preprocessed. Many parameters like device selection, carrying position and recording parameters such as recording frequency or filters influence how the measurement is transformed into the final results and largely affect the following estimation of activity (Burchartz, Anedda, et al., 2020; Rowlands et al., 2018). Understanding the PA behavior of children and adolescents is severely limited by the presence of many sets of intensity-based limits for different brands and models of accelerometers. Trost (2007) used the term "cut-point conundrum" already in the early days of device-based assessment of PA to describe the bewildering number of calibration studies for different individual populations with a wide variety of devices and settings that have been published without a generalized and standardized approach, which has made a comparison between individual studies difficult, if not impossible. Today, there are even more approaches to analyze accelerometer data, and for these reasons, it is important to document all decisions made in recording and processing the data to allow later comparison with other studies (Arvidsson et al., 2019; Burchartz, Anedda, et al., 2020; Migueles et al., 2017). In an earlier review, Cain et al. (2013) reported that only about half of all studies report all decisions used during the process of data processing. In a recent review analyzing the approach of accessibility and use of novel analytic techniques for classifying physical activity intensity using raw or count-based accelerometer data, Pfeiffer et al. (2022) found that less than half of the models developed between 2000 and 2021 are not even publicly available. Therefore, it is not surprising that so many models are not used by other researchers.

One of the first things to look at when preprocessing the data is the accuracy of different algorithms for detecting the time the accelerometer is not worn, the so-called non-wear time (NWT) (Migueles et al., 2017). However, determining NWT has received little attention in the literature, although it sometimes accounts for a large portion of the activity data in the 24-hour activity behavior cycle (Syed et al., 2020). To evaluate compliance with the study, determining the time during which the accelerometer was not worn is very important. For an accurate calculation of summary statistics, such as minutes spent in SED or LPA, MPA, and VPA, NWT has a profound impact (Syed et al., 2020). Second, the effect of different thresholds for PA intensity classification and of choosing the right epoch length for the target population based on age is of high interest as long as proprietary counts are used (Dencker et al., 2012), since classifying activity is done by accumulating counts in a specific time interval, i.e., the epoch length. Some studies Banda et al. (2016), Bornstein et al. (2011), Breau et al. (2022), Dias Moura et al. (2019), Leppänen et al. (2022), Logan et al. (2016), Migueles et al. (2019), and Xing et al. (2021) already did compare the effects of different cut points for children in the estimates of physical behavior. However, to our knowledge, the specific combination of algorithms used in MoMo have not yet been compared as a collective. Migueles et al. (2019) provide a comprehensive comparison of different cut-point algorithms in overweight children. In doing so, they demonstrate that it is currently not and probably never will be possible to determine the prevalence of meeting PA guidelines based on accelerometer data. This is based on the fact that the apparent differences found range from almost zero to almost all participants meeting the guidelines, depending on the algorithm used for the evaluation. Giurgiu et al. (2022) found that most validation studies did not meet recommended quality principles when performing validation of wearables in real-world conditions and therefore recommend using standardized protocols to document all technical decisions for reproducibility. This is another reason why it is important that this point of the analysis in MoMo be adequately documented. Ultimately, the comparability of results from different studies in Germany stands or falls with the algorithms used to analyze the data.

Many cohort studies use sensors or generations of software that have been replaced by newer versions over time, making it difficult to draw conclusions about changes in current technology. In the early days of accelerometry, only vertical acceleration was measured by the devices. Nowadays, sensors are much smaller and cheaper than earlier devices and thus it has become state of the art to measure acceleration in all three axes. Although this enables for even more accurate measurement, it also comes with new challenges. Newer cut-point intensity classification based on vector magnitude (the magnitude of a three-dimensional vector as the length of the entire acceleration, respectively, the movement in all three axes) produces higher MVPA time compared to estimations based on older ENMO (Euclidean Norm Minus One, see Bakrania et al. (2016)) or vertical axis algorithms, independent of the place of attachment (Migueles et al., 2019).

As a consequence of the factors described above that prevented reasonable comparability of previous studies, it is now recommended to analyze multiple cut-point algorithms and identify the one where the validation sample best fits the target population at hand (Breau et al., 2022; Burchartz, Anedda, et al., 2020; Migueles et al., 2017; Syed et al., 2020). This means that from the large number of available validation studies, one selects the study or algorithm that best fits one's own study. It is important that the following criteria are particularly similar to the validation study: Age frame, gender, device, location, recording frequency, filter, epoch length, valid days. By capturing accelerometer data from 2014 to 2020, the nationwide Motorik-Modul study (MoMo) collected representative data on the physical behavior of children and adolescents in Germany in two waves (MoMo Wave 2 and MoMo Wave 3). The objective of this study is to transparently examine the acceleration data of MoMo wave 3 (2018-2022) under the various aspects of epoch lengths, NWT, and cut point sets as a combined overall construct.. In particular, the influence of five different epoch lengths on two different NWT algorithms was examined, as well as the differences when interpreting intensity classifications by two sets of cut-points for different age groups. The results will help German researchers to understand how to set the appropriate parameters and for what reasons when evaluating their own accelerometer data, in order to be able to subsequently compare their data with the results of MoMo.

4.3 Methods

4.3.1 Study design

The German Health Interview and Examination Survey for Children and Adolescents (KiGGS) is part of the Federal Health Monitoring System conducted by the Robert Koch-Institute (RKI) and consists of regularly conducted nationwide surveys among children, adolescents, and young adults aged 0 to 29 years and living in Germany since 2003. MoMo is a submodule of the KiGGS study and aims to assess physical fitness, PA, and determinants of PA in children and adolescents (Woll et al., 2017). The entire study sample was drawn from the German resident population aged 4 to 17 years using a two-stage cluster sampling approach. Furthermore, participants from the baseline study (2003-2006), wave 1 (2009–2012), and wave 2 (2015-2017) were invited back for wave 3 (2018-2021). A detailed description of the study design and sampling procedure can be found elsewhere (Hoffmann et al., 2018; Mauz et al., 2019; Woll et al., 2017). KiGGS and MoMo provide nationally representative data on PA and SED of children, adolescents, and young adults living in Germany. A positive vote of the Ethics Committee of the Karlsruhe Institute of Technology on 23 September 2014 is available for the study.

4.3.2 Sample description

For the current analysis, only cross-sectional data from participants aged 6 to 17 years from MoMo Wave 3 (2018–2022) were used (n = 1,525). All data used in this study had been collected before the first Covid-19-related lockdown in March 2020. The participation in the study was voluntary and the guardians of the participants gave their written consent.

4.3.3 Device-based, measured PA data

The technical and methodological details of the present study have previously been published (Burchartz, Manz, et al., 2020; Burchartz et al., 2021). In summary, for the assessment of device-based measured PA, ActiGraph GT3X+ and wGT3X-BT accelerometers (Actigraph, LLC, Pensacola, FL, USA) were used for eight consecutive days. Participants were instructed to place the accelerometer on the right hip and wear it continuously, except during water activities or while sleeping. Data was sampled at a frequency of 30 Hz, downloaded as ActiLife GT3X raw device files, and stored for offline analysis. The GT3X files were then accumulated in ActiGraph count-based AGD files with a 1s epoch length and converted to a Matlab-readable format.
For the present study, we analyzed all accelerometry data using the Matlab software version R2017a (The MathWorks Inc., Natick, Massachusetts, USA) to automate the data processing workflow without having to rely on the ActiLife 6.13.4 graphical user interface software (Actigraph, LLC, Pensacola, FL, USA). The recordings of the first day were not considered for data analysis because the participants received the devices at different times throughout the day. Additionally, the first day served as an adaptation period for the participants. In total, data for 10,557 days were recorded in n = 1,525 participants, with data available for all 7 days in 97.2% of the sample. Data was analyzed for individual days to investigate the effect of the parameters for example on wear time in the context of the valid day criteria. The analysis focuses on methodological differences rather than test subjects' results. This allows to examine how different the results can be when evaluating them using different methods.

The epoch length data set of 1s was therefore reintegrated into another four data sets with epoch length of 5s, 15, 30s, and 60s, respectively. Afterward, each of these five epoch length datasets were analyzed by specifying the two different non-wear criteria of Troiano et al. (2008) and Choi et al. (2011) resulting in 10 different configurations. Those criteria were chosen since in the standard software ActiLife offers only three wear time validation options to users: Troiano, Choi and a daily/hourly algorithm by ActiGraph (ActiGraph, 2020). Thus, for the inexperienced user there is only the possibility to choose from these two algorithms. Two vertical-axis cut-point algorithms for classification into different activity classes (Evenson et al. (2008) and Romanzini et al. (2014), Table 4.1) were then applied to each of these 10 epoch length/NWT datasets, resulting in 20 final datasets including all possible combinations of the specific epoch lengths, NWT and cut points. This decision is oriented to the ICAD specifications and sample specifications. Evenson is also used in ICAD as an evaluation routine, also only the variant with vertical axis. However, our age range is larger than that of Evenson (validated with 5-9 years). That is why Romanzini was used for older children in MoMo (validated with 10-15 years, recommended also from 12-19 years by Migueles et al. (2017)).

Table 4.1: Count cutoff ranges for one-second epochs of vertical-axis intensity algorithms.

Count cutoffs	Calibration Population	SED	LPA	MPA	VPA
Evenson et al. (2008)	5-9 years	≤ 2	>2 to ≤ 38	>38 to ≤ 67	>67
Romanzini et al. (2014)	10-15 years	≤ 3	>2 to ≤ 40	>40 to ≤ 54	>54

Note: **SED** Sedentary behavior, **LPA** light physical activity, **MPA** moderate physical activity, **VPA** vigorous physical activity

The Troiano et al. (2008), Choi et al. (2011), Evenson et al. (2008), and Romanzini et al. (2014) algorithms were implemented in Matlab programming language according to their published descriptions and, where available, the published code (Choi et al., 2011). The Troiano et al. (2008) and Choi et al. (2011) implementations were validated against their implementations in the ActiLife Software. Cut-point limits of the Evenson et al. (2008) and Romanzini et al. (2014) algorithms were independently verified by two authors.

4.3.4 Statistics

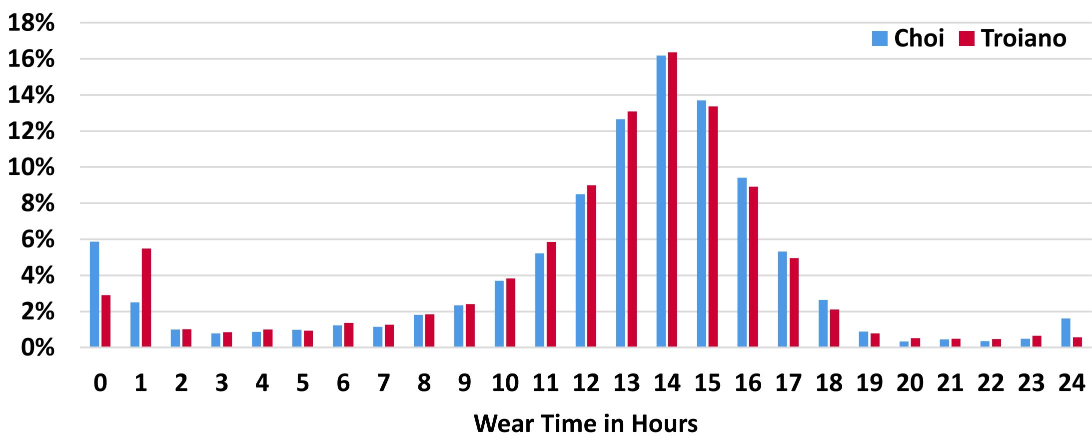
All statistical tests were conducted using IBM SPSS 28 (IBM Corporation, Armonk, NY, USA). Descriptive analyzes, stratified by NWT algorithms, EL, and cut-point algorithms, were performed and means (M), standard deviations (SD), and percentages were reported, as well as the corresponding inference statistical parameters, including 95 % confidence intervals for differences (CI). A one-way analysis of variance with repeated measurements (ANOVA) was used to analyze differences between groups of different epoch lengths and algorithms. The P values were adjusted using the Bonferroni method for multiple comparisons. The Games-Howell post hoc test was used. Additionally, the differences between days regarding the two cut-point algorithms were calculated and t-tests were used to analyze differences between intensity classifications. For all analyses, the level of significance was set to $\alpha = 0.01$.

4.4 Results

4.4.1 Differences in NWT

Regarding the wear time for each day of the week, there were no differences between the algorithms except in the tenth range (Figure 4.1, Supplement Table A.1). While applying the Choi algorithm revealed more full-day NWT, applying the Troiano algorithm resulted in more NWT classification at less than 1 hour. For both algorithms, the results indicated a wear time between 8 and 16 hours (Choi = 73.5 %, Troiano = 74.6 %) for almost three-quarters of the days. Wear times of 0-2 hours were classified by both algorithms for only 9.4 % of the days of the week. However, descriptive analysis of the NWT algorithms revealed that the mean wear time of Choi was independently constant in all epoch lengths and evaluation algorithms (M_C=825.01, SD_C=165.26). For Troiano, however, different epoch lengths had a significant influence on the result. Longer epoch length resulted in more NWT and therefore less valid days (M_{T1s}=822.48, S_{T1s}=162.38; M_{T5s}=819.54, S_{T5s}=160.16; M_{T15s}=813.80, S_{T15s}=155.43; M_{T30s}=806.42, S_{T30s}=149.36; M_{T60s}=805.55, S_{T60s}=145.58). The homogeneity of the variances was asserted using the Levene test, which showed that equal variances could be assumed for Choi ($p_C > .05$) but not for Troiano ($p_T < .001$). Therefore, the wear time did not differ for the different epoch length for Choi ($F_{4,88495} = 0.00$, $p > .05$), but for Troiano (Welch's $F_{4,43987.00} = 42.52$, $p < .001$) (Figure 4.2). Games-Howell's post-hoc analysis revealed a significant difference ($p < .001$) between all epoch length groups (except for 1s to 5s and 30s to 60s) for Troiano. The mean wear time decreased from 5 to 1s (-2.94, 95% CI [-7.62, 1.74], $p >$.05), from 15 to 5s (-5.74, 95% CI [-10.32, -1.15], $p <$.05), from 30 to 15s (-7.38, 95% CI [-11.82, -2.95], $p <$.05), and from 60 to 15s (-0.86, 95% CI [-5.16, -.34], $p >$.05).

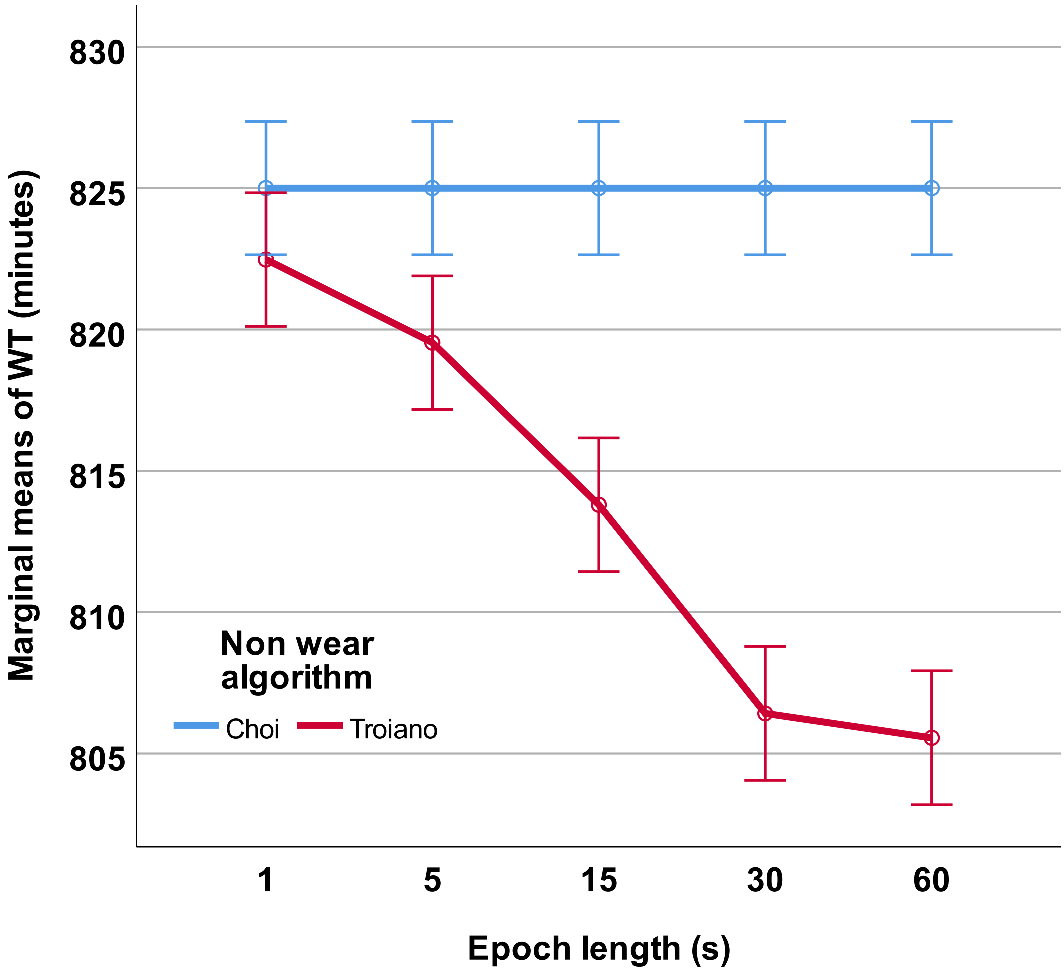
Figure 4.1: Histogram of wear time per hour, for both NWT algorithms in % of datasets.



4.4.2 Differences in valid days by valid hours of WT

The proportion of valid days in the data set depending on the criterion of valid hours required decreases rapidly above 10 h of wear time (Table 4.2). For the analysis of the two cut-point algorithms, only the Choi data set with more than 8h of weartime was used . Differences in mean minutes per day were analyzed for all cut-point levels of both algorithms and significant differences were found between the five epoch lengths (Figure 3). For SED, a 196 min (29.5 %) and a 152 min (22.4 %) decrease were observed

Figure 4.2: Marginal means of wear time for different epoch lengths and non-wear time algorithms (with 95% CI).



Note: Analysis for days with a wear time of more than 8 hours.

between the 1s and 60s epochs for the cut-points of Evenson et al. (2008) and Romanzini et al. (2014), respectively (Figure 4.3A). Significant differences ($p < .01$) in sedentary time were observed between each activity output for all epoch length. A decrease in SED time occurred as epoch length increased for both cut-points. For LPA, an increase of 213 min (216.3 %) and a decrease of 170 min (185.3 %) were observed between the 1st and 60s epochs for Evenson et al. (2008) and Romanzini et al. (2014) Cut-points, respectively (Figure 4.3B). Significant differences ($p < .01$) in LPA time were observed between each activity output for all epoch lengths. An increase in LPA time occurred as the epoch length increased for both cut-points. An increase of 0.6 min (1.8 %) and 1 min (7.9 %) was observed between the 1 and 60s epochs using Evenson et al. (2008) and Romanzini et al. (2014) cut-points, respectively (Figure 4.3C). There was an overall decrease in MPA between the 15 and 1s as well as between the 15 and 60s epochs. The decrease was 4 min (12.2 %) and 3 min (10.6 %) for Evenson et al. (2008) and 2 min (14.4 %) and 1 min (7.7 %) for the cut-points of Romanzini et al. (2014), respectively (Figure 4.3C). Significant differences ($p < .01$) in MPA time were observed between each activity output for all epoch length except for 5s Evenson et al. (2008) with 30s and 5s Romanzini et al. (2014) with 60s ($p > .05$). For VPA, a 17 min (58.9 %) and

Table 4.2: Valid days depending on valid hour criterion, epoch length, and non-wear time algorithm

NWT	EL	8h	10h	12h	13h	14h	15h	16h
Choi	**1s**							
	5s							
	15s	83.83%*	77.84%*	64.26%*	51.62%*	35.43%*	21.76%*	12.17%*
	30s							
	60s							
Troiano	**1s**	83.77%	77.83%	63.74%	51.11%	34.63%	21.06%	11.85%
	5s	83.60%	77.56%	63.28%	50.41%	34.03%	20.52%	11.40%
	15s	83.30%	77.10%	62.19%	49.20%	32.84%	19.48%	10.63%
	30s	83.01%	76.61%	61.12%	47.61%	31.33%	18.20%	9.61%
	60s	83.12%	76.72%	61.50%	48.04%	31.67%	18.44%	9.48%

Note: *The values of the different epoch lengths are uniform for the Choi algorithm within each valid hour criterion. **NWT** non-wear time algorithm, **EL** epoch length in s, % of n=21,114 valid days overall

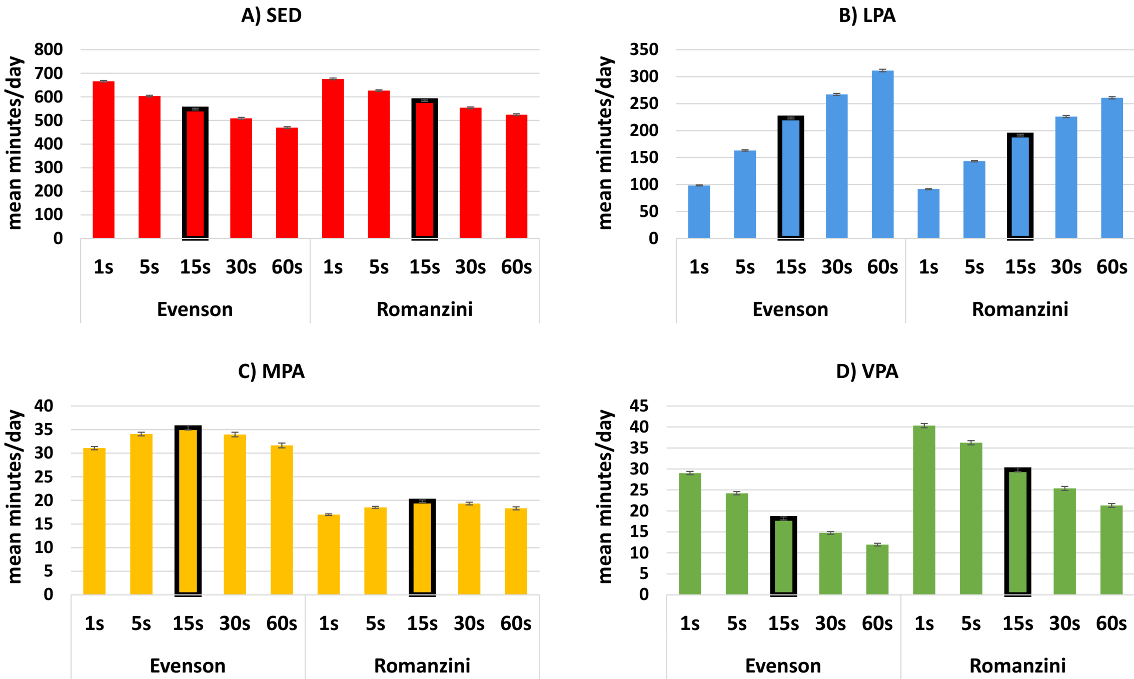
a 19 min (47.2%) decrease were observed between the 1s and 60s epochs for the cut-points of Evenson et al. (2008) and Romanzini et al. (2014), respectively (Figure 4.3D). Significant differences ($p < .01$) were observed between the VPA times spent within and between all epoch length. A decrease in VPA time occurred as epoch length increased for both cut-points. For moderate-to-vigorous physical activity (MVPA), a 17 min (27.5 %) and a 18 min (30.9 %) decrease were observed between the 1s and 60s epochs for Evenson et al. (2008) and Romanzini et al. (2014) Cut-points, respectively (Figure 4.4). The mean minutes of MVPA per day between both the cut-point algorithms and all epoch length are statistically significantly different ($p < .01$). A decrease in MVPA time occurred as the epoch length increased for both cut-points.

For further analysis, only the data set with an epoch length of 15 s was analyzed since this was also used as a parameter in the validation study of the two cut-point algorithms Evenson et al. (2008) and Romanzini et al. (2014). When examining the intensities of individual days (without days with 0 min WT; n=9,937) for normal distribution, all tests were significant due to the high number of observations, but visual inspection of the Q-Q graphs showed that mainly some outliers in the high and low ranges are responsible (Supplement Figure A.2, A.3, A.4, A.5). Skewness and kurtosis showed that the distribution with SED was minimally left-skewed (Evenson: .19, Romanzini: .08) and values below the mean were more likely to be obtained (Evenson: .91, Romanzini: .99), and that marginal areas are more pronounced. For LPA, the distribution was almost symmetric (Evenson: -.04, Romanzini: .10) for both algorithms but equally downward sloping (Evenson: -.27, Romanzini: -.27). MPA and VPA were no longer normally distributed for both algorithms, with a majority of days with low values.

4.4.3 Differences in mean min for SED, LPA, MPA & VPA by EL and cut point algo

The results of the t-test showed that MVPA was significantly lower for Romanzini (M = 42.89; SD = 34.19) compared to Evenson (M = 46.28; SD = 36.22, $t(1056)$ = 130.79; $p < .001$; d=1.27). Taking each step individually, MPA was significantly lower

Figure 4.3: Mean minutes spend in a) sedentary time, b) light, c) moderate and d) vigorous physical activity per day interpreted using Evenson et al. & Romanzini et al. algorithms at five different epoch settings (95% CI).



Note: Epoch length of 15s is highlighted with a black edge since it is the one it was validated for.

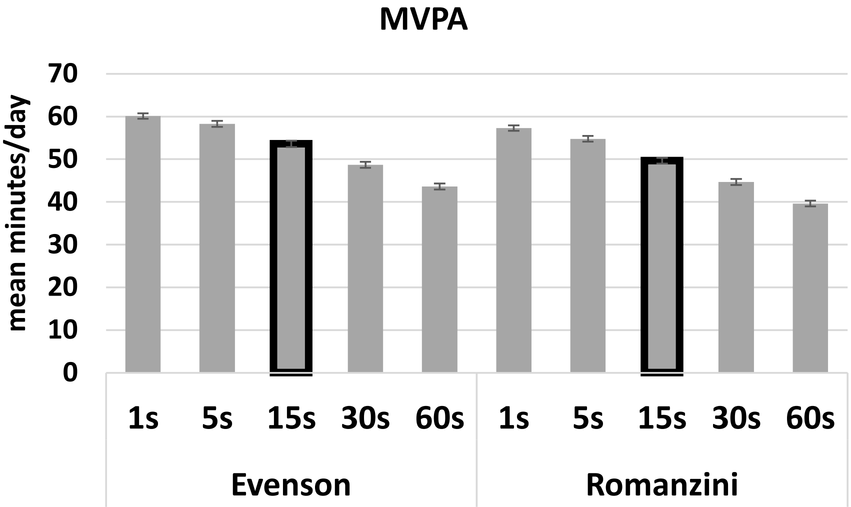
for Romanzini (M = 17.11; SD = 13.06) compared to Evenson (M = 30.53; SD = 22.95, $t(10556)$ = 131.69; $p < .001$; d=1.28). Whereas VPA was significantly higher for Romanzini (M = 25.78; SD = 23.74) compared to Evenson (M = 15.75; SD = 17.14, $t(1056)$ = 122.00; $p < .001$; d=1.19). According to Cohen (1988), all the differences were large.

Calculating active days based on the WHO guidelines on PA (World Health Organization, 2010), which means more than 60 min of MVPA per day, there were some differences based on the epoch length and cut-point algorithm. Of the 10,557 recorded days, 4,033 (38.2 %) and 3,686 (34.9 %) were considered active days at a 1s epoch length in Evenson et al. and Romanzini et al. cut-points, respectively. This number was reduced to 3,189 (30.2 %) and 2,789 (26.4 %) active days at an epoch length of 15s, and 2257 (21.4 %) and 1909 (18.1 %) at an epoch length of 60s for Evenson et al. and Romanzini et al. cut-points, respectively.

4.5 Discussion

The lack of a standardized preprocessing process for accelerometer data has challenged research efforts to gain a comprehensive understanding of children's and adolescents' PA and how they can be related to their health behaviors. It is recommended to analyze multiple accelerometer data preprocessing algorithms and identify those in which the validation sample best fits the sample at hand. Therefore, the present study aimed to examine the accelerometer data of MoMo Wave 3 (2018–2022) concerning different data processing approaches used specifically in this scenario. We sequentially assessed three

Figure 4.4: Mean minutes spend in moderate-to-vigorous physical activity per day interpreted using Evenson et al. & Romanzini et al. algorithms at five different epoch settings (95% CI).



Note: Epoch length of 15s is highlighted with a black edge since it is the one it was validated for.

factors that based on previous research may influence the quantification of children's and adolescents' PA using accelerometers: 1) non-wear time algorithm selection, 2) epoch length selection, and 3) cut-point selection. In our study, we identified problems when evaluating data analyzed with different epoch lengths. Specifically, when using different NWT algorithms and definitions of activity cut-points with epoch lengths that were not originally used to validate these algorithms, the resulting differences in the estimated SED and PA values turned out to be very large.

4.5.1 Differences in non-wear time by epoch length

The NWT has a profound impact on the accurate calculation of summary statistics, such as minutes spent in SED or LPA (Syed et al., 2020; Vanhelst et al., 2019). Except for the edges of the data, both NWT algorithms applied in this study differed only in the tenth range (Figure 4.1, Supplement Table A.1). For both algorithms, almost three-quarters of the recorded days had a wear time between 8 and 16 hours (Choi = 73.7 %, Troiano = 74.7 %). Although there were only small differences between the algorithms in general, different epoch lengths had a significant influence on the results when using the Troiano algorithm, as was also found by others (Banda et al., 2016). The reason for these differences may be attributable to the way the NWT is accumulated by different epoch lengths during the allowance periods since individual counts are always accumulated to the full epoch length, whether it is 10s or 60s. The 2-minute allowance period of non-zero counts in the Troiano algorithm is getting filled up by counts occurring during the epochs in this 2-minute window. Counts occurring during a smaller epoch length like 10-seconds only contribute 10 seconds to the allowance window, the same counts during a 60-second epoch contribute 60 seconds to the 2-minute window. Longer epoch lengths, therefore, lead to less NWT since the 2-minute allowance window is much faster overreached for the Troiano algorithm. Banda et al. (2016) on the other hand described that the Choi algorithm accumulates all non-60-second epoch data into 60-second epoch data before the Choi NWT algorithm is applied on the data. They analyzed the R- package provided by Choi et al. (2011) and did a comparison with the implementation in the ActiLife software and found that both implementations work the same way using only accumulated 60-second epochs. Since the epoch length has also to be taken into account in further data processing (for example, the intensity algorithms for different age groups can differ in the specified epoch lengths), it makes sense to use a NWT algorithm that remains constant. The Choi algorithm has also been validated for an age range of 10-67 years (Choi et al., 2011). Aadland et al. (2018) found differences in various accelerometer NWT criteria and recommend that NWT algorithms should be standardized across studies to avoid confounding and improve the comparability of children's PA levels. Studies with adolescents older than 11 years have shown that there were differences that influenced the evaluation of SED time depending on the choice of NWT rules (Aadland et al., 2018; Vanhelst et al., 2019). For these, it is recommended to choose algorithms with shorter periods of continuous zeros for defining NWT to improve the accuracy in determining the activity levels. Vanhelst et al. (2019) recommend 30 minutes of continuous zeros because Choi et al. (2011) and Troiano et al. (2008), with longer zero criteria, among others, overestimate time with SED. However, this overestimation could be due to their choice to use a 1s epoch length in their analysis, since smaller epoch lengths lead to more SED classifications (Figure 4.3A). Although Vanhelst et al. (2019) conducted a methodologically very well-designed NWT study, and the recorded data were examined and compared with accurate log books, it comes down to small details in data reduction decisions such as the selection of an appropriate epoch length. Two other studies suggested that a 20-minute window is too short and recommended a longer window of at least 60 minutes of consecutive zeros as a realistic NWT criterion for younger children aged 7-13 years (Banda et al., 2016; Chinapaw et al., 2014). However, these studies had small numbers of participants while in MoMo, the complete cross-section of the society is represented. In MoMo, one of the goals was to include as many different children as possible. Toftager et al. (2013) have shown that overweight children drop out of the data set more often when using shorter window sizes because their behavior is more sedentary and sometimes misclassified as non-wear time resulting in not reaching the required wearing time. Since Choi uses a 90 min window of consecutive zero/nonzero counts instead of, for example, a smaller window of 30 or 60 min, overweight children should be included more often in the MoMo dataset instead of dropping out. The longitudinal design and the long-term study duration of almost 20 years of the MoMo study, with a total of 4 survey waves (baseline, wave 1 to 3), resulted in a very large age range and heterogeneity sample of participants. Although a large part of the data is drawn from a cross-sectional sample of participants aged 6 to 17 years each wave, evaluations of the repeating and now adult participants aged 18 to 32 years are planned soon. To be able to use validated algorithms for children, adolescents, and adults with different epoch lengths while still maintaining the comparability of the data, we decided to use Choi's algorithm for calculating the NWT in MoMo.

4.5.2 Valid wear time criterion

The results revealed declining valid days for longer wear time criteria (Table 4.2). 5 % fewer valid days when changing from an 8h to a 10h wear time criterion may still be acceptable, but 20 % less when compared with a 12h wear time criterion would have a large loss of participants as a consequence. This is especially true in light of the large social background spectrum from which MoMo participants are recruited; as many children and adolescents as possible were included. Therefore, the MoMo and KiGGS accelerometer data sets were considered valid with a minimum wear time of 8 hours of recordings on 4 weekdays and 1 weekend day. These scoring policies are also consistent with the requirements for inclusion in the International Children's Accelerometry Database (ICAD) (Sherar et al., 2011). Furthermore, the literature suggests that 4 days is a reasonable measurement time, thus reducing the burden on participants and making it easier for researchers to collect sufficient data for the formation of recommendations related to general health guidelines (Colley et al., 2010; Mâsse et al., 2005; Matthews et al., 2002; Toftager et al., 2013), although more days are even better (Chinapaw et al., 2014). To archive the highest possible number of valid data sets, the accelerometer should be worn for 7 days following the day of the examination in the study center. The assessment period of at least seven days ensured the inclusion of weekdays and weekend days. This inclusion is recommended due to differences in PA between weekdays and weekends (Burchartz et al., 2022; Chinapaw et al., 2014; Donaldson et al., 2016; Matthews et al., 2002).

4.5.3 Differences in Cut-point Algorithms

We found similar patterns of results for the two cut-point algorithms considered in MoMo for the average minutes per day classified in the intensity levels SED and PA. For both algorithms, the estimates for SED and VPA decreased with increasing epoch length, while the estimates for LPA increased and MPA peaked at 15 s. A large amount of SED was reclassified as LPA when longer epoch lengths were used, regardless of the cut-point definitions. This can be seen in Figure 4.3, and similar results were found in other studies (Banda et al., 2016; Logan et al., 2016). It should be noted that the wear time resulting from the sum of all assigned intensities did not differ between Evenson et al. (2008) and Romanzini et al. (2014) when different epoch lengths were used. The differences arise only from the allocation of the activity to a certain intensity level.

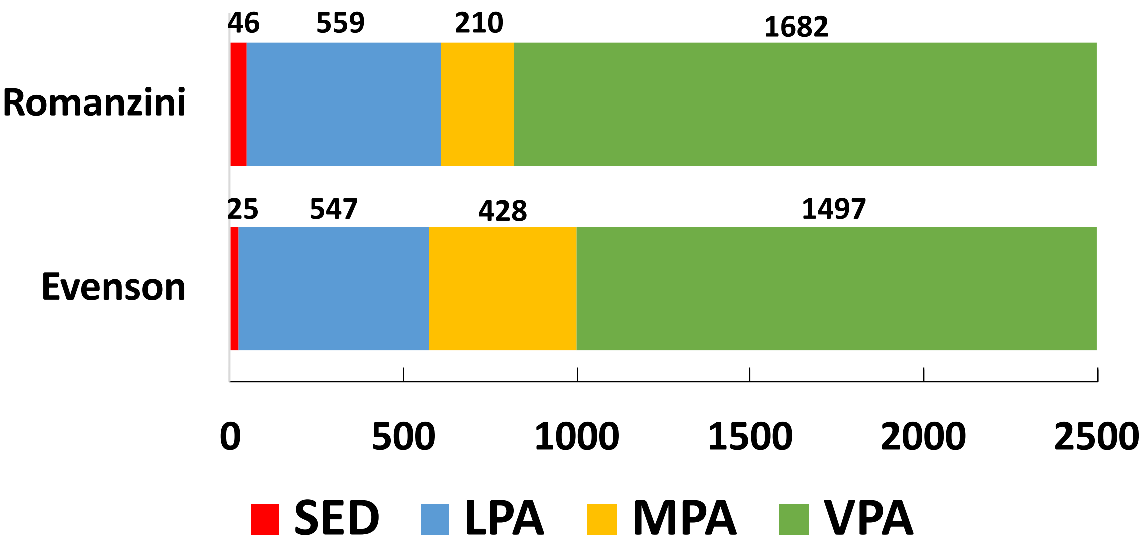
4.5.4 Which EL should be used?

For children and adolescents, an epoch length between 1s and 5s or the shortest possible epoch length is recommended by Bornstein et al. (2011), Migueles et al. (2017), and Vanhelst et al. (2012) due to the sporadic activity of children. It must be acknowledged, however, that when deciding which epoch length to use, one must carefully weigh the potential benefit of this decision for a small epoch length (e.g., 1 second) against the validity of a change in the original activity cut-point definitions (e.g., a cut-point validated at an epoch length of 15 seconds). The large differences observed in the estimates of WT, SED, and PA used in converting the validated cut-point epochs in the present study suggest the use of validated epoch length in future studies as long as there are no new validation studies with short epoch lengths for children. Assuming that this is possible, at a minimum, open documentation of data processing decisions must be provided to help other studies interpret the results. This is even more important to nationally representative studies such as MoMo, since other studies with special subsamples (e.g., children with certain diseases) from Germany, for example, are using these results to compare their data.

4.5.5 Which cut-point algorithm should be used?

Banda et al. (2016) and Logan et al. (2016) also examined the differences between cut-points of Evenson and Romanzini and found similar results. Both compared the algorithm for the vector magnitude and not the vertical axis of Romanzini with the algorithm of the vertical axis of Evenson. Therefore, comparability is rather complicated because different raw data sets (vertical axis vs. vector magnitude) were considered. Another study by Breau et al. (2022) also compared Evenson with newer algorithms for children under 6 years of age and found significant variations in times spent in different intensities of PA. The choice as to which cut-points to apply to young children's data to calculate times in different PA intensities, in turn, affects the proportion of children who meet the guidelines, which is also suggested by results from other studies (Banda et al., 2016; Breau et al., 2022; Logan et al., 2016). Thus, comparisons of movement intensities should not be made across studies with the same study population using different sets of cut-points. We again emphasize the need for additional studies validating WT algorithms and cut-points using smaller epoch lengths in children and adolescents, as already suggested elsewhere (Giurgiu et al., 2022).

Figure 4.5: Visualization of the cut point ranges for 15s epochs of vertical axis algorithms



Note: **SED** Sedentary behavior, **LPA** light physical activity, **MPA** moderate physical activity, **VPA** vigorous physical activity.

On the other hand, Migueles et al. (2017) recommended using different cut-point criteria for different age groups. Therefore, and because the MoMo sample has such a wide age range from 6-32 years, we still decided to divide the sample into different age groups. In this study, the two groups of children (6-10 years) and adolescents (11-17 years) and the most suitable cut-points of Evenson (5-9 years) and Romanzini (10-15 years) were compared. The advantage of these two algorithms is that they fit very closely with the two age groups in MoMo, had similar validation protocols, and, in addition, both used an epoch length of 15s. Table 4.1 and Figure 4.5 also show similar cut-point ranges of the two algorithms. This also explains the slight differences in the classification of SED and PA times. The cut-point range for SED is larger in Romanzini, leading to an increased classification of SED in contrast to Evenson. In comparison, there is an earlier classification of LPA and a larger range for the classification of MPA times in Evenson. VPA again takes an earlier and larger part of the classification in Romanzini than Evenson, resulting in higher amounts of VPA when using the Romanzini cut-points.

Small differences in the applied cut-point range should be noted when comparing these two age groups. However, in general, we do not consider this to be a problem, since SED is correlated with age. As children get older, their school curricular activities change from very short and unpredictable activity bouts to longer and more constant physical behavior sessions (Crane et al., 2018). This is also confirmed by PA behavior captured with MoMo questionnaires, which also showed more SED for older children (Burchartz et al., 2021). The differences in the classification of MPA and VPA compensate each other to the extent that both are often considered together under the sum term MVPA anyway. Therefore, the more frequent classification of MPA in Evenson would be compensated by the more frequent classification of VPA in Romanzini and vice versa (Figure 4.4). However, when examining individual classifications such as VPA, this factor should be kept in mind, especially since the percentage of VPA on an average day has been shown to remain stable across age groups, while the absolute time spent on VPA increases at the same time due to longer waking hours among adolescents (Burchartz et al., 2022).

The present study adds important knowledge based on its investigation of a large sample of children and adolescents who were evaluated under free-living conditions, even if the evaluation was not performed at the subject, gender, or age-specific level. It provides important results for the data processing of NWT, epoch length, and intensity classification on data from the same subjects on the same days. With its comprehensive documentation of all technical decisions, the present study improves replicability. However, the following limitations of this study should be acknowledged: a) Cut-points based on a vertical axis instead of vector magnitude are known to produce significantly lower estimations of time spent in MVPA (Migueles et al., 2019). The two cut-points under analysis in this study only used the vertical axis. The MVPA values of other vector magnitude cut-points might therefore be higher. Although recent studies (Breau et al., 2022; Leppänen et al., 2022) have investigated other cut-points, we are not aware of any study that has investigated the specific algorithms used in MoMo; b) we did not identify sleep/wake states since the participants were told to remove the device for sleeping. Therefore, separate detection of sleep phases was not possible; and c) we were unable to evaluate the precision of the individual cut-point limits, as we did not have a comparison criterion measure available in the present study. In current device-based PA research, the results are reported with different criteria for collection and processing, as no consensus has been found yet. Reporting includes a wide variety of methods that require a high level of technical expertise with sometimes detailed explanations of nuanced differences. These can overwhelm the unsophisticated recipient and lead to poorly selected criteria. Pooling accelerometer data, the merging and combining of multiple datasets from different studies, and re-analyzing them in an uniform way across studies may be the best long-term solution to overcome inconsistencies in processing criteria, as implemented by ICAD, for example (Steene-Johannessen et al., 2020). By using standardized methods, the data can then be processed in a consistent manner to enable the creation of comparable accelerometer variables. Recently, newer open-source metrics are being used more and more often to analyze accelerometer data (Leppänen et al., 2022). For a long time, count-based evaluation has only been possible with a proprietary algorithm because different manufacturers used different and unknown proprietary definitions to calculate them. Although methods are now known to calculate strong correlating counts produced by ActiGraphs for accelerometers from other manufacturers (Brønd & Arvidsson, 2016; Brondeel et al., 2021; Clevenger et al., 2022), the trend in evaluation is to use the original raw data as a basis for calculating certain metrics. Faced with this development, ActiGraph has recently published the complete algorithm for calculating the device-internal counts, so that raw data from other accelerometers can now officially be converted into ActiGraph counts (Neishabouri et al., 2022). Open-source analyzes could improve comparability between studies by comparing different device models and help further rationalize data analysis (Clevenger et al., 2022).

4.6 Conclusion

This study demonstrates the extent of variation in the results for PA levels across epoch lengths, different non-wear algorithms, and between different cut-point algorithms. We identified problems when evaluating data analyzed with different epoch lengths, specifically when different NWT algorithms and definitions of activity cut-points are used with epoch lengths that were not originally used to validate these algorithms. As a consequence, the resulting differences in the results of estimated SED and PA values can become very large.

Therefore, our results confirm previous studies and extend their findings to a sample of children and adolescents in Germany. We propose to pool the data and evaluate them in a harmonized way in the future. In addition to new validation studies with short epoch lengths for young children (e.g., 1 or 5 s), we also propose to conduct meta-analyses using data from multiple studies to validate cutoff points to propose a consensual set of cutoff points that can be used in different settings and projects.

Authors' contributions

AB was responsible for the general conception and design of this manuscript. SS supported in the statistical analysis. AB, CN, SS, SK were responsible for the acquisition of the data. SK, SS, BHM, CN, AW were responsible for the critical revision. CN & AW contributed to the design of the study. All authors read and approved the final manuscript.

References Methods-Paper

Aadland, E., Andersen, L. B., Anderssen, S. A., & Resaland, G. K. (2018). A comparison of 10 accelerometer non-wear time criteria and logbooks in children. *BMC public health*, *18*(1), 323. https://doi.org/10.1186/s12889-018-5212-4

Arvidsson, D., Fridolfsson, J., Börjesson, M., Andersen, L. B., Ekblom, Ö., Dencker, M., & Brønd, J. C. (2019). Re-examination of accelerometer data processing and calibration for the assessment of physical activity intensity. *Scandinavian Journal of Medicine & Science in Sports*, *29*(10), 1442–1452. https://doi.org/10.1111/sms.13470

Bakrania, K., Yates, T., Rowlands, A. V., Esliger, D. W., Bunnewell, S., Sanders, J., Davies, M. J., Khunti, K., & Edwardson, C. L. (2016). Intensity Thresholds on Raw Acceleration Data: Euclidean Norm Minus One (ENMO) and Mean Amplitude Deviation (MAD) Approaches. *PLOS ONE*, *11*(10), e0164045. https://doi.org/10.1371/journal.pone.0164045

Banda, J. A., Haydel, K. F., Davila, T., Desai, M., Bryson, S., Haskell, W. L., Matheson, D., & Robinson, T. N. (2016). Effects of Varying Epoch Lengths, Wear Time Algorithms, and Activity Cut-Points on Estimates of Child Sedentary Behavior and Physical Activity from Accelerometer Data. *PLOS ONE*, *11*(3), e0150534. https://doi.org/10.1371/journal.pone.0150534

Bornstein, D. B., Beets, M. W., Byun, W., & McIver, K. L. (2011). Accelerometer-derived physical activity levels of preschoolers: a meta-analysis. *Journal of Science and Medicine in Sport*, *14*(6), 504–511. https://doi.org/10.1016/j.jsams.2011.05.007

Breau, B., Coyle-Asbil, H. J., Haines, J., Ma, D. W. L., & Vallis, L. A. (2022). ActiGraph Cutpoints Impact Physical Activity and Sedentary Behavior Outcomes in Young Children. *Journal for the Measurement of Physical Behaviour*, *5*(2), 85–96. https://doi.org/10.1123/jmpb.2021-0042

Brønd, J. C., & Arvidsson, D. (2016). Sampling frequency affects the processing of Actigraph raw acceleration data to activity counts. *Journal of applied physiology (Bethesda, Md. : 1985)*, *120*(3), 362–369. https://doi.org/10.1152/japplphysiol.00628.2015

Brondeel, R., Kestens, Y., Rahimipour Anaraki, J., Stanley, K., Thierry, B., & Fuller, D. (2021). Converting Raw Accelerometer Data to Activity Counts Using Open-Source Code: Implementing a MATLAB Code in Python and R, and Comparing the Results to ActiLife. *Journal for the Measurement of Physical Behaviour*, *4*(3), 205–211. https://doi.org/10.1123/jmpb.2019-0063

Burchartz, A., Anedda, B., Auerswald, T., Giurgiu, M., Hill, H., Ketelhut, S., Kolb, S., Mall, C., Manz, K., Nigg, C. R., Reichert, M., Sprengeler, O., Wunsch, K., & Matthews, C. E. (2020). Assessing physical behavior through accelerometry –

State of the science, best practices and future directions. *Psychology of Sport and Exercise*, *49*, 101703. https://doi.org/10.1016/j.psychsport.2020.101703

Burchartz, A., Manz, K., Anedda, B., Niessner, C., Oriwol, D., Schmidt, S. C. E., & Woll, A. (2020). Measurement of Physical Activity and Sedentary Behavior by Accelerometry Among a Nationwide Sample from the KiGGS and MoMo Study: Study Protocol. *JMIR Research Protocols*, *9*(7), e14370. https://doi.org/10.2196/14370

Burchartz, A., Oriwol, D., Kolb, S., Schmidt, S. C. E., von Haaren-Mack, B., Niessner, C., & Woll, A. (2022). Impact of weekdays versus weekend days on accelerometer measured physical behavior among children and adolescents: results from the MoMo study. *German Journal of Exercise and Sport Research*, *52*(2), 218–227. https://doi.org/10.1007/s12662-022-00811-4

Burchartz, A., Oriwol, D., Kolb, S., Schmidt, S. C. E., Wunsch, K., Manz, K., Niessner, C., & Woll, A. (2021). Comparison of self-reported & device-based, measured physical activity among children in Germany. *BMC public health*, *21*(1), 1081. https://doi.org/10.1186/s12889-021-11114-y

Cain, K. L., Sallis, J. F., Conway, T. L., van Dyck, D., & Calhoon, L. (2013). Using Accelerometers in Youth Physical Activity Studies: A Review of Methods. *Journal of Physical Activity and Health*, *10*(3), 437–450. https://doi.org/10.1123/jpah.10.3.437

Chinapaw, M. J. M., de Niet, M., Verloigne, M., de Bourdeaudhuij, I., Brug, J., & Altenburg, T. M. (2014). From sedentary time to sedentary patterns: accelerometer data reduction decisions in youth. *PLOS ONE*, *9*(11), e111205. https://doi.org/10.1371/journal.pone.0111205

Choi, L., Liu, Z., Matthews, C. E., & Buchowski, M. S. (2011). Validation of accelerometer wear and nonwear time classification algorithm. *Medicine and science in sports and exercise*, *43*(2), 357–364. https://doi.org/10.1249/MSS.0b013e3181ed61a3

Clevenger, K. A., Mackintosh, K. A., McNarry, M. A., Pfeiffer, K. A., Montoye, A. H. K., & Brønd, J. C. (2022). Comparison of Child and Adolescent Physical Activity Levels From Open-Source Versus ActiGraph Counts. *Journal for the Measurement of Physical Behaviour*, *5*(2), 120–128. https://doi.org/10.1123/jmpb.2021-0057

Colley, R. C., Connor-Gorber, S., & Tremblay, M. S. (2010). Quality control and data reduction procedures for accelerometry-derived measures of physical activity. *Health Reports*, *21*(1), 63–69.

Crane, J. R., Naylor, P.-J., & Temple, V. A. (2018). The Physical Activity and Sedentary Behaviour Patterns of Children in Kindergarten and Grade 2. *Children (Basel, Switzerland)*, *5*(10). https://doi.org/10.3390/children5100131

Dencker, M., Svensson, J., El-Naaman, B., Bugge, A., & Andersen, L. B. (2012). Importance of epoch length and registration time on accelerometer measurements in younger children. *The Journal of sports medicine and physical fitness*, *52*(2), 115–121.

Dias Moura, I. R., Barbosa, A. O., Da Silva, I. C. M., Romanzini, M., Prazeres Filho, A., & de Farias Júnior, J. C. (2019). Impact of cutoff points on adolescent sedentary behavior measured by accelerometer. *Revista Brasileira de Atividade Física & Saúde*, *24*, 1–7. https://doi.org/10.12820/rbafs.24e0071

Donaldson, S. C., Montoye, A. H. K., Tuttle, M. S., & Kaminsky, L. A. (2016). Variability of Objectively Measured Sedentary Behavior. *Medicine and science in sports and exercise*, *48*(4), 755–761. https://doi.org/10.1249/MSS.0000000000000828

Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S., & McMurray, R. G. (2008). Calibration of two objective measures of physical activity for children. *Journal of sports sciences*, *26*(14), 1557–1565. https://doi.org/10.1080/02640410802334196

Giurgiu, M., Kolb, S., Nigg, C., Burchartz, A., Timm, I., Becker, M., Rulf, E., Doster, A.-K., Koch, E. D., Bussmann, J. B. J., Nigg, C. R., Ebner-Priemer, U. W., & Woll, A. (2022). Assessment of 24-hour physical behaviour in children and adolescents via wearables: a systematic review of free-living validation studies. *BMJ Open Sport & Exercise Medicine*, *8*(2), e001267. https://doi.org/10.1136/bmjsem-2021-001267

Hoffmann, R., Lange, M., Butschalowsky, H., Houben, R., Allen, J., Kuhnert, R., Schaffrath Rosario, A., & Gößwald, A. (2018). Querschnitterhebung von KiGGS Welle 2: Teilnehmendengewinnung, Response und Repräsentativität (Journal of Health Monitoring, Ed.). https://doi.org/10.17886/RKI-GBE-2018-015

Leppänen, M. H., Migueles, J. H., Abdollahi, A. M., Engberg, E., Ortega, F. B., & Roos, E. (2022). Comparing estimates of physical activity in children across different cut-points and the associations with weight status. *Scandinavian Journal of Medicine & Science in Sports*, *32*(6), 971–983. https://doi.org/10.1111/sms.14147

Logan, G. R. M., Duncan, S., Harris, N. K., Hinckson, E. A., & Schofield, G. (2016). Adolescent physical activity levels: discrepancies with accelerometer data analysis. *Journal of sports sciences*, *34*(21), 2047–2053. https://doi.org/10.1080/02640414.2016.1150599

Mâsse, L. C., Fuemmeler, B. F., Anderson, C. B., Matthews, C. E., Trost, S. G., Catellier, D. J., & Treuth, M. S. (2005). Accelerometer data reduction: a comparison of four reduction algorithms on select outcome variables. *Medicine and science in sports and exercise*, *37*(11 Suppl), S544–54. https://doi.org/10.1249/01.mss.0000185674.09066.8a

Matthews, C. E., Ainsworth, B. E., Thompson, R. W., & Bassett, D. R. (2002). Sources of variance in daily physical activity levels as measured by an accelerometer. *Medicine and science in sports and exercise*, *34*(8), 1376–1381. https://doi.org/10.1097/00005768-200208000-00021

Mauz, E., Lange, M., Houben, R., Hoffmann, R., Allen, J., Gößwald, A., Hölling, H., Lampert, T., Lange, C., Poethko-Müller, C., Richter, A., Rosario, A. S., von Schenck, U., Ziese, T., & Kurth, B.-M. (2019). Cohort profile: KiGGS cohort longitudinal study on the health of children, adolescents and young adults in Germany. *International journal of epidemiology*, *49*(2), 375–375k. https://doi.org/10.1093/ije/dyz231

Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Delisle Nyström, C., Mora-Gonzalez, J., Löf, M., Labayen, I., Ruiz, J. R., & Ortega, F. B. (2017). Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports medicine (Auckland, N.Z.)*, *47*(9), 1821–1845. https://doi.org/10.1007/s40279-017-0716-0

Migueles, J. H., Cadenas-Sanchez, C., Tudor-Locke, C. E., Löf, M., Esteban-Cornejo, I., Molina-Garcia, P., Mora-Gonzalez, J., Rodriguez-Ayllon, M., Garcia-Marmol, E., Ekelund, U., & Ortega, F. B. (2019). Comparability of published cut-points for the assessment of physical activity: Implications for data harmonization. *Scandinavian Journal of Medicine & Science in Sports*, *29*(4), 566–574. https://doi.org/10.1111/sms.13356

Neishabouri, A., Nguyen, J., Samuelsson, J., Guthrie, T., Biggs, M., Wyatt, J., Cross, D., Karas, M., Migueles, J. H., Khan, S., & Guo, C. C. (2022). Quantification of acceleration as activity counts in ActiGraph wearable. *Scientific Reports*, *12*(1), 11958. https://doi.org/10.1038/s41598-022-16003-x

Pfeiffer, K. A., Clevenger, K. A., Kaplan, A., van Camp, C. A., Strath, S. J., & Montoye, A. H. K. (2022). Accessibility and use of novel methods for predicting physical activity and energy expenditure using accelerometry: a scoping review. *Physiological Measurement*, *43*(9). https://doi.org/10.1088/1361-6579/ac89ca

Romanzini, M., Petroski, E. L., Ohara, D., Dourado, A. C., & Reichert, F. F. (2014). Calibration of ActiGraph GT3X, Actical and RT3 accelerometers in adolescents. *European journal of sport science*, *14*(1), 91–99. https://doi.org/10.1080/17461391.2012.732614

Rowlands, A. V., Mirkes, E. M., Yates, T., Clemes, S., Davies, M. J., Khunti, K., & Edwardson, C. L. (2018). Accelerometer-assessed Physical Activity in Epidemiology: Are Monitors Equivalent? *Medicine and science in sports and exercise*, *50*(2), 257–265. https://doi.org/10.1249/mss.0000000000001435

Sherar, L. B., Griew, P., Esliger, D. W., Cooper, A. R., Ekelund, U., Judge, K., & Riddoch, C. J. (2011). International children's accelerometry database (ICAD): Design and methods. *BMC public health*, *11*, 485. https://doi.org/10.1186/1471-2458-11-485

Steene-Johannessen, J., Hansen, B. H., Dalene, K. E., Kolle, E., Northstone, K., Møller, N. C., Grøntved, A., Wedderkopp, N., Kriemler, S., Page, A. S., Puder, J. J., Reilly, J. J., Sardinha, L. B., van Sluijs, E. M. F., Andersen, L. B., van der Ploeg, H. P., Ahrens, W., Flexeder, C., Standl, M., . . . Ekelund, U. (2020). Variations in accelerometry measured physical activity and sedentary time across Europe - harmonized analyses of 47,497 children and adolescents. *International Journal of Behavioral Nutrition and Physical Activity*, *17*(1), 38. https://doi.org/10.1186/s12966-020-00930-x

Syed, S., Morseth, B., Hopstock, L. A., & Horsch, A. (2020). Evaluating the performance of raw and epoch non-wear algorithms using multiple accelerometers and electrocardiogram recordings. *Scientific Reports*, *10*(1), 5866. https://doi.org/10.1038/s41598-020-62821-2

Toftager, M., Kristensen, P. L., Oliver, M., Duncan, S., Christiansen, L. B., Boyle, E., Brønd, J. C., & Troelsen, J. (2013). Accelerometer data reduction in adolescents: effects on sample retention and bias. *International Journal of Behavioral Nutrition and Physical Activity*, *10*, 140. https://doi.org/10.1186/1479-5868-10-140

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine and science in sports and exercise*, *40*(1), 181–188. https://doi.org/10.1249/mss.0b013e31815a51b3

Trost, S. G. (2007). State of the Art Reviews: Measurement of Physical Activity in Children and Adolescents. *American Journal of Lifestyle Medicine*, *1*(4), 299–314. https://doi.org/10.1177/1559827607301686

Vanhelst, J., Béghin, L., Duhamel, A., Bergman, P., Sjöström, M., & Gottrand, F. (2012). Comparison of uniaxial and triaxial accelerometry in the assessment of physical activity among adolescents under free-living conditions: The HELENA study. *BMC medical research methodology*, *12*(1), 26. https://doi.org/10.1186/1471-2288-12-26

Vanhelst, J., Vidal, F., Drumez, E., Béghin, L., Baudelet, J.-B., Coopman, S., & Gottrand, F. (2019). Comparison and validation of accelerometer wear time and non-wear time algorithms for assessing physical activity levels in children and adolescents. *BMC medical research methodology*, *19*(1), 72. https://doi.org/10.1186/s12874-019-0712-1

Woll, A., Albrecht, C., & Worth, A. (2017). Motorik-Module (MoMo) – the KiGGS wave 2 module to survey motor performance and physical activity. *Journal of Health Monitoring*, *2*(S3), 66–73. https://doi.org/10.17886/RKI-GBE-2017-110

World Health Organization. (2010). *Global recommendations on physical activity for health*.

Xing, R., Huang, W. Y., & Sit, C. H.-P. (2021). Validity of accelerometry for predicting physical activity and sedentary time in ambulatory children and young adults with cerebral palsy. *Journal of exercise science and fitness*, *19*(1), 19–24. https://doi.org/10.1016/j.jesf.2020.06.006

5. Comparison of self-reported & device-based, measured physical activity

Slightly modified version of the 3^{rd} published article

Burchartz, Alexander; Oriwol, Doris; Kolb, Simon; Schmidt, Steffen C. E.; Wunsch, Kathrin; Manz, Kristin; Niessner, Claudia; Woll, Alexander (2021): Comparison of self-reported & device-based, measured physical activity among children in Germany.

BMC public health 21 (1).
DOI: 10.1186/s12889-021-11114-y.

5.1 Abstract

Background: As children show a more complex but less structured physical behavior than adults, assessment of their many spontaneous and impulsive movements is a challenge for physical activity (PA) assessment. Since neither questionnaires nor accelerometers enable optimal detection of all facets of PA, a multimodal, combined approach of self-reported and device-based methods is recommended. Based on the number of days on which the participants reached the physical activity (PA) values given in the WHO guideline, this study examines 1) the difference between self-reported and device-based, measured PA and 2) whether PA differences between age and gender groups obtained by two methods are comparable.
Methods: Participants aged 6–17 years were randomly chosen and data were collected representatively at 167 sample points throughout Germany within the Motorik-Modul Study. PA of n = 2694 participants (52% female) was measured using the ActiGraph accelerometer (ACC) and a physical activity questionnaire (PAQ). The sample was divided into three age groups (6–10 yrs. n = 788, 11–13 yrs. n = 823, 14–17 yrs. n = 1083). Numbers of days per week with at least 60 min moderate to vigorous PA (MVPA) were analyzed for both methods.

Results: Only every 25th respondent (4%) reaches the WHO standard of 60 min MVPA every day if measured with ACC. Self-reported PA was slightly higher (9%) (mean$_{PAQ}$ = 3.82 days; mean$_{ACC}$ = 2.34 days; F_{method} = 915.85; p = <.001; f_{Cohen} = .64). The differences between the methods are significantly smaller in younger children than in the older age groups (F_{age} = 264.2, $p < .001$; f_{Cohen} = .48). The older the subjects are, the lower is the proportion of those who meet the WHO guideline on each day, with girls meeting the guideline less frequently than boys in all age groups.
Conclusion: Children and adolescents living in Germany show a very low adherence to the WHO guideline on PA. While younger children are much more active with their free play, especially children over 10 years of age and especially girls should be the target of programs to increase PA.

5.2 Background

As children show a more complex but less structured physical behavior than adults (Gabrys et al., 2015; Graf et al., 2013), assessment of their many spontaneous and impulsive movements is a challenge for physical activity (PA) assessment (Müller et al., 2011). To date, questionnaires have been the most commonly used method to assess PA in large, epidemiological studies. One of the biggest advantages of questionnaires lies in their versatility. In addition to the assessment of duration, frequency, and intensity, self-report methods provide information about the type of PA, which is not feasible with common device-based methods like accelerometry without complex diaries or ambulatory assessment. Furthermore, epidemiological research requires larger samples, which is why questionnaires often are the more feasible alternative (Jekauc et al., 2013). However, many studies have shown that the level of PA assessed by self-reports is often overestimated (Jekauc et al., 2013; Slootmaker et al., 2009; Wijndaele et al., 2015). Since this overestimation is higher in children and adolescents compared to adults, it is particularly important to make accurate investigations in the former group (Slootmaker et al., 2009). Unstructured and irregular activities in everyday life are difficult to remember correctly. Accelerometers were used more frequently to measure PA in recent large-scale studies (Troiano et al., 2014), as they became more feasible, more accurate, and more affordable in the last decade. By measuring movement acceleration, everyday PA, including PA intensities and patterns, can be recorded in more detail than by self-reports. In conclusion, no single procedure provides for optimal detection of all facets of PA, which is why a multimodal, combined approach of self-reported and devicebased methods is recommended (Jekauc et al., 2014; Schutz et al., 2001). To the best of our knowledge, this study is the first attempt to compare self-reported and device-based, measured PA guideline adherence in a nationwide sample of children and adolescents living in Germany. Based on the number of days on which the participants reached the value given in the World Health Organization (WHO) guideline (World Health Organization, 2010), this study examines 1) the difference between self-reported and device-based, measured PA, and 2) whether PA differences between age and gender groups are comparable in the two methods.

5.3 Methods

5.3.1 Study design

The German Health Interview and Examination Survey for Children and Adolescents (KiGGS) is part of the Federal Health Monitoring System conducted by the Robert

Koch-Institute (RKI) and consists of regularly conducted nationwide surveys among children, adolescents, and young adults aged 0 to 29 years and living in Germany. KiGGS Wave 2 was conducted between 2014 and 2017. The Motorik-Modul Study (MoMo) is a submodule of the KiGGS study and aims to assess physical fitness, PA, as well as determinants of PA in children and adolescents (Woll et al., 2017). The whole study sample was drawn from the German resident population aged 4 to 17 years using a two-stage cluster sampling approach. Informed consent to participate in the study was obtained from all parents of the participants. Also, participants from the baseline study (2003–2006) and Wave 1 (2009–2012) were reinvited. A detailed description of the study design and sampling procedure can be found elsewhere (Hoffmann et al., 2018; Mauz et al., 2019; Woll et al., 2017). KiGGS and MoMo provide nationally representative data of PA and sedentary behavior of children, adolescents, and young adults living in Germany. A positive vote of the ethics committee of Karlsruhe Institute of Technology of September 23, 2014, is available for the study.

5.3.2 Sample description

For the current analysis, cross-sectional data of participants aged 6 to 17 years from KiGGS and MoMo Wave 2 (2014–2017) were used (n = 2743). A detailed dropout description can be found elsewhere (Burchartz et al., 2020). To investigate the direct comparison of both measurement methods within each participant, only participants with complete valid device-based (using accelerometer) data as well as self-reported (via physical activity questionnaire) PA data were included in the analyses. The final sample consisted of n = 2236 children and adolescents (meanage = 12.5, SD = 3.3; Table 5.1). The sample was divided into three age groups (6–10 yrs. n = 698, 11–13 yrs. n = 694, 14–17 yrs. n = 844) as well as two gender groups (boys n = 1050, girls n = 1186). The sample revealed no gender differences in age, weight, height, or BMI.

5.3.3 Measures

5.3.3.1 Self-reported PA data – physical activity questionnaire (PAQ)

The MoMo-PAQ is a self-administered questionnaire with 28 items (Bös et al., 2002). The reliability and validity of the MoMo-PAQ were found to be comparable to those of other activity questionnaires internationally used to assess PA for the same age group. Data obtained with the MoMo-PAQ are sufficiently reliable (test-retest reliability: ICC = 0.68), but correlation coefficients with accelerometry data are low (r = 0.29) (Jekauc et al., 2013). The goal of the MoMo-PAQ is a domain-specific quantification of PA in minutes per week at different intensity levels (low, moderate, and vigorous) in four domains: PA in everyday life, PA in school, leisure time in sports clubs, and leisure time outside of sports clubs. For the present study, the guideline adherence question “On how many days of a normal week were you/was your child physically active for at least 60 minutes” was analyzed (Prochaska et al., 2001). Parents (or legal guardians) of the 6- to 10-year-olds helped to complete the self-administered questionnaires, the 11- to 17-year-olds did so themselves. The answer categories ranged from 0 to 7 days. This question refers to the internationally agreed PA criterion of at least 60 min of moderate to vigorous PA per day that is promoted by the World Health Organization (WHO) (Corbin et al., 2004; World Health Organization, 2010).

Table 5.1: Characteristics of participants: Number of participants as n(%). Age in years, weight in kg, height in cm, body mass index in $kg \cdot m^{-2}$, data are presented as mean ± SD

Variable	Overall			6–10 years old		11–13 years old		14–17 years old	
	All	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
n, (%)	2236 (100)	1050 (47)	1186 (53)	350 (15.7)	348 (15.6)	321 (14.4)	373 (16.7)	379 (16.9)	465 (20.8)
Age, yr	12.5 ± 3.3	12.4 ± 3.3	12.7 ± 3.3	8.5 ± 1.4	8.5 ± 1.4	12.5 ± 0.9	12.5 ± 0.8	15.9 ± 1.1	16.0 ± 1.1
Weight, kg	48.4 ± 17.5	48.4 ± 19.0	46.7 ± 15.7	30.1 ± 7.3	30.1 ± 8.3	48.3 ± 11.8	47.2 ± 10.6	67.1 ± 13.7	59.4 ± 11.0
Height, cm	154.2 ± 18.0	155.5 ± 20.0	152.9 ± 16.0	133.7 ± 9.6	133.2 ± 10.9	157.4 ± 9.9	156.6 ± 8.6	176.0 ± 8.9	165.5 ± 6.3
BMI, $kg \cdot m^{-2}$	19.3 ± 3.8	19.2 ± 3.8	19.3 ± 3.8	16.7 ± 2.4	16.7 ± 2.7	19.3 ± 3.3	19.1 ± 3.2	21.6 ± 3.8	21.7 ± 3.5

5.3.3.2 Device-based, measured PA data - accelerometer

For assessment of device-based, measured PA, Acti- Graph GT3X accelerometers were used, as they were found to be reliable and valid devices to monitor PA in children and adolescents (O'Neil et al., 2014; Puyau et al., 2002; Treuth et al., 2004; Tryon, 2005; Wood et al., 2008). The technical and methodological details of the present study, which are required when accelerometers are used in epidemiological studies (suggested by Migueles et al. (2017)), can be found elsewhere (Burchartz et al., 2020). A modified summary of the Burchartz et al. (2020) setup procedures can be found in Table 5.2. For the present study, minutes of moderate to vigorous-intensity physical activity (MVPA) per day were calculated. Each day was categorized as either meeting the guideline (MVPA $>$ 60 min.) or not meeting the guideline. The resulting variable, therefore, ranges from 0 to 7.

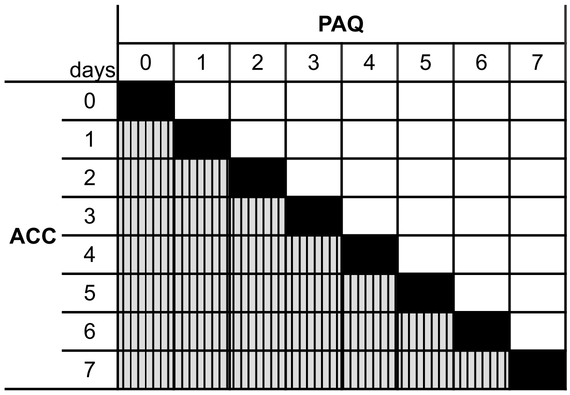
Table 5.2: Expanded list of accelerometer criteria used in KiGGS and MoMo, modified from Burchartz et al. (2020)

Criteria	Definitions within this study
Accelerometer devices	ActiGraph (models: GT3X+, wGT3X-BT)
Placement of the device	Laterally on top of the right anterior superior iliac spine
Sampling frequency	30 hertz
Filter	Normal ActiGraph GT3X filter
Epoch lengths	1 second with possibility to convert to 5, 10, 15, 30, and 60 seconds
Non-wear time definition	Choi et al. (2011): 90-minute time window for consecutive zero/nonzero counts; allowance of 2-minute intervals of nonzero counts with an up/downstream 30-minute consecutive zero counts window
Valid days/valid weeks	8 h of recordings on four weekdays and one further weekend day when wearing the device for 7 days
Population age range	Children, adolescents and young adults from 6 to 17 years
Sedentary and physical activity intensity classification and cut point algorithms	6-10 years: Evenson et al. (2008) 11-17 years: Romanzini et al. (2014)

5.3.4 Statistical analysis

Statistical analyses were performed using SPSS Statistics 24 (IBM Corporation, Armonk, NY). Differences in the number of days with 60 min of MVPA (dependent variable) between the assessment methods (independent variables) were calculated. Repeated measures ANOVAs were calculated for assessment methods (PAQ and ACC) with between-subjects factors age and gender to reveal the effects of assessment and interaction. Mean values, standard deviations, as well as p-values, and effect sizes (f) are given for the analysis of variance. 0.1 stands for a small effect, 0.25 for a medium, and 0.4 for a large effect (Cohen, 1988). The statistical significance level was set to .05. 95% CI were calculated for the comparison of all age and gender groups. Crosstable heat maps were chosen to display two-dimensionally varying PAQ and ACC values of WHO guideline adherence by the same subject. Those heat maps are color-coded (white-grayblack for low to high values) for each cell to visually highlight patterns in row-column interactions. Figure 1 symbolizes compliance with the standards as shown by the PAQ and the ACC. When a participant met the criteria outlined in the PAQ, he/she can be found on the black-colored diagonal. When the participant reached more days of at least 60 min of MVPA with the ACC than given in the PAQ, he/she will be found below the diagonal in one of the vertically lined cells. Participants reaching 60 min MVPA on fewer days than stated in the PAQ will be found in the white cells above the diagonal. Besides, the difference in days between the two methods was calculated and a t-test was used to detect differences between genders in all age groups. Here the level of significance was set to $\alpha = 0.01$.

Figure 5.1: Cross table heat map - black cells: Same results in the PAQ and ACC; white cells: More days with the PAQ; vertically lined cells: More days with the ACC



5.4 Results

Reaching the WHO Guidelines of at least 60 min MVPA every day is significantly lower when measured with the ACC than with the PAQ. 9% of the participants met the WHO Guidelines on each day per week when measured with the PAQ, 4% with the ACC. According to the PAQ, more than 60 min MVPA is achieved on a mean$_{PAQ}$ = 3.82 days. When measured with ACC, this goal is achieved on a mean$_{ACC}$ = 2.34 days (F_{method} = 915.85; p = <.001; f_{Cohen} = .64) (compare Table 5.3). Besides the differences in the detection method, smaller but still large effect sizes can be found between age groups (F_{age} = 264.2, p < .001; f_{Cohen} = .480). Especially the 6–10-year-olds adhere to the WHO Guidelines significantly more often (16% PAQ, 11% ACC, p = .000) on every day of the week compared to 11–13-year-olds (7% PAQ, 2% ACC, p < .001) and 14–17-year-olds (4% PAQ, 1% ACC, p < .001). In gender, medium effect sizes can be found. 11% (PAQ) and 7% (ACC) of boys reached the WHO standard on each day of a week. Girls reach the target on a much lower percentage of days (7% PAQ, 2% ACC, F_{gender} = 134.8, p < .001; f_{Cohen} = .25). In the PAQ 68% of all participants stated that they were active on more days than detected with the ACC (upper right corner of the heat map). For 13% of the participants, the answers from the PAQ and the ACC matched, whereas 19% met the Guidelines on more days than stated in the PAQ (see Fig. 5.2a)). The youngest age group shows a different pattern in the heat maps. The remaining groups show patterns similar to that of the overall sample, which is why these two are shown here. Heat maps for all groups differentiated by age and gender can be found in the supplementary material. The distribution in the cross table heat maps shows that boys and girls as well as the 11–13-year-olds and the 14–17-year-olds (see supplementary material) did not differ much in distribution from the whole sample. As noted above, girls reached the Guideline on fewer days than boys. Figure 3 shows the differences in the various age groups. The non-overlapping confidence intervals in Table 3 show that gender has a significant influence on the results, but a much bigger influence, and the larger variance can be explained by age. The 6–10-year-olds differ most from the other groups (see Fig. 5.2b and supplementary material). 52% of these youngest participants stated more days with 60 min of PA in the PAQ than measured by the ACC. That is 16% less than the average across all participants. Also, the proportion of participants aged 6–10 years with more PA measured by the ACC as stated in the PAQ almost doubled to 33% in contrast to the older participants. The distribution is much smoother with peaks now lying in the estimation corridor (number of days PAQ = number of days ACC), deviations of $\pm$ one day, and only a slight overestimation for the PAQ. Figure 3 shows more detailed results for the mean differences of the age groups separated by genders. Here, the difference between boys and girls (6–10 years) in reaching the WHO Guidelines with PAQ and ACC is significant. The youngest boys (meanDiff$_{boys}$ = 0.3d; SD = 2.5) nearly reach their PAQ results with the ACC. The difference in girls (meanDiff$_{girls}$ = 1.2d; SD = 2.6) is almost one day higher in the PAQ than in boys. A significant difference of half a day can be found between genders for 11–13-year-olds (meanDiff$_{boys}$ = 1.4d; meanDiff$_{girls}$ = 2.0d;). No significant difference was found between 14 and 17-year-old boys and girls (meanDiff$_{boys}$ = 1.9d; meanDiff$_{girls}$ = 2.0d).

Table 5.3: Numbers of days with 60 min of moderate- to vigorous-intensity physical activity (MVPA) as obtained from accelerometer (ACC) measurements and self-reported (PAQ) PA, data are presented as mean and 95%-CI

Group	ACC		PAQ	
	Mean	**95%-CI**	**Mean**	**95%-CI**
Overall	2.42	(2.35–2.50)	3.86	(3.79–3.93)
Boys	2.89	(2.78–3.00)	4.05	(3.95–4.15)
Girls	1.95	(1.85–2.06)	3.67	(3.57–3.76)
6–10 years	3.66	(3.53–3.79)	4.39	(4.26–4.51)
11–13 years	2.08	(1.95–2.22)	3.76	(3.63–3.88)
14–17 years	1.53	(1.40–1.65)	3.43	(3.32–3.54)
Boys				
6–10 years	4.29	(4.10–4.48)	4.54	(4.37–4.72)
11–13 years	2.53	(2.33–2.73)	3.94	(3.76–4.13)
14–17 years	1.85	(1.67–2.03)	3.67	(3.50–3.83)
Girls				
6–10 years	3.03	(2.84–3.22)	4.23	(4.06–4.41)
11–13 years	1.63	(1.45–1.82)	3.57	(3.40–3.74)
14–17 years	1.20	(1.03–1.36)	3.20	(3.05–3.35)

Figure 5.2: Cross table heat map – numbers of days with MVPA ≥60 min; ACC*PAQ a) overall sample b) age: 6–10-year-olds; % of participants; heat maps are color-coded (white/min < gray/moderate < black/max)

a)

		PAQ								
	days	0	1	2	3	4	5	6	7	n
ACC	0	1%	3%	6%	6%	4%	4%	1%	1%	25%
	1	0%	1%	4%	5%	4%	3%	2%	1%	20%
	2	0%	1%	3%	4%	3%	2%	1%	1%	16%
	3	0%	1%	1%	3%	3%	2%	1%	1%	13%
	4	0%	0%	1%	2%	2%	2%	2%	2%	11%
	5	0%	0%	1%	1%	2%	2%	1%	1%	8%
	6	0%	0%	0%	1%	1%	1%	1%	1%	5%
	7	0%	0%	0%	1%	0%	1%	0%	1%	3%
	n	1%	6%	16%	23%	18%	17%	9%	9%	100%

max

min

b)

		PAQ								
	days	0	1	2	3	4	5	6	7	n
ACC	0	0%	1%	1%	2%	1%	2%	1%	1%	9%
	1	0%	0%	1%	3%	3%	3%	2%	1%	12%
	2	0%	1%	2%	2%	2%	3%	1%	2%	12%
	3	0%	1%	1%	3%	2%	2%	1%	1%	11%
	4	0%	0%	2%	4%	2%	3%	3%	3%	16%
	5	0%	1%	1%	2%	3%	3%	2%	3%	15%
	6	0%	0%	1%	2%	2%	3%	2%	3%	14%
	7	0%	0%	1%	2%	2%	2%	1%	2%	11%
	n	1%	4%	9%	20%	18%	19%	13%	16%	100%

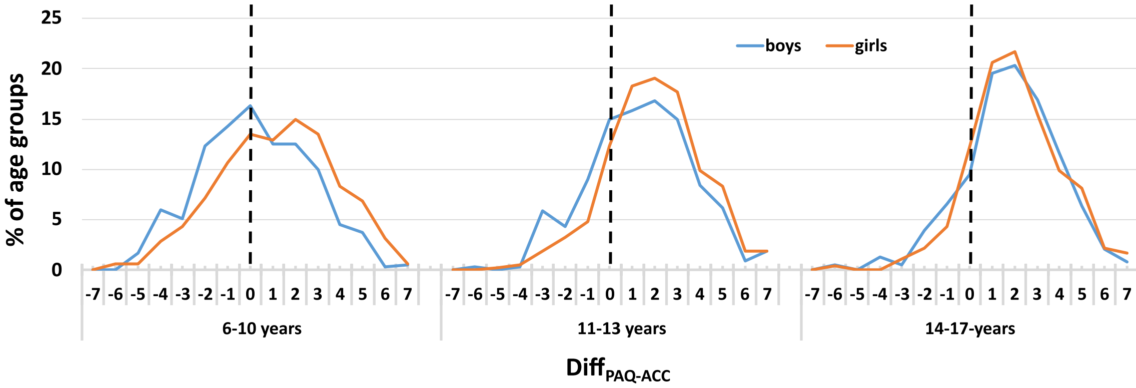
max

min

5.5 Discussion

The study aimed to investigate how children and adolescents in Germany differ in reaching the WHO Guideline of at least 60 min MVPA per day depending on whether PA was self-reported or measured by accelerometer. As expected, the PAQ values were higher than those measured by the ACC but still, both values are alarmingly low. The low overall adherence to the PA Guideline as obvious from the PAQ can also be found in the results of a recently published pooled data study by the WHO, according to which less than 15% of school-going adolescents aged 11–17 did meet the Guidelines (Guthold et al., 2020). The lower PA Guideline adherence as measured by the ACC is consistent with the findings of (Adamo et al., 2009; Dyrstad et al., 2014; Troiano et al., 2014) who also reported less light to moderate PA determined by device-based measurements compared to questionnaires. The difference between self-reported and device-based measured PA might be because the PAQ only asks for physically active time, which is subjective and depends on the physical fitness of the participant, as was stated by other studies before (Jekauc et al., 2013; Slootmaker et al., 2009; Wijndaele et al., 2015). Additional, well-known factors that influence the validity of self-reports are recall bias and social desirability (Dyrstad et al., 2014). The differences between the results of both methods

Figure 5.3: Differences between numbers of days with MVPA ≥60 min measured by PAQ and ACC (Diff$_{PAQ-ACC}$) for the three age groups by gender in %, mean difference for boys and girls, including SD. Zero on the X-axis means the same number of days with MVPA ≥60 min measured by PAQ and ACC



are significantly smaller for younger children than in the older age groups. Differences between age and gender groups are found in both methods. The older the subjects are, the lower is the proportion of those who meet the WHO Guideline on each day, with girls meeting the guideline less frequently than boys in all age groups. A closer look at the differences between genders in the youngest age group reveals that boys almost match their answers given in the questionnaire with a difference of 0.2 days. The previously mentioned overestimation of PA by the questionnaires (Jekauc et al., 2013; Slootmaker et al., 2009; Wijndaele et al., 2015) cannot be found here. The significant difference between the youngest age group and the older ones may be caused by the fact that in this group an external observer (usually a parent) fills out the questionnaire together with the child and may therefore be better able to assess the activity. (Sallis et al., 2002) found that parent-reported MVPA corresponds to one to twothirds of the child's activity measured by the accelerometer. This could be a clue why the gap between the methods is smaller. A more plausible explanation is that the activity patterns of children are more spontaneous, impulsive, and of shorter duration (Gabrys et al., 2015; Graf et al., 2013). These short activities, when measured in total, often result in a small amount of light or moderate activity and are poorly captured by questionnaires (Shephard, 2003). Accelerometers, by contrast, register these short and spontaneous movements which are sometimes overlooked when filling out the questionnaire. This does not lead to an overestimation in the questionnaire for small children, but an underestimation due to the short, spontaneous movements not recorded. Since accelerometers measure these movements, this would explain the reduced difference between the two methods. This is also reflected by the increasing difference between the methods used in the older age groups. The older the participants are, the more structured is their everyday life and the less spontaneous movements are registered by the accelerometer. The familiar overestimation by the questionnaire reoccurs. The older the participants were, the less often they reached the WHO recommendation of 60 min MVPA per day. The lower adherence in older age groups is confirmed by findings of (Cooper et al., 2015; Riddoch et al., 2004) in other European countries and might be due to longer times spent at school or work. In comparison to boys, girls reach the 60 min MVPA on fewer days in all age groups, but the difference decreases to about half of the initial value with increasing age. The difference in gender and the lower adherence for girls is consistent with the worldwide gender gap of physical activity reported in Guthold et al. (2020), Mielke et al. (2018), and Riddoch et al. (2004). Mielke et al. (2018) found a similar prevalence of inactivity in women and men in a study based on worldwide data of the WHO. Normally, girls and boys might be expected to be equally active until puberty, with the gap starting to open up at this point in time. However, different interests may probably be the reason for this earlier gap - girls tend to be sociable and do esthetic sports, while boys tend to romp, scuffle, and do run-intensive sports. Further examination of the data on PA intensity and sport disciplines in MoMo could give a more detailed answer as to where this difference comes from and where interventions could be useful to close this gap.

5.5.1 Strength and limitations

The present study is limited to its observational nature and we do not intend to infer causality from paralleled trends or significant correlations. The main goal of MoMo is to track and report PA and fitness of children and adolescents in a nationwide sample, and significant effort was put into collecting representative data from 167 sample points all over the country. A major strength is the large number of participants and recording of physical activity of each participant by PAQ and ACC. However, this also leads to the restriction that the PAQ assessed PA of an average week, whereas the ACC measured PA during one specific week. An additional comprehensive and elaborate diary was avoided during the week by wearing an ACC. Study participants carried the ACC following the completion of the already very time-consuming fitness test and surveys on activity and health. Even though the wearing times were very long on average, some participants told us that wearing had been prohibited in some sports competitions like soccer. Wearing electronic devices was forbidden to prevent trainers from having an unfair advantage in knowledge. Even if documented by the non-wear time protocol (Burchartz et al., 2020), the unrecorded activities could not be taken into account retrospectively. The manual input of the data from the handwritten non-wear protocol is very timeconsuming. Besides, the information in the protocols is very inconsistent and manual input would distort the acceleration data. This missing data could be another link to the difference between PAQ and ACC results in this study. A wearing time of 24 h and a consistent ambulatory assessment for the non-wear time could solve this problem in future studies. A check of the WHO guidelines is easy to implement with an accelerometer at first glance. However, evaluation results in a multitude of possibilities for implementation. When examining the average time spent with physical activity each week (as now recommended by the new WHO Guidelines of 2020 (Chaput et al., 2020)), days with activity times longer than 60 min would compensate for those with less activity (Colley et al., 2017). Still, the daily stimulus is very important in children (Dwyer et al., 1983). This study determined whether the subject was active for at least 60 min or not on each day individually. To look at the exact times spent with MVPA on every single day will result in fewer days of at least 60 min MVPA when both evaluation methods are compared (Colley et al., 2017). The main reason, however, since the study was already planned and started in 2014, the questions in the questionnaire still referred to the 2010 WHO Guidelines (World Health Organization, 2010). Only now the recommendations on youth activity have changed from a recommendation of at least 60 min per day to a recommendation of an average of 60 min per day (Chaput et al., 2020). This adaptation will require changes in survey questions and sampling methods for future monitoring. However, changing the question wording is unlikely to address the need for PA monitoring among children who, especially at young ages, are unable to answer a complex question about average behavior over the past few days, weeks, or months. In the future, this may require the use of proxy reports from multiple respondents, including parents and teachers, though both may also miss observing large portions of the day (Troiano et al., 2020). The alternative of asking daily duration for an entire week may be more accurate but increases survey response time. Therefore, measuring daily PA remains a strength of the portable devices for now, and adapting the questionnaires to the new WHO guideline remains a real challenge. However, we have looked at the accelerometer data with the background of the new WHO guideline. It should be noted, however, that these results cannot be compared with the results of the question used in this study about the number of days on which the subjects have MVPA for more than 60 min. However, if one compares the number of subjects meeting the old versus the new guideline based on the accelerometer data, we see that the percentage increases from 3% to a full 34% of the study participants. This means that 31% of the participants who do not reach 60 min MVPA on all days still have days in the week on which they do so much physical activity that these outweigh the remaining days under the new guideline compared to the old one. This drastically reduces the proportion of children and adolescents who are too inactive, which is also likely to cause some political controversy in the future. By using an epoch length of one second in MoMo, short activities can be recorded with the accelerometers. This could be another reason why younger children have a more consistent PA outcome with both methods. These short activities are less frequent for older children, which is associated with the fact that an increasing number of older children only practice organized sports. According to MoMo data from previous waves, organized PA in extracurricular activities and sports clubs increased by 8 %, while unorganized PA decreased by 7 % (Schmidt et al., 2017). Apart from PA, many other parameters were collected in MoMo. This results in a multitude of evaluation options, and further examination of the data in MoMo (such as PA intensity, sports disciplines, socioeconomic status, migration status, etc.) will give a more detailed answer as to the reasons behind the differences regarding age and gender.

5.6 Conclusion

Children and adolescents living in Germany and examined within the MoMo and KiGGS studies show a very low adherence to the recommendations given in the 2010 WHO Guideline (World Health Organization, 2010). These results were confirmed by both survey methods. Surprisingly, the differences in meeting the Guideline between the measurement methods are much smaller for younger children than for older age groups. Future studies should take a deeper look into the underlying cause and verify whether short and spontaneous movements reduce the gap between the methods. Continuous overestimation of the self-report in contrast to the accelerometer was observed in all other age groups. With increasing age, the percentage of compliance with the 2010 WHO Guidelines was found to decreasing, with girls reaching the target with a significantly lower percentage in all age groups. As this study only used PA data, no statements can be made for the underlying cause. The large number of participants that did not reach the WHO Guidelines, however, suggests that PA interventions and further monitoring as well as further analysis of the MoMo data are required. This also is the conclusion drawn by the WHO. It recommends all countries adopt policies and programs to increase the PA of children and adolescents, especially girls (Guthold et al., 2020). Further examination of the data on PA intensity and sport disciplines in MoMo could give a more detailed answer to the reasons behind the differences regarding age and gender. Having both self-reported and device-based, measured data will help explain the observed population differences (Corder & van Sluijs, 2010). However, the true value of physical activity probably lies somewhere in between these two methods. The WHO Guideline adaptation in 2020 (Chaput et al., 2020) will require changes in survey questions and sampling methods for future monitoring. To record a more accurate activity profile, a combination of both methods might be a solution. This could be an algorithm to subtract or add the methodological difference if only one method is used. A solution not to loose the activity during non-wear time could be the use of ambulatory assessment in combination with 24 h recording. Triggered e-diaries may ask the subjects for the type of activity performed after certain events have been detected (e.g. device not worn, periods of high activity, or sedentary behavior). Then, non-wear times, their reasons, and activities performed while not wearing the device can be considered uniformly. The participant may be given feedback on how much activity was not recorded and how much was missing to reach the Guidelines. Ambulatory assessment can also be used to ensure that the activity time frames for both methods are consistent by answering the PAQ on the mobile phone at the end of the survey. In this way, the strengths of one method would compensate for the limitations of the other method.

Supplementary Information

This article contains supplementary material available in Chapter A.1.

Authors' contributions

AB was responsible for the overall conception and design of this manuscript. DO & SS supported the process of writing the manuscript as well as the statistical analysis and interpretation of data. AB, CN, DO, SS, KM were responsible for the acquisition of the data. SK, KW, KM, CN were responsible for the critical revision. SK was also involved in drafting the manuscript. CN & AW contributed to the design of the study. All authors read and approved the final manuscript.

References Comparison Article

Adamo, K. B., Prince, S. A., Tricco, A. C., Connor-Gorber, S., & Tremblay, M. S. (2009). A comparison of indirect versus direct measures for assessing physical activity in the pediatric population: A systematic review. *International journal of pediatric obesity : IJPO : an official journal of the International Association for the Study of Obesity*, *4*(1), 2–27. https://doi.org/10.1080/17477160802315010

Bös, K., Opper, E., & Woll, A. (2002). *Fitness in der Grundschule: Förderung von körperlich-sportlicher Aktivität, Haltung und Fitness zum Zwecke der Gesundheitsförderung und Unfallverhütung; Endbericht*. Bundesarbeitsgemeinschaft für Haltungs-u. Bewegungsförderung eV. https://www.schulsport-nrw.de/fileadmin/user_upload/sicherheits_und_gesundheitsfoerderung/pdf/Fitness_in_der_Grundschule.pdf

Burchartz, A., Anedda, B., Auerswald, T., Giurgiu, M., Hill, H., Ketelhut, S., Kolb, S., Mall, C., Manz, K., Nigg, C. R., Reichert, M., Sprengeler, O., Wunsch, K., & Matthews, C. E. (2020). Assessing physical behavior through accelerometry – State of the science, best practices and future directions. *Psychology of Sport and Exercise*, *49*, 101703. https://doi.org/10.1016/j.psychsport.2020.101703

Chaput, J.-P., Willumsen, J. F., Bull, F. C., Chou, R., Ekelund, U., Firth, J., Jago, R., Ortega, F. B., & Katzmarzyk, P. T. (2020). 2020 WHO guidelines on physical activity and sedentary behaviour for children and adolescents aged 5-17 years: summary of the evidence. *The international journal of behavioral nutrition and physical activity*, *17*(1), 141. https://doi.org/10.1186/s12966-020-01037-z

Choi, L., Liu, Z., Matthews, C. E., & Buchowski, M. S. (2011). Validation of accelerometer wear and nonwear time classification algorithm. *Medicine and science in sports and exercise*, *43*(2), 357–364. https://doi.org/10.1249/MSS.0b013e3181ed61a3

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. ed.). Erlbaum. http://www.loc.gov/catdir/enhancements/fy0731/88012110-d.html

Colley, R. C., Carson, V., Garriguet, D., Janssen, I., Roberts, K. C., & Tremblay, M. S. (2017). *Physical activity of Canadian children and youth, 2007 to 2015*. Statistics Canada.

Cooper, A. R., Goodman, A., Page, A. S., Sherar, L. B., Esliger, D. W., van Sluijs, E. M. F., Andersen, L. B., Anderssen, S. A., Cardon, G., Davey, R., Froberg, K., Hallal, P. C., Janz, K. F., Kordas, K., Kreimler, S., Pate, R. R., Puder, J. J., Reilly, J. J., Salmon, J., … Ekelund, U. (2015). Objectively measured physical activity and sedentary time in youth: the International children's accelerometry database (ICAD). *The international journal of behavioral nutrition and physical activity*, *12*, 113. https://doi.org/10.1186/s12966-015-0274-5

Corbin, C. B., Pangrazi, R. P., & Le Masurier, G. C. (2004). Physical activity for children: Current patterns and guidelines. *Journal of Physical Activity and Health*, *1*(3), 281.

Corder, K., & van Sluijs, E. M. F. (2010). Invited commentary: comparing physical activity across countries–current strengths and weaknesses. *American journal of epidemiology*, *171*(10), 1065–1068. https://doi.org/10.1093/aje/kwq068

Dwyer, T., Coonan, W. E., Leitch, D. R., Hetzel, B. S., & Baghurst, R. A. (1983). An investigation of the effects of daily physical activity on the health of primary school students in South Australia. *International journal of epidemiology*, *12*(3), 308–313. https://doi.org/10.1093/ije/12.3.308

Dyrstad, S. M., Hansen, B. H., Holme, I. M., & Anderssen, S. A. (2014). Comparison of self-reported versus accelerometer-measured physical activity. *Medicine and science in sports and exercise*, *46*(1), 99–106. https://doi.org/10.1249/MSS.0b013e3182a0595f

Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S., & McMurray, R. G. (2008). Calibration of two objective measures of physical activity for children. *Journal of sports sciences*, *26*(14), 1557–1565. https://doi.org/10.1080/02640410802334196

Gabrys, L., Thiel, C., Tallner, A., Wilms, B., Müller, C., Kahlert, D., Jekauc, D., Frick, F., Schulz, H., Sprengeler, O., Hey, S., Kobel, S., & Vogt, L. (2015). Akzelerometrie zur Erfassung körperlicher Aktivität. *Sportwissenschaft*, *45*(1), 1–9. https://doi.org/10.1007/s12662-014-0349-5

Graf, C., Beneke, R., Bloch, W., Bucksch, J., Dordel, S., Eiser, S., Ferrari, N., Koch, B., Krug, S., Lawrenz, W., Manz, K., Naul, R., Oberhoffer, R., Quilling, E., Schulz, H., Stemper, T., Stibbe, G., Tokarski, W., Völker, K., & Woll, A. (2013). Vorschläge zur Förderung der körperlichen Aktivität von Kindern und Jugendlichen in Deutschland. *Monatsschrift Kinderheilkunde*, *161*(5), 439–446. https://doi.org/10.1007/s00112-012-2863-6

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1·6 million participants. *The Lancet Child & Adolescent Health*, *4*(1), 23–35. https://doi.org/10.1016/S2352-4642(19)30323-2

Hoffmann, R., Lange, M., Butschalowsky, H., Houben, R., Allen, J., Kuhnert, R., Schaffrath Rosario, A., & Gößwald, A. (2018). Querschnitterhebung von KiGGS Welle 2: Teilnehmendengewinnung, Response und Repräsentativität (Journal of Health Monitoring, Ed.). https://doi.org/10.17886/RKI-GBE-2018-015

Jekauc, D., Reimers, A. K., & Woll, A. (2014). Methoden der Aktivitätsmessung bei Kindern und Jugendlichen. *B&G Bewegungstherapie und Gesundheitssport*, *30*(02), 79–82. https://doi.org/10.1055/s-0033-1361578

Jekauc, D., Wagner, M. O., Kahlert, D., & Woll, A. (2013). Reliabilität und Validität des MoMo-Aktivitätsfragebogens für Jugendliche (MoMo-AFB). *Diagnostica*, *59*(2), 100–111. https://doi.org/10.1026/0012-1924/a000083

Mauz, E., Lange, M., Houben, R., Hoffmann, R., Allen, J., Gößwald, A., Hölling, H., Lampert, T., Lange, C., Poethko-Müller, C., Richter, A., Rosario, A. S., von Schenck, U., Ziese, T., & Kurth, B.-M. (2019). Cohort profile: KiGGS cohort longitudinal study on the health of children, adolescents and young adults in Germany. *International journal of epidemiology*, *49*(2), 375–375k. https://doi.org/10.1093/ije/dyz231

Mielke, G. I., da Silva, I. C. M., Kolbe-Alexander, T. L., & Brown, W. J. (2018). Shifting the Physical Inactivity Curve Worldwide by Closing the Gender Gap. *Sports medicine (Auckland, N.Z.)*, *48*(2), 481–489. https://doi.org/10.1007/s40279-017-0754-7

Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Delisle Nyström, C., Mora-Gonzalez, J., Löf, M., Labayen, I., Ruiz, J. R., & Ortega, F. B. (2017). Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports medicine (Auckland, N.Z.)*, *47*(9), 1821–1845. https://doi.org/10.1007/s40279-017-0716-0

Müller, C., Winter, C., Mogwitz, M.-S., & Rosenbaum, D. (2011). Validierung von neun Bewegungssensoren bei Kindern und jungen Erwachsenen. *Sportwissenschaft*, *41*(1), 8–15. https://doi.org/10.1007/s12662-011-0167-y

O'Neil, M. E., Fragala-Pinkham, M. A., Forman, J. L., & Trost, S. G. (2014). Measuring reliability and validity of the ActiGraph GT3X accelerometer for children with cerebral palsy: A feasibility study. *Journal of pediatric rehabilitation medicine*, *7*(3), 233–240. https://doi.org/10.3233/PRM-140292

Prochaska, J. J., Sallis, J. F., & Long, B. (2001). A Physical Activity Screening Measure for Use With Adolescents in Primary Care. *Archives of Pediatrics & Adolescent Medicine*, *155*(5), 554. https://doi.org/10.1001/archpedi.155.5.554

Puyau, M. R., Adolph, A. L., Vohra, F. A., & Butte, N. F. (2002). Validation and calibration of physical activity monitors in children. *Obesity research*, *10*(3), 150–157. https://doi.org/10.1038/oby.2002.24

Riddoch, C. J., Bo Andersen, L., Wedderkopp, N., Harro, M., Klasson-Heggebø, L., Sardinha, L. B., Cooper, A. R., & Ekelund, U. (2004). Physical activity levels and patterns of 9- and 15-yr-old European children. *Medicine and science in sports and exercise*, *36*(1), 86–92. https://doi.org/10.1249/01.MSS.0000106174.43932.92

Romanzini, M., Petroski, E. L., Ohara, D., Dourado, A. C., & Reichert, F. F. (2014). Calibration of ActiGraph GT3X, Actical and RT3 accelerometers in adolescents. *European journal of sport science*, *14*(1), 91–99. https://doi.org/10.1080/17461391.2012.732614

Sallis, J. F., Taylor, W. C., Dowda, M., Freedson, P. S., & Pate, R. R. (2002). Correlates of Vigorous Physical Activity for Children in Grades 1 through 12: Comparing Parent-Reported and Objectively Measured Physical Activity. *Pediatric Exercise Science*, *14*(1), 30–44. https://doi.org/10.1123/pes.14.1.30

Schmidt, S. C. E., Henn, A., Albrecht, C., & Woll, A. (2017). Physical Activity of German Children and Adolescents 2003-2012: The MoMo-Study. *International journal of environmental research and public health*, *14*(11). https://doi.org/10.3390/ijerph14111375

Schutz, Y., Weinsier, R. L., & Hunter, G. R. (2001). Assessment of free-living physical activity in humans: An overview of currently available and proposed new measures. *Obesity research*, *9*(6), 368–379. https://doi.org/10.1038/oby.2001.48

Shephard, R. J. (2003). Limits to the measurement of habitual physical activity by questionnaires * Commentary. *British journal of sports medicine*, *37*(>3), 197–206. https://doi.org/10.1136/bjsm.37.3.197

Slootmaker, S. M., Schuit, A. J., Chinapaw, M. J., Seidell, J. C., & van Mechelen, W. (2009). Disagreement in physical activity assessed by accelerometer and self-report in subgroups of age, gender, education and weight status. *The international* *journal of behavioral nutrition and physical activity*, *6*, 17. https://doi.org/10.1186/1479-5868-6-17

Treuth, M. S., Schmitz, K., Catellier, D. J., McMurray, R. G., Murray, D. M., Almeida, M. J., Going, S., Norman, J. E., & Pate, R. R. (2004). Defining accelerometer thresholds for activity intensities in adolescent girls. *Medicine and science in sports and exercise*, *36*(7), 1259–1266.

Troiano, R. P., McClain, J. J., Brychta, R. J., & Chen, K. Y. (2014). Evolution of accelerometer methods for physical activity research. *British journal of sports medicine*, *48*(13), 1019–1023. https://doi.org/10.1136/bjsports-2014-093546

Troiano, R. P., Stamatakis, E., & Bull, F. C. (2020). How can global physical activity surveillance adapt to evolving physical activity guidelines? Needs, challenges and future directions. *British journal of sports medicine*, *54*(24), 1468–1473. https://doi.org/10.1136/bjsports-2020-102621

Tryon, W. W. (2005). The reliability and validity of two ambulatory monitoring actigraphs. *Behavior Research Methods*, *37*(3), 492–497.

Wijndaele, K., Westgate, K., Stephens, S. K., Blair, S. N., Bull, F. C., Chastin, S. F. M., Dunstan, D. W., Ekelund, U., Esliger, D. W., Freedson, P. S., Granat, M. H., Matthews, C. E., Owen, N., Rowlands, A. V., Sherar, L. B., Tremblay, M. S., Troiano, R. P., Brage, S., & Healy, G. N. (2015). Utilization and Harmonization of Adult Accelerometry Data: Review and Expert Consensus. *Medicine and science in sports and exercise*, *47*(10), 2129–2139. https://doi.org/10.1249/MSS.0000000000000661.

Woll, A., Albrecht, C., & Worth, A. (2017). Motorik-Module (MoMo) – the KiGGS wave 2 module to survey motor performance and physical activity. *Journal of Health Monitoring*, *2*(S3), 66–73. https://doi.org/10.17886/RKI-GBE-2017-110

Wood, A. C., Kuntsi, J., Asherson, P., & Saudino, K. J. (2008). Actigraph data are reliable, with functional reliability increasing with aggregation. *Behavior Research Methods*, *40*(3), 873–878.

World Health Organization. (2010). *Global recommendations on physical activity for health*.

6. Typical day and influence of weekend on physical activity

Slightly modified version of the 4^{th} published article

Burchartz, Alexander; Oriwol, D.; Kolb, S.; Schmidt, S. C. E.; Haaren-Mack, B. von; Niessner, C.; & Woll, A. (2022). Impact of weekdays versus weekend days on accelerometer measured physical behavior among children and adolescents: results from the MoMo study.

German Journal of Exercise and Sport Research, 52(2), 218–227.
DOI: 10.1007/s12662-022-00811-4

6.1 Abstract

Structured activities, in which children participate, for example, at school, are consistent and limited in scope. On the contrary, after school or weekend activities involve a wider range of behaviors. Studies have shown that physical activity (PA), measured by accelerometers, is lower on weekends compared to weekdays or school days, while PA does not differ between weekdays. In the present study, we examined accelerometer data of children and adolescents living in Germany for different days of the week and weekends.weekends. The current analysis used cross-sectional data from participants (n = 2743) between 6 and 17 years of age collected between 2014 and 2017. The final valid sample consisted of 2278 children and adolescents divided into three age groups (6–10 years, n = 713; 11–13 years, n = 706; 14–17 years, n = 859) and two gender groups (1072 boys, 1206 girls). Physical behavior, including sedentary behavior, as well as light, moderate, vigorous PA, and wear time were analyzed. The absolute and percentage intensity distributions were evaluated daily. The average wear time was 807min daily from Monday–Thursday with significant deviations from the mean on Friday (+38min), Saturday (–76min), and Sunday (–141min). Absolute moderate to vigorous PA times were lower on weekends than during the week. However, the percentage intensity distribution remained constant over all days. Girls were less physically active and

more sedentary than boys ($F_{1,2272} = 38.3$; p<0.01) and adolescents were significantly less active than younger children ($F_{2,2272} = 138.6$; p< 0.01). Waking times increased with age ($F_{2,2272} = 138.6$; p< 0.01). Shorter awake periods limit possible active times on weekends, resulting in lower PA and sedentary behavior compared to weekdays. The percentage distributions of the different categories of intensity of physical behavior are similar on all weekdays and weekend days. We were unable to find a justification for specific weekend interventions. Instead, interventions should generally try to shift activity away from sedentary behavior toward a more active lifestyle.

Key Words: Wear time, School days, Weekend interventions, Physical exercise, Health behavior

6.2 Background

Recently, the World Health Organization (WHO) conducted a pooled data study (Guthold et al., 2020) and analyzed global trends in insufficient physical activity (PA). Using cross-sectional studies from 146 countries, they analyzed 1.6 million participants (ages 11-17 years) and found that a total of 81.0% were not physically active sufficiently worldwide world in 2016, including 77.6% boys and 84.7% girls. Therefore, a large proportion of adolescents do not follow the current recommendations for daily PA, which can affect their current and future health status. Due to the high level of physical inactivity, the WHO has published the action plan 'More active people for a healthier world' for the period from 2018 to 2030(World Health Organization, 2019). Among other things, it has set itself the goal of reducing physical inactivity by 15% by 2030.
Still, little is known about which time periods should be targeted to effectively prevent obesity and improve PA and fitness of school-aged children. So far, the most successful intervention studies for school PA programs have used questionnaires that are subject to recall bias and often lead to overestimation of PA (Dobbins et al., 2013; van Sluijs et al., 2007). However, the use of device-measured physical behavior (PB) is increasing for tracking school-based interventions (Mannocci et al., 2020). Device-based PB recording makes the collected data less vulnerable to social desirability, although it no longer captures the type of PB, but intensity, duration, and frequency (Burchartz et al., 2021).
These intervention studies address activity behaviors on school days, but often focus on changes in moderate to vigorous PA (MVPA) (Ajja et al., 2021; Kolle et al., 2020; Wang, 2019), thus following the WHO guidelines (Chaput et al., 2020; World Health Organization, 2010) that recommend at least 60 min MVPA per day. Time-related factors can be problematic in the school setting, such as competing demands of the education curriculum, the potential overload of teachers through additional PA breaks, or a lack of PA resources and a lack of a PA-supportive school climate. These factors can affect implementation (Naylor et al., 2015). While physical activities during school hours are more consistent, but limited in scope, weekend activities enable a much wider range of behaviors (Fairclough et al., 2012; Mannocci et al., 2020). Studies have shown that children get most of their MVPA during non-structured times, also at school Bailey et al. (2012). Furthermore, large cohort studies found a lower MVPA on weekends compared to weekdays or school days, while PA was comparable on the different weekdays (Brazendale et al., 2021; Corder et al., 2013; Fairclough et al., 2015). Therefore, these studies recommend future PA interventions to target less structured days, such as weekend days, and to provide additional PA opportunities for children. However, PA intervention studies

targeting children on weekends have a lower reach, are less feasible, and may reinforce social disparities, since children with higher socioeconomic status participate in more types of MVPA activity than peers with lower socioeconomic status (Aibar et al., 2014; Love et al., 2019). For this reason, interventions aimed only at weekends should be critically evaluated. As mentioned earlier, the focus of PA interventions often is on changing MVPA levels according to the 2020 WHO guidelines. Regarding the time of the week, Drenowatz et al. (2016) found that longer sleep times on weekends were associated with less time spent sedentary. These extended sleep times indicate shorter awake times on weekends, which should not be ignored.

Hence, the question arises of whether shorter MVPA times on the weekend are simply caused by the shorter awake phase. Or, more specifically, can it be assumed that waking phase during the week is largely determined by the very structured daily school routine, while the weekend time is influenced by the changed sleeping routines? On weekends, the structure of the day is much more self-determined, but getting up later can lead to shorter waking phases. This leaves less time for activity. For example, more children do not reach 60 min of absolute MVPA on the weekend compared to schooldays as reported by Brazendale et al. (2021), Corder et al. (2013), and Fairclough et al. (2015).

More research is needed to better understand how device-based measured PB differs between weekdays/school days and weekends for children and adolescents. In addition, the complete range of PB intensity should be analyzed - not only MVPA but also the proportion of other activities and sedentary intensity. By capturing accelerometer data from 2014–2017, the nationwide Motorik-Modul study (MoMo) collected representative data on PB from children and adolescents in Germany. On this basis, we can now present the first detailed distribution of device-based measured PB levels in Germany. This study examines accelerometer data to determine the differences between weekdays and weekend days. In particular, the different daily patterns of intensity distributions on weekdays versus weekends, as the well as absolute and relative times spent in these intensities, are investigated and compared.

6.3 Methods

6.3.1 Study design

The German Health Interview and Examination Survey for Children and Adolescents (KiGGS) is part of the Federal Health Monitoring System run by the Robert Koch Institute (RKI) and consists of regularly conducted nationwide surveys among children, adolescents, and young adults living in Germany. KiGGS Wave 2 was conducted between 2014 and 2017. The Motorik-Modul study (MoMo) is an in-depth module of KiGGS and aims to assess physical fitness, PA, as well as determinants of PA in children and adolescents (Woll et al., 2017).

The whole study sample was drawn from the German resident population using a two-stage cluster sampling approach. Informed consent to participate in the study was obtained from all parents of the participants. Participants of earlier surveys (baseline study [2003–2006] and Wave 1 [2009–2012]) were re-invited. A detailed description of the study design and sampling procedure can be found elsewhere (Hoffmann et al., 2018; Mauz et al., 2019; Woll et al., 2017). KiGGS and MoMo provide nationally representative data of PA and sedentary behavior (SED) of children, adolescents, and young adults living in Germany. A positive vote of the ethics committee of Karlsruhe Institute of

Technology (KIT) of September 23, 2014, is available for the study. The STROBE statement (Strengthening the Reporting of Observational Studies in Epidemiology) guided the reporting of this study (Vandenbroucke et al., 2007).

6.3.2 Sample description

For the current analysis, cross-sectional data of participants aged 6-to-17 years of KiGGS and MoMo Wave 2 (2014–2017) were used (n = 2743). Children under 6 years of age did not wear an accelerometer. A detailed dropout description can be found elsewhere (Burchartz et al., 2020). The final valid sample that reaches the set wear time threshold (WT) of 8 h on at least four weekdays and one weekend day consists of n = 2278 children and adolescents (cf. Table 6.1). The sample was divided into three age groups (6-10 years. n = 713, 11-13 years. n = 706, 14-17 years. n = 859), as well as two gender groups (boys n = 1,072, girls n = 1,206). Gender was almost equally distributed with females representing 52.9 % of the sample population.

Table 6.1: Participant characteristics

Age group	N	Sex (% female)	Age M ± s (years)	Days with WT >8h M ± s	WT per day M ± s (min.)
6-10 years	713	49.9	08.46 ± 1.43	6.66 ± 0.58	774.2 ± 67.1
11-13 years	706	53.7	12.48 ± 0.85	6.65 ± 0.58	809.6 ± 70.4
14-17 years	859	54.8	15.93 ± 1.12	6.60 ± 0.61	836.1 ± 79.3
All participants	2,278	52.9	12.52 ± 3.30	6.64 ± 0.59	808.5 ± 77.3

Note: **N** Number of participants, **M** mean, **s** standard deviation, **WT** wear time in number of days and minutes per day, **min** minutes.

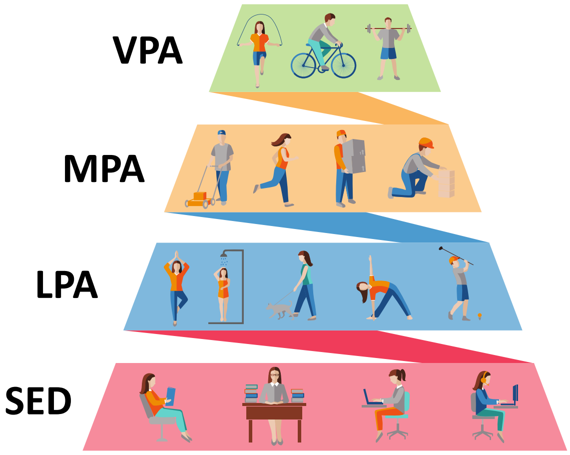
6.3.3 Device-based PA data

ActiGraph GT3X+/wGT3X-BT accelerometers (Actigraph, LLC, Pensacola, FL, USA) were used to assess PA. The technical and methodological details of the present study are described elsewhere (Burchartz et al., 2020; Burchartz et al., 2021). For the present study, mean minutes of SED, light (LPA), moderate (MPA), and vigorous-intensity PA (VPA) were calculated per weekday as well as a percentage distribution of WT (cf. Fig. 6.1). The final data sets included in the analysis are of high quality with an average of 6.64 valid days (WT > 8h) and an average WT (seven days) of 808.5±77.3 minutes per day (more than 13 hours). Subjects were instructed to remove the accelerometers only to sleep or when in contact with water (Burchartz et al., 2020). Therefore, we assumed that the wearing time was almost equivalent to the awake phase of the day. Due to the interrelationship between PA, SED, and sleep, the present manuscript uses the term 'physical behavior' (PB) to refer to these three behaviors (Bussmann & van den Berg-Emons, 2013).

6.3.4 Statistical analysis

All statistical analyzes were performed using IBM SPSS 28 (IBM Corporation, Armonk, NY, USA). Descriptive analyzes stratified by age and gender were performed. Means

Figure 6.1: MoMo – Activity pyramid that associates different physical behaviors to their intensities.



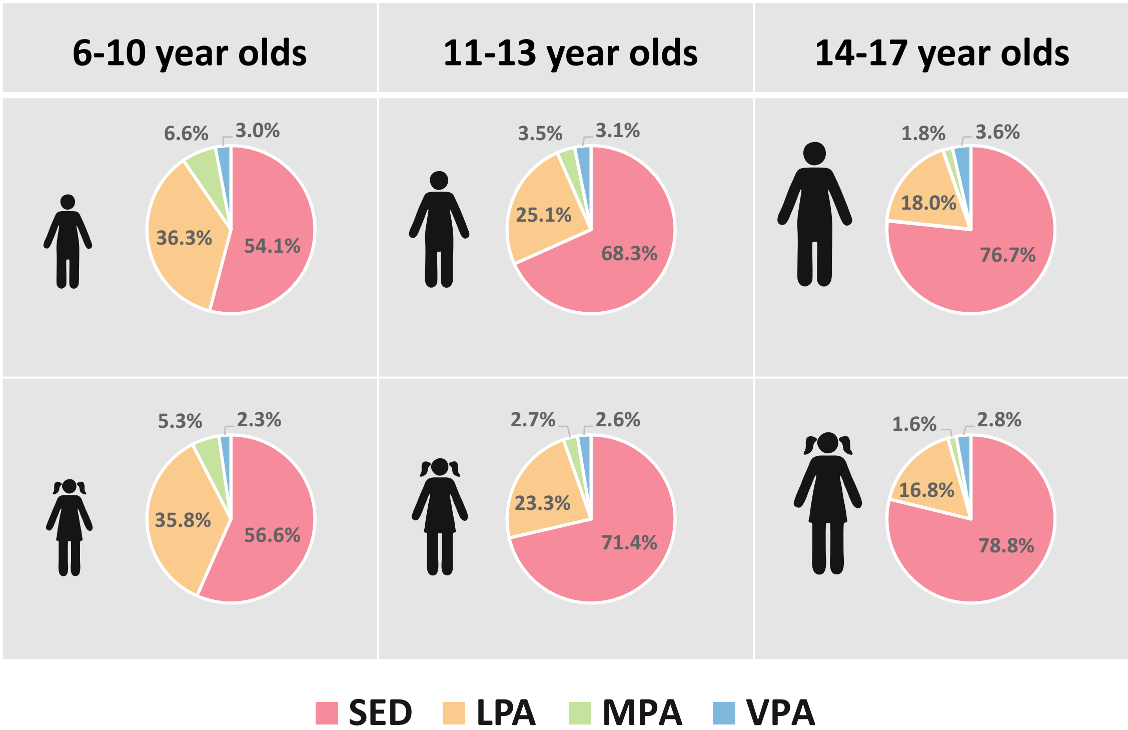
Note: **SED** Sedentary behavior, **LPA** light physical activity, **MPA** moderate physical activity, **VPA** vigorous physical activity. The pictograms represent activities in different intensities. Due to the wide age range in the MoMo study, age-unspecific representations were chosen to the extent possible (e.g.,mowing the lawn symbolizes gardening, bricklaying symbolizes physical work in occupation or education, carrying packages symbolizes carrying lighter objects in contrast to weight-lifting which symbolizes lifting heavier objects).

(M), standard deviations (SD), and percentages were reported, as well as the corresponding inference statistical parameters, including 95 % confidence intervals for differences (CI). Two-way analysis of variance (ANOVA) with post hoc t-test was used to analyze differences between age groups and gender. P values were adjusted using the Bonferroni method for multiple comparisons. For comparison of weekdays and weekend days, Student's t tests for paired samples were used. The effects sizes were calculated using Cohen's d (Cohen, 1988) with $|d| = 0.2$ representing small effects, $| d | = 0.5$ representing medium and $| d | = 0{,}8$ representing large effects.

6.4 Results

WT, in MoMo, the assumed awake phase of their day, increased with age in our sample. Accordingly, the effect of age for WT was significant ($F_{2,2272} = 138.6$; $p<.01$ and explained 10.9% of the variance in WT. No significant difference in WT was found between gender. The Leven test for homoscedasticity was positive ($W = 9.8$) due to the theory-compliant deviation of the standard deviation and the high number of participants in each group.
The average WT was 806 min during all weekdays from Monday to Thursday, except Friday. The difference to Friday was +38 min ($t = 9.85$; $df = 2.276$; $p < .01$; 95% CI: [-46.09, -30.78]; $|d| = .21$). The difference to Saturday was -76 min (medium effect for 14-17 years, $|d| = .51$) and -141 min to Sunday (medium effect for 6-10 years $|d| = .67$, large effects for other age groups, $|d| =.83/.84$; cf. Table 6.2). The average distribution of intensities over the day was: 69% SED, 25% LPA, 3% MPA, 3% VPA (cf. Fig. 6.2). There was a 1% shift from SED to LPA on Saturday and from VPA to SED on Sunday (cf. Table 6.3). Absolute MVPA times were lower on weekends compared to the rest of the week (cf. Fig. 6.3, Table 6.3 and 6.4). However, when looking at the intensity distribution relative to the WT, the intensity remained constant for all days (cf.

Figure 6.2: Typical day (average over all 7 days) of a participant



Note: Percentage distribution of activity intensities during the awake phase of the day for the three age groups 6–10 years ($N_{male} = 357, N_{female} = 356$), 11–13 years ($N_{male} = 327, N_{female} = 379$), and 14–17 years ($N_{male} = 388, N_{female} = 471$) and both genders. **SED** Sedentary behavior, **LPA** light physical activity, **MPA** moderate physical activity, **VPA** vigorous physical activity,

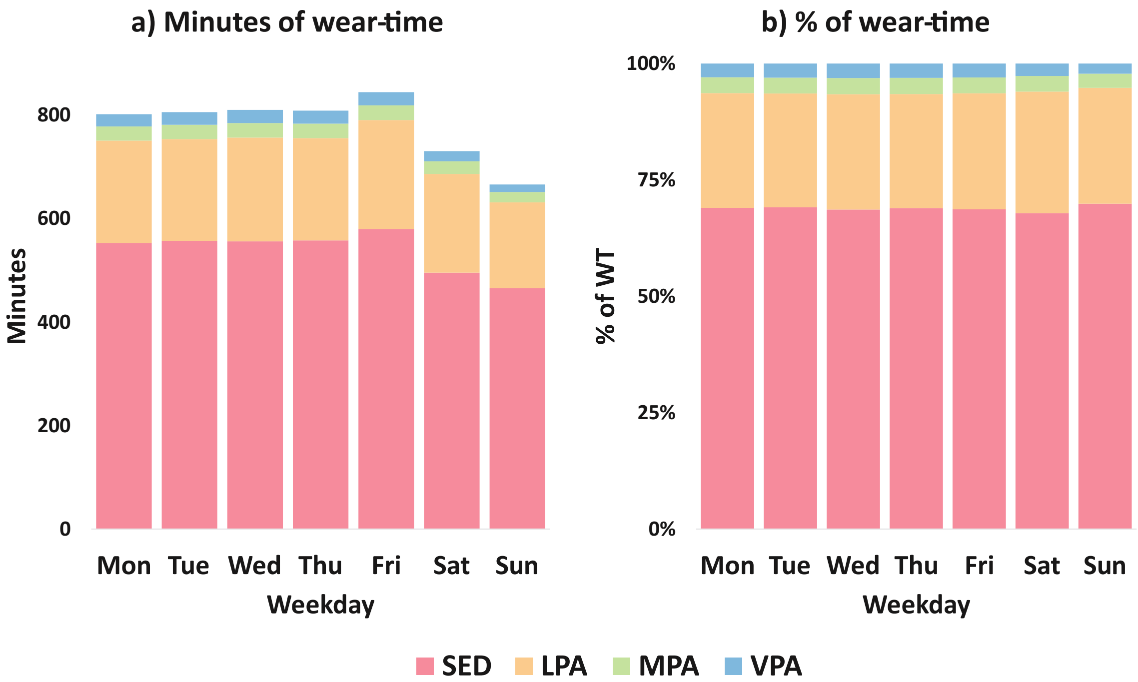
Fig. 6.3 and Table 6.3). The youngest children (6-10 years) spent just over half of the day sedentary, whereas the oldest adolescents (14-17 years) spent almost three-quarters of the day in SED (see Fig. 6.2 and Table 6.3). Age explained 56.7% of the variance of SED ($F_{2,2272} = 1489.7$; $p<.01$). Girls were significantly more sedentary than boys ($F_{1,2272} = 38.3$; $p<.01$). The oldest age group had the longest WT and the highest amount of sedentary time (cf. Fig. 6.4). Younger children had less overall WT, but spent more time in higher intensity activity (cf. Fig. 6.4).

6.5 Discussion

The purpose of this study was to obtain a better understanding of how device-based PB data differ between weekdays and between school days and weekends for children and adolescents. To cover the complete range of PA intensity, not only MVPA, but also the proportion percentage of all aspects of PB intensity as well as WT were calculated. Behavioral patterns were determined per day, gender, and age, and WT was calculated for each day of the week.

The present study revealed main differences between weekdays and weekend days. In general, the absolute WT for weekdays (Monday-Thursday) is 13.4h. While the awake period on Friday is over half an hour longer (14.0h), children and adolescents have less waking time on Saturday (12.2h) and Sunday (11.1h). These differences can be explained by different aspects. During structured school days, participants get up and

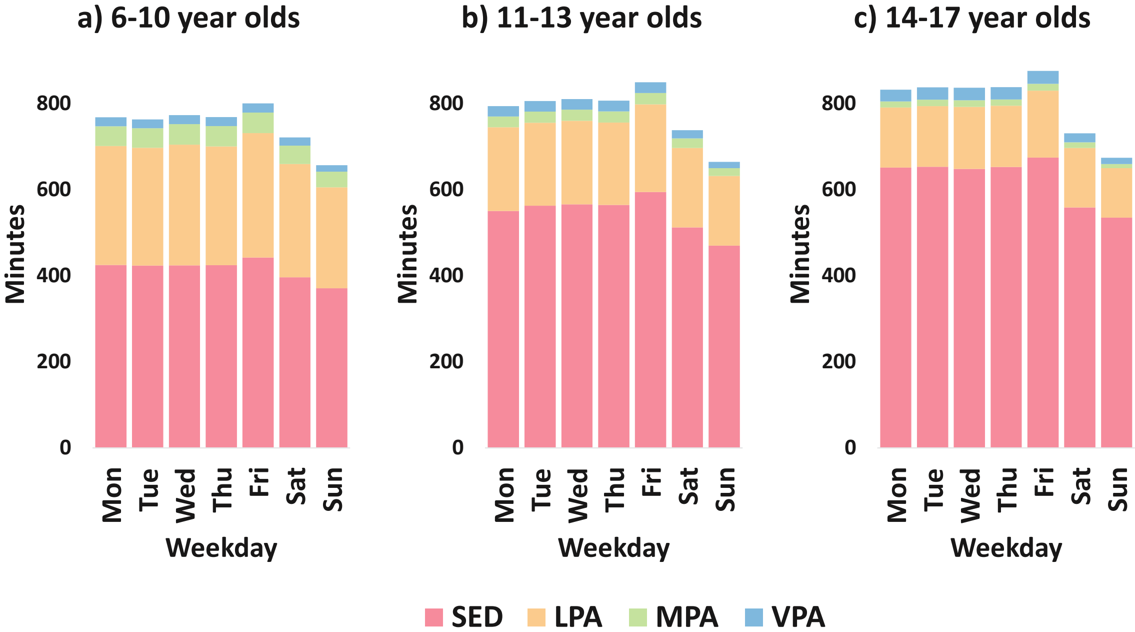
Figure 6.3: Overall (N=2278) mean minutes in different intensities for all weekdays.



Note: **SED** Sedentary behavior, **LPA** light physical activity, **MPA** moderate physical activity, **VPA** vigorous physical activity, **Minutes of wear time** shows the absolute minutes of wear time among intensity classes on weekdays. **% of wear time (WT)** shows the percentage distribution of wear time among intensity classes.

go to bed at similar times (Gradisar et al., 2011). Friday is an exception in terms of going to bed late due to the following weekend Noland et al. (2009). Getting up late on Saturday is relativized by going late to bed (Gradisar et al., 2011; Mishra et al., 2017; Noland et al., 2009). Sunday, however, usually shows the shortest period of awakeness, since after sleeping until late in the morning, the evening ends early to get enough sleep for the following school day and has similar sleep times compared to school nights (Noland et al., 2009). Based on these results, future studies should consider evaluating PB based on the sleep–wake cycle rather than on a 24 hour basis or on the external construct of time of day. In this case, the night from Friday to Saturday may already be counted as a weekend day, and the night from Sunday to Monday as a school-related day. This could have further positive implications in view of the current 24h PB debate. One option would be to determine threshold parameters for the separation of sleep and waking phases, which would then no longer require the use of fixed 24 hour evaluation windows, but would allow the actual influence of sleep and waking times on PB to be investigated. For example, it would be possible to make even more realistic assumptions on how sleep behavior affects psychological factors such as mood and other parts of PB. Furthermore, it could be examined how much influence the day before has in contrast to the day following the sleep phase. Questions that could be addressed in this connection include: Is the night after a day off more attributable to the following school day or to the day off weekend? What is the influence of the night from Friday to Saturday or before a day off, which starts later but also lasts longer, compared to that of a night between two school days or workdays? This would allow for different sleeping patterns to be considered. In contrast to fixed-day analyzes that use periods from midnight to

Figure 6.4: Overall mean minutes in different intensities over weekdays for three age groups.



Note: a) shows the absolute minutes among intensity classes on weekdays for 6–10 years ($N_{male} = 357$,$N_{female} = 356$), b) 11–13 years ($N_{male} = 327$,$N_{female} = 379$), and c) 14–17 years ($N_{male} = 388$,$N_{female} = 471$). **SED** Sedentary behavior, **LPA** light physical activity, **MPA** moderate physical activity, **VPA** vigorous physical activity.

midnight, the actual daily sleep phases could then be taken into account. Therefore, changes in sleep patterns, such as different bedtimes on different weekdays, could be considered.

However, this would require data for the entire 24 hour PB cycle over one or more weeks and cannot be done with the data underlying this study.

As a second result, this study showed similar day patterns for all age groups in terms of intensity distribution. The pattern of the waking phase of the day in the form of WT is also consistent for all groups. The total minutes of WT per day increased with age, which is in line with the observed decrease in sleep hours with increasing age (Chaput et al., 2018). With less sleep at night, waking time during the day increases. The distribution of activity intensity over a day is consistent with previous studies. A 25% increase in SED correlated with age coincided with changes in children's school curricular activities (Crane et al., 2018; Troiano et al., 2008; Williams et al., 2013). Interestingly, the percentage of VPA remains consistent in all age groups at approximately 3%, with girls consistently engaging in fewer VPA than boys. This must be emphasized in the context of the longer waking time of older subjects, as this increases the absolute minutes of VPA by 5 minutes (cf. Fig. 6.3 & Table 6.3). These results are in line with a cross-sectional study by Steele et al. (2010) who found no difference in percentage time of VPA for weekdays and weekends in children measured by accelerometers aged 9-10 years. The same was found by Drenowatz et al. (2016) for young adults and by Aibar et al. (2014) for Spanish and French children. This positive trend in the relatively small increase in absolute VPA is also compensated by a reduction from 6% to below <2% for MPA and a decrease to half of the LPA proportion from 36% to 17% when comparing younger participants (6-10 years) with older ones (14-17 years).

No significant differences in WT were found between boys and girls. However, a sig-

Table 6.2: Differences in minutes of wear time on weekdays and weekend

Age		N	Monday - Thursday	Sunday	Difference				
(years)			**(M ± s)**	**(M ± s)**	**diff.**	**95% CI**	**t**	**p**	**d**
6-10	Ø	713	768.4 ± 83.0	657.0 ± 157.8	99.26	123.63	17.96	<.001	**0.67**
11-13	Ø	706	804.8 ± 89.1	665.5 ± 168.0	126.85	151.62	22.07	<.001	**0.83**
14-17	Ø	859	837.2 ± 109.5	674.4 ± 192.6	149.87	175.68	24.75	<.001	**0.84**
All	Ø	2,278	805.6 ± 99.7	666.2 ± 174.8	132.09	146.73	37.35	<.001	**0.78**

Age		N	Friday	Saturday	Difference				
(years)			**(M ± s)**	**(M ± s)**	**diff.**	**95% CI**	**t**	**p**	**d**
6-10	Ø	713	800.4 ± 179.0	721.6 ± 188.5	62.83	94.81	9.68	<.001	0.36
11-13	Ø	706	850.3 ± 193.2	738.1 ± 227.9	91.87	132.60	10.82	<.001	0.41
14-17	Ø	859	875.1 ± 188.8	731.2 ± 212.4	127.19	160.61	16.90	<.001	**0.58**
All	Ø	2,278	844.1 ± 189.7	730.3 ± 210.3	103.44	123.99	21.71	<.001	0.46

Age		N	Monday - Thursday	Saturday	Difference				
(years)			**(M ± s)**	**(M ± s)**	**diff.**	**95% CI**	**t**	**p**	**d**
6-10	Ø	713	768.4 ± 83.0	721.6 ± 188.5	32.97	60.62	6.64	<.001	0.25
11-13	Ø	706	804.8 ± 89.1	738.1 ± 227.9	50.36	83.03	8.02	<.001	0.30
14-17	Ø	859	837.2 ± 109.5	731.2 ± 212.4	92.17	119.78	15.07	<.001	**0.51**
All	Ø	2,278	805.6 ± 99.7	730.3 ± 210.3	66.78	83.78	17.36	<.001	0.36

Note: Medium and large effects are marked in **bold**.
M mean, **s** standard deviation, **diff.** difference, **95% CI** 95% confidence interval, **t** posthoc t-test; **p** Significant difference, **d** effect sizes calculated using Cohen's d.

nificantly higher proportion of SED and lower proportions of LPA, MPA and VPA were observed on the daily average for girls (see Fig. 6.1 and Table 6.3), with this gender gap repeatedly reported by many studies (Berglind & Tynelius, 2017; Guthold et al., 2018, 2020; Padmapriya et al., 2021; Steele et al., 2010).
Considering our study's results in the context of previous studies, most studies showed longer MVPA time during the week compared to the weekend. Brazendale et al. (2021) found similar results in a study using pooled accelerometer data from the ICAD project (International Children's Accelerometry Database), which revealed that children accumulated approximately 10 minutes more MVPA on school days (up to 17 minutes more in European children) compared to weekend days. Similar results were reported by Corder et al. (2013) and Steele et al. (2010). Some studies found also less time spent in SED on weekend days compared to weekdays (Corder et al., 2013; Frago-Calvo et al., 2018) and concluded that this may be explained by a lack of typically structured SED associated with school lessons. Until now, we are not aware of any PA study that considers significantly lower awake periods on weekends. The present results suggest that the percentage intensity distribution of SED, LPA, MPA, and VPA on the weekend is similar to that during the week, including Friday, despite the lower WT (cf. Fig. 6.2). As can be seen in detail in Table 6.3, there is a small 1% change from SED to LPA on Saturday and a 1% reduction in VPA on Sunday that can be attributed to the increase in SED

compared to weekdays. Hence, lower MVPA and SED times on the weekend cannot be attributed to a fundamental change in PA behavior. On the contrary, the PB patterns observed during the week also prevailed over the weekend. Since the shorter waking phase on the weekend indicates that less time is available overall, this also results in less time spent for the different PB intensities. With this in mind, we were unable to find any justification for specific weekend interventions based on PA distribution.

Instead, we recommend interventions to generally change activity behaviors from SED to a more active lifestyle (more LPA, MPA, and VPA). Due to the distinct differences between the levels of activity during the week and on weekends, previous studies recommended collecting data on both weekdays and weekend days (Trost et al., 2005; Vanhelst et al., 2014). However, some authors disagreed with the inclusion of weekend data in PA analysis (Wolff-Hughes et al., 2016). Although the inclusion of weekend data can introduce bias into PA estimates, most studies typically use measurement periods of four out of seven valid days, including at least one weekend day, to calculate PA (Skender et al., 2016). Due to the results of the present study that the show that durations of the waking phases differ on Saturday and Sunday, we assume that the evaluation of PA may be influenced by the type of weekend day included in the analysis. We suggest that the evaluations continue to include weekend days. Due to the high proportion of valid days (m= 6.64) in the MoMo study, we consider the data presented to be robust. However, studies with a small sample should consider the potential influence of different weekend days in their analysis (if quality assessment specifications such as 4 + 1 are used: four valid weekdays and one [random] weekend day). In addition to absolute minutes, relative proportions of MVPA should be reported in future studies.

The present study is limited due to its cross-sectional nature and we do not intend to infer causality. The main goal of MoMo is to track and report PA and fitness of children and adolescents in a nationwide sample, and significant effort was put into collecting representative data from 167 sample points all over the country.

In the light of the recent development to assess 24 h PB, the present study is subject to the restriction that the accelerometer measures PA during a specific week and during waking hours. Keeping an additional comprehensive and elaborate diary was avoided during the week by wearing an accelerometer. Study participants carried the accelerometer after completion of the already very time-consuming fitness test and surveys on activity and health. A WT of 24 h could solve this problem in future studies. Although we attempted to solve most of the methodological problems in our accelerometer-measured PA, difficulties remained, such as calibration for a wide age range and the occurrence of non-wear times during sports competitions.

This study's advantages are the large and nationwide sample (N= 2278) and the recording of PA of each participant by a PA questionnaire as well as by accelerometry. We collected the sample at 167 locations across Germany throughout the year to account for seasonal effects. To our knowledge, this study was one of the first to look at individual wake-up phases of each day in a week and combine it with an analysis of absolute and relative intensity data of PA measured by accelerometers. Apart from PA, many other parameters were collected in MoMo. This results in a multitude of further evaluation options. Closer examination of the data in MoMo (e.g. fitness levels, sports disciplines, socioeconomic status, migration status) will give a more detailed answer as to the reasons behind the differences regarding age and gender.

6.6 Conclusion

This study was the first to provide detailed information on device-based PA behavior on a national level in Germany. The results regarding PA patterns on weekdays and weekends allow the conclusion to be drawn that the WT of the participants increases with age because the waking phase during the day increases. However, despite the longer waking periods on Friday and Saturday (in terms of absolute minutes), the percentage distribution of time spent in the different intensities of activity is similar on all weekdays and weekends. Since shorter waking periods limit the possible absolute active time on weekends, PA and SED times are lower on the weekend compared to weekdays. The visualization of PA on a typical day for an average MoMo participant may help researchers compare their data. We could not find any justification for specific weekend interventions from PA-level distributions alone. Instead, we believe that interventions should generally try to change activity behavior from SED toward a more active lifestyle. We recommend that low-threshold sports, play, and exercise facilities for children and adolescents should be increased and promoted on the municipal level. Here, open public spaces and exercise areas (e.g., paths, squares, courtyards, green spaces, forests, playgrounds, football fields and skate parks) should be created to support a generally more active lifestyle. Joint efforts to ensure physical activity and sports programs in schools, clubs, and leisure time are necessary at the federal, state, and local levels and should be extensively funded within a federal pact to promote PA (Woll et al., 2021). In the present study, we provided daily patterns for intensity distributions of PA levels. We suggest that PA analysis from wake-up to wake-up may not be a fixed 24 hour PB cycle. Due to sleeping patterns, for example, there is more time to be active on Friday compared to Saturday, which is why the comparability of these two days is limited. In the future, this should be considered when interventions are made and evaluated. Future studies of PB should consider using the sleep–wake cycle rather than the 24-hour basis.

Authors' contributions

AB was responsible for the general conception and design of this manuscript. CN & DO supported the process of writing the manuscript. SS supported in the statistical analysis. AB, CN, DO, SS, SK were responsible for the acquisition of the data. DO, SK, SS, BHM, CN, AW were responsible for the critical revision. CN & AW contributed to the design of the study. All authors read and approved the final manuscript.

Table 6.3: Overall mean minutes (%) of wear time (WT) in different intensities over weekdays for three age groups.

Age (years)	N	Intensities	Monday min (%)		Thuesday min (%)		Wednesday min (%)		Thursday min (%)		Friday min (%)		Saturday min (%)		Sunday min (%)		Week min (%)	
6-10	713	SED	425.2	(55%)	423.8	(56%)	424.4	(55%)	424.5	(55%)	442.5	(55%)	396.6	(55%)	371.2	(57%)	2908.1	(55%)
		LPA	276.4	(36%)	273.6	(36%)	280.5	(36%)	276.0	(36%)	289.0	(36%)	263.3	(36%)	234.6	(36%)	1893.4	(36%)
		MPA	45.8	(6%)	45.4	(6%)	47.3	(6%)	47.2	(6%)	47.6	(6%)	42.4	(6%)	36.0	(5%)	311.8	(6%)
		VPA	20.8	(3%)	20.5	(3%)	21.2	(3%)	21.0	(3%)	21.3	(3%)	19.3	(3%)	15.2	(2%)	139.3	(3%)
		WT	768.2	(100%)	763.3	(100%)	773.4	(100%)	768.7	(100%)	800.4	(100%)	721.6	(100%)	656.9	(100%)	5252.5	(100%)
		MVPA	66.6	(9%)	65.9	(9%)	68.5	(9%)	68.2	(9%)	68.9	(9%)	61.7	(9%)	51.2	(8%)	451.0	(9%)
11-13	706	SED	550.7	(69%)	562.9	(70%)	565.8	(70%)	564.5	(70%)	594.8	(70%)	512.6	(69%)	470.3	(71%)	3821.7	(70%)
		LPA	194.3	(24%)	192.9	(24%)	194.1	(24%)	191.7	(24%)	203.4	(24%)	184.6	(25%)	161.7	(24%)	1322.7	(24%)
		MPA	25.2	(3%)	25.9	(3%)	26.0	(3%)	25.7	(3%)	26.4	(3%)	22.1	(3%)	17.9	(3%)	169.2	(3%)
		VPA	23.8	(3%)	24.2	(3%)	24.7	(3%)	24.9	(3%)	25.2	(3%)	18.9	(3%)	14.9	(2%)	156.7	(3%)
		WT	794.0	(100%)	805.9	(100%)	810.6	(100%)	806.9	(100%)	849.8	(100%)	738.2	(100%)	664.9	(100%)	5470.3	(100%)
		MVPA	49.0	(6%)	50.1	(6%)	50.7	(6%)	50.6	(6%)	51.6	(6%)	41.1	(6%)	32.8	(5%)	325.9	(6%)
14-17	859	SED	651.5	(78%)	653.2	(78%)	648.2	(77%)	653.0	(78%)	674.6	(77%)	558.9	(76%)	535.2	(79%)	4374.6	(78%)
		LPA	139.6	(17%)	141.1	(17%)	144.2	(17%)	142.2	(17%)	155.8	(18%)	138.2	(19%)	115.0	(17%)	976.1	(17%)
		MPA	14.2	(2%)	14.8	(2%)	15.3	(2%)	14.7	(2%)	15.9	(2%)	12.9	(2%)	9.7	(1%)	97.5	(2%)
		VPA	26.9	(3%)	28.6	(3%)	29.3	(4%)	28.7	(3%)	29.7	(3%)	20.8	(3%)	14.5	(2%)	178.5	(3%)
		WT	832.2	(100%)	837.8	(100%)	837.0	(100%)	838.5	(100%)	875.9	(100%)	730.8	(100%)	674.4	(100%)	5626.6	(100%)
		MVPA	41.1	(5%)	43.5	(5%)	44.6	(5%)	43.3	(5%)	45.5	(5%)	33.8	(5%)	24.2	(4%)	276.0	(5%)

Note: **SED** Sedentary behavior, **LPA** light physical activity, **MPA** moderate physical activity, **VPA** vigorous physical activity, **min (%)** mean minutes (%) of wear time in different intensities for each weekday and the whole week.

Table 6.4: Minutes of wear time on weekdays and weekend days

Age (years)		N	Monday - Thursday (M ± s)	Friday (M ± s)	Difference diff.	95% CI	t	p	d
6-10	m	357	771.6 ± 84.1	807.9 ± 186.7	-55.16	-17.46	-3.79	<.01	-0.20
	f	356	765.2 ± 81.9	792.9 ± 170.8	-45.28	-10.18	-3.11	<.01	-0.17
	Ø	713	768.4 ± 83.0	800.4 ± 179.0	-44.88	-19.18	-4.89	<.01	-0.18
11-13	m	327	798.7 ± 91.0	843.1 ± 202.6	-66.01	-22.82	-4.05	<.01	-0.22
	f	379	810.1 ± 87.2	856.6 ± 184.7	-64.65	-28.37	-5.04	<.01	-0.26
	Ø	706	804.8 ± 89.1	850.3 ± 193.2	-59.46	-31.61	-6.42	<.01	-0.24
14-17	m	388	830.3 ± 120.3	874.9 ± 189.9	-64.29	-24.96	-4.46	<.01	-0.23
	f	471	842.9 ± 99.5	875.3 ± 188.2	-49.66	-15.13	-3.69	<.01	-0.17
	Ø	859	837.2 ± 109.5	875.1 ± 188.8	-50.88	-24.96	-5.74	<.01	-0.20
All	m	1,072	801.1 ± 103.5	842.9 ± 194.7	-53.31	-30.27	-7.12	<.01	-0.22
	f	1,206	809.6 ± 96.1	845.1 ± 185.2	-45.67	-25.24	-6.81	<.01	-0.20
	Ø	2,278	805.6 ± 99.7	844.1 ± 189.7	-46.09	-30.78	-9.85	<.01	-0.21

Age (years)		N	Saturday (M ± s)	Sunday (M ± s)	Difference diff.	95% CI	t	p	d
6-10	m	357	728.2 ± 194.1	659.7 ± 166.9	43.89	93.09	5.48	<.01	0.29
	f	356	715.0 ± 182.8	654.2 ± 148.4	38.43	83.18	5.35	<.01	0.28
	Ø	713	721.6 ± 188.5	657.0 ± 157.8	48.06	81.24	7.65	<.01	0.29
11-13	m	327	740.0 ± 230.4	655.6 ± 191.5	54.45	114.32	5.55	<.01	0.31
	f	379	736.5 ± 226.0	674.2 ± 144.3	37.04	87.60	4.85	<.01	0.25
	Ø	706	738.1 ± 227.9	665.5 ± 168.0	53.17	91.91	7.35	<.01	0.28
14-17	m	388	728.0 ± 223.0	669.3 ± 190.3	32.61	84.77	4.42	<.01	0.23
	f	471	733.9 ± 203.4	678.7 ± 194.5	32.75	77.73	4.83	<.01	0.22
	Ø	859	731.2 ± 212.4	674.4 ± 192.6	39.78	73.82	6.55	<.01	0.22
All	m	1,072	731.7 ± 216.0	661.9 ± 183.1	54.35	85.23	8.87	<.01	0.27
	f	1,206	729.1 ± 205.2	670.0 ± 167.0	45.59	72.63	8.58	<.01	0.25
	Ø	2,278	730.3 ± 210.3	666.2 ± 174.8	53.94	74.33	12.34	<.01	0.26

Note: **Ø** mean of males and females, **M** mean, **s** standard deviation, **diff.** difference, **95% CI** 95% confidence interval, **t** posthoc t-test; **p** Significant difference, **d** effect sizes calculated using Cohen's d.

References Typical Day Article

Aibar, A., Bois, J. E., Zaragoza Casterad, J., Generelo, E., Paillard, T., & Fairclough, S. J. (2014). Weekday and weekend physical activity patterns of French and Spanish adolescents. *European journal of sport science*, *14*(5), 500–509. https://doi.org/10.1080/17461391.2013.829127

Ajja, R., Wikkeling-Scott, L. F., Brazendale, K., Hijazi, R., & Abdulle, A. (2021). Accelerometer measured physical activity patterns of children during segmented school day in Abu Dhabi. *BMC pediatrics*, *21*(1), 182. https://doi.org/10.1186/s12887-021-02639-7

Bailey, D. P., Fairclough, S. J., Savory, L. A., Denton, S. J., Pang, D., Deane, C. S., & Kerr, C. J. (2012). Accelerometry-assessed sedentary behaviour and physical activity levels during the segmented school day in 10-14-year-old children: the HAPPY study. *European journal of pediatrics*, *171*(12), 1805–1813. https://doi.org/10.1007/s00431-012-1827-0

Berglind, D., & Tynelius, P. (2017). Objectively measured physical activity patterns, sedentary time and parent-reported screen-time across the day in four-year-old Swedish children. *BMC public health*, *18*(1), 69. https://doi.org/10.1186/s12889-017-4600-5

Brazendale, K., Beets, M. W., Armstrong, B., Weaver, R. G., Hunt, E. T., Pate, R. R., Brusseau, T. A., Bohnert, A. M., Olds, T., Tassitano, R. M., Tenorio, M. C. M., Garcia, J., Andersen, L. B., Davey, R., Hallal, P. C., Jago, R., Kolle, E., Kriemler, S., Kristensen, P. L., ... van Sluijs, E. M. F. (2021). Children's moderate-to-vigorous physical activity on weekdays versus weekend days: a multi-country analysis. *International Journal of Behavioral Nutrition and Physical Activity*, *18*(1), 28. https://doi.org/10.1186/s12966-021-01095-x

Burchartz, A., Manz, K., Anedda, B., Niessner, C., Oriwol, D., Schmidt, S. C. E., & Woll, A. (2020). Measurement of Physical Activity and Sedentary Behavior by Accelerometry Among a Nationwide Sample from the KiGGS and MoMo Study: Study Protocol. *JMIR Research Protocols*, *9*(7), e14370. https://doi.org/10.2196/14370

Burchartz, A., Oriwol, D., Kolb, S., Schmidt, S. C. E., Wunsch, K., Manz, K., Niessner, C., & Woll, A. (2021). Comparison of self-reported & device-based, measured physical activity among children in Germany. *BMC public health*, *21*(1), 1081. https://doi.org/10.1186/s12889-021-11114-y

Bussmann, J. B. J., & van den Berg-Emons, R. J. G. (2013). To total amount of activity..... and beyond: perspectives on measuring physical behavior. *Frontiers in Psychology*, *4*, 463. https://doi.org/10.3389/fpsyg.2013.00463

Chaput, J.-P., Dutil, C., & Sampasa-Kanyinga, H. (2018). Sleeping hours: what is the ideal number and how does age impact this? *Nature and science of sleep*, *10*, 421–430. https://doi.org/10.2147/NSS.S163071

Chaput, J.-P., Willumsen, J. F., Bull, F. C., Chou, R., Ekelund, U., Firth, J., Jago, R., Ortega, F. B., & Katzmarzyk, P. T. (2020). 2020 WHO guidelines on physical activity and sedentary behaviour for children and adolescents aged 5-17 years: summary of the evidence. *The international journal of behavioral nutrition and physical activity*, *17*(1), 141. https://doi.org/10.1186/s12966-020-01037-z

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. ed.). Erlbaum. http://www.loc.gov/catdir/enhancements/fy0731/88012110-d.html

Corder, K., Craggs, C., Jones, A. P., Ekelund, U., Griffin, S. J., & van Sluijs, E. M. F. (2013). Predictors of change differ for moderate and vigorous intensity physical activity and for weekdays and weekends: a longitudinal analysis. *International Journal of Behavioral Nutrition and Physical Activity*, *10*(1), 69. https://doi.org/10.1186/1479-5868-10-69

Crane, J. R., Naylor, P.-J., & Temple, V. A. (2018). The Physical Activity and Sedentary Behaviour Patterns of Children in Kindergarten and Grade 2. *Children (Basel, Switzerland)*, *5*(10). https://doi.org/10.3390/children5100131

Dobbins, M., Husson, H., DeCorby, K., & LaRocca, R. L. (2013). School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. *The Cochrane database of systematic reviews*, (2), CD007651. https://doi.org/10.1002/14651858.CD007651.pub2

Drenowatz, C., Gribben, N., Wirth, M. D., Hand, G. A., Shook, R. P., Burgess, S., & Blair, S. N. (2016). The Association of Physical Activity during Weekdays and Weekend with Body Composition in Young Adults. *Journal of obesity*, *2016*, 8236439. https://doi.org/10.1155/2016/8236439

Fairclough, S. J., Beighle, A., Erwin, H., & Ridgers, N. D. (2012). School day segmented physical activity patterns of high and low active children. *BMC public health*, *12*(1), 406. https://doi.org/10.1186/1471-2458-12-406

Fairclough, S. J., Boddy, L. M., Mackintosh, K. A., Valencia-Peris, A., & Ramirez-Rico, E. (2015). Weekday and weekend sedentary time and physical activity in differentially active children. *Journal of Science and Medicine in Sport*, *18*(4), 444–449. https://doi.org/10.1016/j.jsams.2014.06.005

Frago-Calvo, J. M., Aibar, A., Ibor, E., Generelo, E., & Zaragoza Casterad, J. (2018). Weekday and Weekend Days correlates of Sedentary Time and Screen-Based Behaviors in Children. *European Journal of Human Movement*, (40), 47–65. http://www.eurjhm.com/index.php/eurjhm/article/view/438/624

Gradisar, M., Gardner, G., & Dohnt, H. (2011). Recent worldwide sleep patterns and problems during adolescence: a review and meta-analysis of age, region, and sleep. *Sleep medicine*, *12*(2), 110–118. https://doi.org/10.1016/j.sleep.2010.11.008

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2018). Worldwide trends in insufficient physical activity from 2001 to 2016: A pooled analysis of 358 population-based surveys with 1·9 million participants. *The Lancet Global Health*, *6*(10), e1077–e1086. https://doi.org/10.1016/S2214-109X(18)30357-7

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1·6 million participants. *The Lancet Child & Adolescent Health*, *4*(1), 23–35. https://doi.org/10.1016/S2352-4642(19)30323-2

Hoffmann, R., Lange, M., Butschalowsky, H., Houben, R., Allen, J., Kuhnert, R., Schaffrath Rosario, A., & Gößwald, A. (2018). Querschnitterhebung von KiGGS Welle 2: Teilnehmendengewinnung, Response und Repräsentativität (Journal of Health Monitoring, Ed.). https://doi.org/10.17886/RKI-GBE-2018-015

Kolle, E., Solberg, R. B., Säfvenbom, R., Dyrstad, S. M., Berntsen, S., Resaland, G. K., Ekelund, U., Anderssen, S. A., Steene-Johannessen, J., & Grydeland, M. (2020). The effect of a school-based intervention on physical activity, cardiorespiratory fitness and muscle strength: the School in Motion cluster randomized trial. *International Journal of Behavioral Nutrition and Physical Activity*, *17*(1), 154. https://doi.org/10.1186/s12966-020-01060-0

Love, R., Adams, J., & van Sluijs, E. M. F. (2019). Are school-based physical activity interventions effective and equitable? A meta-analysis of cluster randomized controlled trials with accelerometer-assessed activity. *Obesity Reviews*, *20*(6), 859–870. https://doi.org/10.1111/obr.12823

Mannocci, A., D'Egidio, V., Backhaus, I., Federici, A., Sinopoli, A., Ramirez Varela, A., Villari, P., & La Torre, G. (2020). Are There Effective Interventions to Increase Physical Activity in Children and Young People? An Umbrella Review. *International journal of environmental research and public health*, *17*(10). https://doi.org/10.3390/ijerph17103528

Mauz, E., Lange, M., Houben, R., Hoffmann, R., Allen, J., Gößwald, A., Hölling, H., Lampert, T., Lange, C., Poethko-Müller, C., Richter, A., Rosario, A. S., von Schenck, U., Ziese, T., & Kurth, B.-M. (2019). Cohort profile: KiGGS cohort longitudinal study on the health of children, adolescents and young adults in Germany. *International journal of epidemiology*, *49*(2), 375–375k. https://doi.org/10.1093/ije/dyz231

Mishra, A., Pandey, R. K., Minz, A., & Arora, V. (2017). Sleeping Habits among School Children and their Effects on Sleep Pattern. *Journal of caring sciences*, *6*(4), 315–323. https://doi.org/10.15171/jcs.2017.030

Naylor, P.-J., Nettlefold, L., Race, D., Hoy, C., Ashe, M. C., Wharf Higgins, J., & McKay, H. A. (2015). Implementation of school based physical activity interventions: a systematic review. *Preventive medicine*, *72*, 95–115. https://doi.org/10.1016/j.ypmed.2014.12.034

Noland, H., Price, J. H., Dake, J., & Telljohann, S. K. (2009). Adolescents' sleep behaviors and perceptions of sleep. *The Journal of school health*, *79*(5), 224–230. https://doi.org/10.1111/j.1746-1561.2009.00402.x

Padmapriya, N., Chen, B., Goh, C. M. J. L., Shek, L. P. C., Chong, Y. S., Tan, K. H., Chan, S.-Y., Yap, F., Godfrey, K. M., Lee, Y. S., Eriksson, J. G., Bernard, J. Y., & Müller-Riemenschneider, F. (2021). 24-hour movement behaviour profiles and their transition in children aged 5.5 and 8 years - findings from a prospective cohort study. *International Journal of Behavioral Nutrition and Physical Activity*, *18*(1), 145. https://doi.org/10.1186/s12966-021-01210-y

Skender, S., Ose, J., Chang-Claude, J., Paskow, M., Brühmann, B., Siegel, E. M., Steindorf, K., & Ulrich, C. M. (2016). Accelerometry and physical activity questionnaires - a systematic review. *BMC public health*, *16*, 515. https://doi.org/10.1186/s12889-016-3172-0

Steele, R. M., van Sluijs, E. M. F., Sharp, S. J., Landsbaugh, J. R., Ekelund, U., & Griffin, S. J. (2010). An investigation of patterns of children's sedentary and vigorous physical activity throughout the week. *The international journal of behavioral nutrition and physical activity*, *7*, 88. https://doi.org/10.1186/1479-5868-7-88

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine and science in sports and exercise*, *40*(1), 181–188. https://doi.org/10.1249/mss.0b013e31815a51b3

Trost, S. G., McIver, K. L., & Pate, R. R. (2005). Conducting accelerometer-based activity assessments in field-based research. *Medicine and science in sports and exercise*, *37*(11 Suppl), S531–43. https://doi.org/10.1249/01.mss.0000185657.86065.98

van Sluijs, E. M. F., McMinn, A. M., & Griffin, S. J. (2007). Effectiveness of interventions to promote physical activity in children and adolescents: systematic review of controlled trials. *BMJ (Clinical research ed.)*, *335*(7622), 703. https://doi.org/10.1136/bmj.39320.843947.BE

Vandenbroucke, J. P., von Elm, E., Altman, D. G., Gøtzsche, P. C., Mulrow, C. D., Pocock, S. J., Poole, C., Schlesselman, J. J., & Egger, M. (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Epidemiology (Cambridge, Mass.)*, *18*(6), 805–835. https://doi.org/10.1097/ede.0b013e3181577511

Vanhelst, J., Fardy, P. S., Duhamel, A., & Béghin, L. (2014). How many days of accelerometer monitoring predict weekly physical activity behaviour in obese youth? *Clinical Physiology and Functional Imaging*, *34*(5), 384–388. https://doi.org/10.1111/cpf.12109

Wang, L. (2019). Accelerometer-determined physical activity of children during segmented school days: The Shanghai perspective. *European Physical Education Review*, *25*(3), 816–829. https://doi.org/10.1177/1356336X18776874

Williams, J. A., Zimmerman, F. J., & Bell, J. F. (2013). Norms and trends of sleep time among US children and adolescents. *JAMA pediatrics*, *167*(1), 55–60. https://doi.org/10.1001/jamapediatrics.2013.423

Wolff-Hughes, D. L., McClain, J. J., Dodd, K. W., Berrigan, D., & Troiano, R. P. (2016). Number of accelerometer monitoring days needed for stable group-level estimates of activity. *Physiological Measurement*, *37*(9), 1447–1455. https://doi.org/10.1088/0967-3334/37/9/1447

Woll, A., Albrecht, C., & Worth, A. (2017). Motorik-Module (MoMo) – the KiGGS wave 2 module to survey motor performance and physical activity. *Journal of Health Monitoring*, *2*(S3), 66–73. https://doi.org/10.17886/RKI-GBE-2017-110

Woll, A., Scharenberg, S., Klos, L., Opper, E., & Niessner, C. (2021). Fünf Thesen und elf Empfehlungen zur Bewegungs- und Sportförderung für Kinder und Jugendliche vor dem Hintergrund der Corona-Pandemie „Es ist bereits fünf nach Zwölf – wir fordern einen Bewegungspakt!“ https://doi.org/10.5445/IR/1000139901/v2

World Health Organization. (2010). *Global recommendations on physical activity for health*.

World Health Organization. (2019). *Global Action Plan on Physical Activity 2018-2030: More Active People for a Healthier World*. https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=5910089

7. Discussion and Conclusion

The aim of this thesis was to investigate the usage of accelerometers to measure PA in the MoMo-Study. The main objectives were to clarify technical decisions that should be made when using accelerometers in large epidemiological studies of activity behavior research, investigate the implications of different evaluation methods in the preprocessing of accelerometer data in general, and provide reference data to practitioners and researchers when comparing their accelerometer-measured physical activity with the MoMo study. The analyzes performed were able to contribute to the achievement of these objectives. Based on the results already reported, a summary of the answers is given below. In five research articles, it was shown that

(1) the "Consensus Article" helps the end user of PA monitoring devices to work through the excessive technical details of accelerometry, use **outlines of best practices in device selection and application**, and apply these learned lessons to quantify the three broad behavioral categories of PA, SED, and sleep (compare Chapter 2);

(2) the study protocol provided an extensive list of considerations to measure physical activity and sedentary behavior by accelerometry in a large epidemiological sample of the MoMo-Study; Researchers planning studies with the same protocol are given a **best practice example and all the information needed for replication** and allows comparability to other epidemiological studies (compare Chapter 3);

(3) there are **challenges when evaluating data analyzed with different epoch lengths** that need to be considered. Specifically, when different non-wear time algorithms and definitions of activity cut-points are used with epoch lengths that were not originally used to validate those algorithms. As a consequence, the resulting differences in the estimated SED and PA values can become very large (compare Chapter 4).

(4) Children and adolescents living in Germany show a **very low adherence to the 2010 WHO Guideline** (World Health Organization, 2010). These results were confirmed by accelerometer and PA questionnaire. Surprisingly, the differences in meeting the Guideline between the measurement methods are much smaller for younger children than for older age groups. **Continuous overestimation of the**

self-report in contrast to the accelerometer was observed in all other age groups (compare Chapter 5);

(5) the percentage distribution of time spent in **different activity intensities is the same across all weekdays and weekend days**. However, since shorter awake periods limit the absolute possible active time on weekends, PA and SED times are lower on weekends than during the week. We could not find a justification for specific weekend interventions; instead, we suggest that the physical behavior analysis from wake up to wake up may not be a fixed 24-hour behavioral cycle. (compare Chapter 6).

Accelerometer devices for measuring physical activity have emerged as important self-monitoring tools in clinical medicine (Lobelo et al., 2018), large-scale health promotion (U.S. Department of Health and Human Services, 2018), and are now commonly used in national surveillance efforts (Burchartz, Manz, et al., 2020; Colley et al., 2011; Matthews et al., 2008; Troiano et al., 2008). The number of yearly publications using search terms 'exercise or physical activity' and 'acceleromet*' increased from 10 or fewer until 1996 to more than 600 in 2012 and 2013 (Troiano et al., 2014), to more than 1400 publications on scopus.com in 2021 (compare Figure 1.5 assessed on Nov. 04, 2022). Despite this large number of publications dealing with the assessment of device-based measured physical activity, there is still no general consensus on a standardized assessment, although it is regularly criticized (Bussmann & van den Berg-Emons, 2013; Migueles et al., 2017; Rowlands, 2007; Troiano et al., 2020).

Accelerometry is nevertheless the "state of the art" when it comes to device-based measurement of PB. The advantage of accelerometry is that it can collect dense data over a long period of time, allowing a detailed examination of daily behavior. From these multidimensional data, a large number of metrics can be derived to capture and describe the unique aspects of PB. In addition to PA, SED and sleep are the most common behaviors being assessed (Diaz et al., 2017; Doherty et al., 2017; German National Cohort Consortium, 2014; Matthews et al., 2016). Various carrying positions and sensors are available for different areas of application. Complex and dense data resulting from device-based measured PB, as well as various options regarding devices, data collection, and data analysis, can also be a challenge for researchers, and there are still no internationally accepted standards for signal processing (Leppänen et al., 2022; Migueles, Cadenas-Sanchez, et al., 2019; Rowlands, 2007; Trost, 2007). Furthermore, the different approaches used in the studies can lead to limited comparability and reproducibility of the study results. The aim of the first article, the Consensus Article of Chapter 2, was to outline the best practices and highlight the necessary considerations when selecting and using devices to quantify the two important behavioral categories PA and SED, which are of general interest to the research community. This helps smaller studies to better navigate the excessive technical details of accelerometry.

It is important to note that device-based PA measurements are often assumed to be accurate and reflect actual PB. However, the values of, for example, accelerometers are still estimates, and in the absence of satisfactory agreement with the ground truth gold standard measures of free-living PA should not be interpreted as 'actual' PA levels. The consortium at the CAPA Workshop proposed that researchers choose their method based on the most valid approach for their given behavioral metric. From this perspective, the method chosen will dictate the type of device and the prediction algorithm. The research community has not yet reached a consensus on the most promising approach in statistical analyses of device-based measured data (Migueles et al., 2021), besides that the inherent multicollinearity within data based on human behavior during a finite period of time should be carefully considered. Each approach has its strengths, limitations, and practical relevance, and some examples are discussed in Chapter 7.3. Therefore, researchers must carefully inspect the approach that fits best with their research goal and data. A repository of Clevenger, Montoye, et al. (2022) will also help researchers find novel methods to estimate physical activity or energy expenditure and the best possible solution for their study. It is possible to find models sorted by population, wear location, or device brand and gives a good overview of the currently published open source models.

The results are another step towards a better understanding of the evaluation of the physical activity behavior of children and adolescents in Germany. However, as mentioned above, the field of research is not yet in consensus, neither in the selection and attachment of the device nor in the signal processing and the following statistical analysis. Therefore, the important issue of documentation and differences in the evaluation of accelerometers, especially the MoMo study, will be examined in more detail next.

7.1 Issue of methods

How is acceleration data analyzed in MoMo and what are the causation and specific underlying characteristics?

Only since 2017 have there been national recommendations on physical activity in Germany (Rütten & Pfeifer, 2017). These are also largely based on representative questionnaire data on physical activity of children and adolescents in Germany collected in previous years by the MoMo and KiGGS studies (Woll et al., 2017). Despite that, two years earlier in 2015, the MoMo study was already again the pioneer in Germany for the next development in data collection and went one step further. MoMo is the only national study in Germany that has collected data on physical activity using accelerometers since 2015 (Burchartz, Manz, et al., 2020). Especially in a study that provides representative results that other studies (with special subsamples, such as children with certain diseases) from Germany use to compare their data, it is important that all technical decisions are well documented. Therefore, for MoMo, it was important to fulfill the following three important criteria (Burchartz, Manz, et al., 2020) when introducing this acquisition method, which was still quite new at the time:

1.) Complete documentation of all technical decisions
2.) Possibility of later integration into the International Children's Accelerometry Database (ICAD)
3.) Delivery of reference data at national level

Therefore, the study protocol of Chapter 3 as the second research article in this thesis provides an extensive list of considerations to measure physical activity and sedentary behavior using accelerometry in the large sample of the MoMo-Study (see Table 7.1). It gives all the details of similar studies for replication and allows comparison with other large studies.

Table 7.1: List of accelerometer criteria used in MoMo, modified from Burchartz, Anedda, et al. (2020)

Criteria	Definitions within this study
Accelerometer devices	ActiGraph (models: GT3X+, wGT3X-BT)
Placement of the device	Laterally on top of the right anterior superior iliac spine
Sampling frequency	30 Hz
Filter	Normal ActiGraph GT3X filter
Epoch lengths	1 s with possibility to convert to 5, 10, 15, 30, and 60 s
Non-wear time definition	Choi et al. (2011): 90-minute time window for consecutive zero/nonzero counts; allowance of 2-minute intervals of nonzero counts with an up/downstream 30-minute consecutive zero counts window
Valid days/valid weeks	8 h of recordings on four weekdays and one further weekend day when wearing the device for 7 days
Population age range	Children and adolescents from 6 to 17 years
Sedentary and physical activity intensity classification and cut point algorithms	6-10 years: Evenson et al. (2008) vertical axis 11-17 years: Romanzini et al. (2014) vertical axis

Furthermore, the specifications for integration into the ICAD database could also be included (Sherar et al., 2011), since the physical activity data are stored in raw accelerometer (.dat) files from a waist-worn ActiGraph accelerometer in children aged 6-17 years and accompanying data of gender, age, measured height, and weight were also collected. Although the original plans for the expansion of ICAD included the addition of new studies, it became clear to the ICAD consortium over time that this would not be possible within the scheduled time frame and staff capacity, so this phase has been put on hold (Atkin et al., 2017). However, discussions on adding new partners to ICAD are expected to resume soon. For this reason, it was not possible to actually integrate the MoMo data into the ICAD database until today. The description of the national reference data will be presented later, as they are part of the two scientific articles in Chapters 5 and 6. Meanwhile, let's look at another discussion point that was frequently raised in the context. The era of accelerometers in physical behavior research could be described as a golden age with tremendous opportunities, but, as we have noticed, there are also many challenges. Although suitable and accurate devices are available for most applications (e.g., interventions, epidemiology, surveillance), the wide variety of devices and prediction algorithms available for a variety of metrics (e.g., step count, energy expenditure, intensity classification, posture, sleep behavior) and the limited information available from rigorous validation studies make it difficult for the average user to understand which options are best suited for each application. A device such as, for example, the ActiGraph in MoMo may not be appropriate for all applications, and users must make informed decisions to optimize the results of each study. As mentioned in Chapters 3 and 4, there are also some methodological problems in MoMo that need to be disscussed.

First, the use of **normal ActiGraph filter** removes signals with a frequency greater than 2.5 Hz. However, while performing vigorous physical activity, the human body produces accelerations on the hip up to a frequency of 3.4 Hz (Cavagna & Franzetti,

1986; Cavagna et al., 1991). Due to this limitation, activities with higher movement frequencies (that is, in the vigorous activity spectrum) may not be correctly assessed. In the context of MoMo and KiGGS, this will not be an issue, because all activities in this frequency range will be classified as vigorous and more detailed investigations are currently not planned. This is only one of the proprietary ActiGraph signal preprocessing algorithms, which was not specified in more detail by the manufacturer processing until recently (Neishabouri et al., 2022). Other technical decisions, such as the use of **sampling frequencies**, can have a great impact on the outcome. For example, Clevenger et al. (2019) reported that the sampling rate affects the determination of intensity classification. They recommended that researchers be consistent and vigilant in reporting these decisions, since higher Hz data resulted in significantly more total generated counts when using hip-worn monitors. Similar conclusions were reached by Brønd et al. (2019) who investigated the variation of different bandpass filters and found a reduced bias in the measurement when running above 10 km/h due to the higher frequency.
Furthermore, there are **subjective decisions** (placement of the device, device selection, study protocol, sampling frequency, choice of epoch length, etc.) that were made in MoMo after thorough research and analysis of existing validation studies, but that influence the non-wear time and also the intensity calculation (compare Chapters 3 and 4. In MoMo in particular, there is also the **restriction** that the physical activity questionnaire asked and assessed physical activity of an average week and the week before the accelerometer measured physical activity during a specific week. An additional comprehensive and elaborate diary was avoided during the week by wearing an accelerometer. Study participants carried the accelerometer after completing the already very time-consuming fitness test and activity and health surveys. Having self-reported and measured data will help explain the observed population differences in Chapter 5. As already said, the true value probably lies somewhere in between these two methods. To record a more accurate activity profile, another combination of both methods might be a solution. This could be an algorithm to subtract or add the methodological difference if only one method is used. A solution to not lose activity during non-wearing time could be the use of ambulatory assessment in combination with 24 h recording (Reichert et al., 2020). Triggered e-diaries may ask the subjects for the type of activity performed after certain events have been detected (e.g. device not worn, periods of high activity, or sedentary behavior). Then, non-wear times, their reasons, and activities performed while not wearing the device can be considered uniformly. The participant can receive feedback on how much activity was not recorded and how much was missing to reach the guidelines. Ambulatory assessment can also be used to ensure that the activity time frames for both methods are consistent by answering the physical activity questionnaire on the mobile phone at the end of the survey. In this way, the strengths of one method would compensate for the limitations of the other method.
Although **wearing times** were very long on average, some participants reported that wearing had been prohibited in some sports competitions, such as soccer. The use of electronic devices was forbidden to prevent trainers from having an unfair advantage in knowledge. Even if documented by the non-wear time protocol (Burchartz, Anedda, et al., 2020), the **unrecorded activities** could not be retrospectively taken into account. Manual input of the data from the handwritten non-wear protocol is very time consuming. Also, the information in the protocols is very inconsistent, and manual input would distort the acceleration data. These missing data could be another link to the difference between the questionnaire and the accelerometer results in Chapter 5. A wear time of 24 h and a consistent ambulatory assessment for the non-wear time could solve this

problem in future studies. Although **compliance** in MoMo was quite good, other large national studies have switched to wrist-worn devices (Doherty et al., 2017; Freedson & John, 2013). We will look at these developments in device selection and placement on the body in the next subsection of the Chapter.
Another topic to discuss is the **large age span** in MoMo. The range of 6 to 17 years and older, for long-term participants who are now adults, complicates the calculation of intensity because validation studies typically use only age groups with more compact age ranges. Therefore, to cover the entire sample, different algorithms must be used, but may not be directly comparable among themselves. We have demonstrated the extent of variation in PA levels when using ActiGraphs in MoMo during Chapter 4. Specifically, when different non-wear time algorithms and definitions of activity cut-points are used with epoch lengths that were not originally used to validate those algorithms as reported by others (Banda et al., 2016; Breau et al., 2022; Leppänen et al., 2022; Logan et al., 2016). As a consequence, the resulting differences in the estimated SED and PA values can become very large. The advantage of these two algorithms used in MoMo is that they fit very closely with the two age groups for children and adolescents of MoMo, because they had similar validation protocols, and in addition, both used an epoch length of 15s. A possible future solution could be through new validation studies with short epoch lengths for young children, as proposed by (Banda et al., 2016; Giurgiu et al., 2022). As explained in detail in Chapters 3 and 4, our results justify the choice of algorithms in MoMo, but in the end also show the general technical problems of large epidemiological studies that still need to be overcome in the future.

In conclusion, documentation is still essential for comparability within a large study such as MoMo but also for comparison with other studies. However, documentation is often not easy to handle and can also be very technical and time consuming when dealing with large data sets. (compare Chapters 2, 3 and 4).

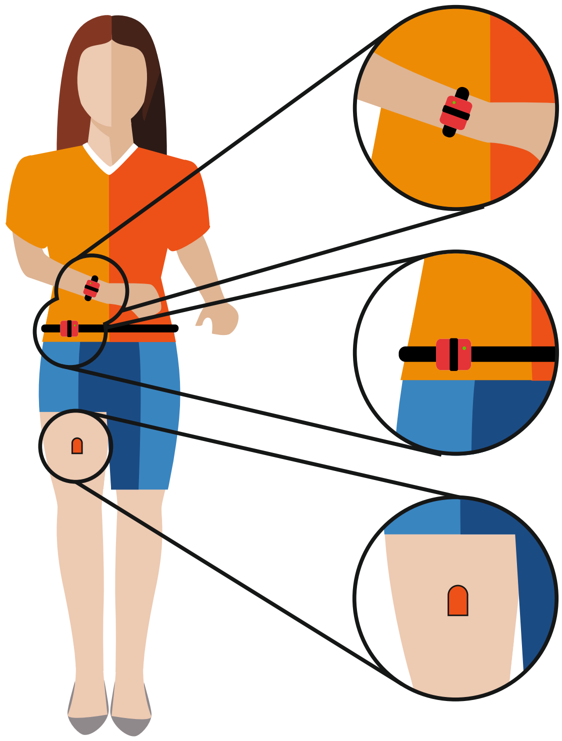
7.2 Issue of Development

How should data be collected in the future?

After analyzing the specific conditions under which the device-based measured physical activity was executed in MoMo, the following section discusses the "lessons learned" and gives an insight into the challenges to be faced in the future when collection device-based physical behavior. As mentioned above, there is still the problem of positioning devices and attaching them to the body. In current research, three positions are used in particular for accelerometers and offer various advantages and disadvantages: the hip, the wrist, and the thigh (Figure 7.1).

As mentioned in Chapter 3, device-based measurement of physical behavior with an accelerometer on the hip has many advantages, but also some limitations. **Placement of the accelerometer** affects which part of the mechanical work of the body is captured. The placement on the hip captures mechanical work and acceleration and deceleration of the body during locomotion relatively accurately; it serves well to measure the intensity of activity consistent with age and stature. However, it should be remembered that mechanical work, such as riding a bicycle or carrying and lifting heavy objects or weights, is not captured very adequately as these activities produce very little movement on the hip. Placements on the thighs and wrists can better capture this physical work performed

Figure 7.1: Device placement on the wrist, the hip and the thigh



Note: Illustrations for ActiGraph on wrist and hip, ActivPAL on thigh.

internally, but do not provide equivalent results to the measurement at the hip (Arvidsson et al., 2019). This is also documented by Tudor-Locke et al. (2015) who found that visual step counts matched better when the accelerometer was attached to the hip than when it was attached to the wrist. Furthermore, Rosenberger et al. (2013) showed higher sensitivity and specificity on the hip compared to attachment to the wrist for motion detection, as other studies did (Chen et al. (2003), Swartz et al. (2000), and Welch et al. (2013)). Lynch et al. (2019), however, showed some evidence in a review suggesting that children are more compliant when placed on the wrist than when wearing devices on the hip. Notably, in the NHANES study, the physical activity monitoring protocol was changed from hip placement to wrist placement. Another large national study using wrist-worn accelerometer devices is the UK Biobank Study with over 100.000 participants (Doherty et al., 2017). This improved adherence to wear time from 40%-70% to 70%-80% (Freedson & John, 2013). The NHANES study assumed at that time that the algorithms could be adjusted to improve the weaker activity estimate based on wrist placement. It is not clear what impact this trade-off between validity and wear time has on the survey data. Furthermore, it should be noted that higher acceleration values are found at the wrist compared to hip placement, resulting in a shorter time in SED and a longer time in light, moderate, and vigorous PA (Routen et al., 2012).
Another limitation is that separate detection of sleep phases was not possible in MoMo, as this would require downloading a second file for each participant in 60 second epochs, with the low frequency extension function selected to process sleep data. Unfortunately, this option was not available. According to (ActiGraph, 2018), this option lowers the amplitude threshold, allowing lower amplitude movements such as when sleeping to be included in the acceleration signal. This would increase the sensitivity to detect the sleep

phase and can be used to remove the sleep phase in accelerometer data (Coyle-Asbil et al., 2020). However, it should be mentioned here that smaller movements during sleep can be better detected by wrist measurements and are often used to assess sleep quality, while hip detection can be used primarily for sleep time (Ancoli-Israel et al., 2003; Full et al., 2018; Smith et al., 2019).

Sedentary Behavior is a major part of the day as has been shown in Chapter 6 but can be a mixture of many different behaviors with low metabolic cost. The Sedentary Behavior Research Network published consensus definitions for different key terms related to these sedentary behaviors (Tremblay et al., 2017). Among other things, a distinction is made between stationary behavior and sedentary behavior, although more differentiated distinctions such as standing, lying, sitting, and reclining are part of these definitions. Consensus definitions clarify the distinction between different behaviors under the broader term "Sedentary Behavior". For example, researchers can now use the term 'stationary time' when reporting data measured by accelerometers that do not measure posture; those who do can differentiate between sitting, standing, reclining, and lying. On the other hand, it is also possible to distinguish different forms of locomotion, such as running, jumping, walking, sneaking, or even climbing, in terms of categories when visual analyzing the participants. Hence, activity classification is the process of identifying what type of activity a person performs based on their body movements. In contrast, intensity recognition refers to the amount of effort a person exerts during an activity. This is very complicated when only the data of the hip-worn accelerometer are available, but works well when using the ActivPAL. This waterproof device, which is worn on the thigh, uses accelerometer-derived information on the position of the thigh in space to determine the beginning and end of each period of changes in sitting / lying, standing, and posture (Edwardson et al. (2017) and compare Figure 7.2).

Figure 7.2: The Different ActivPAL activity classification profiles from www.palt.com

In recent years, **thigh worn accelerometers** such as ActivPALs have become the gold standard for recording especially sedentary behavior, sitting and sleep (Arundell et al., 2019; Blackwood et al., 2022; Carlson et al., 2022; Edwardson et al., 2017; Hurter et al., 2018; Shi et al., 2019; Smits et al., 2022). Since various studies have shown its high validity and reliability in measuring physical activity (Dowd, Harrington, Bourke, et al., 2012; Dowd, Harrington, & Donnelly, 2012; Grant et al., 2006; Kim et al., 2015; Kozey-Keadle et al., 2011), the ActivPAL is increasingly chosen mainly for its accurate discrimination between upright and sedentary postures. This is a huge advantage, especially in the direct output of body posture, in contrast to the calculation of body posture using models from other manufacturers such as ActiGraph or GeneActive, which have not yet provided comparable results (Hurter et al., 2018; Pereira et al., 2020).

When sensors are applied to the thigh with a biopatch, it is also easier to record **24h assessments**, as insufficient wear time leads to data loss and increases the cost of conducting large-scale studies. To date, only a few studies have investigated the impact on compliance and practical benefits of a continuous wear protocol when recording with devices such as the activPAL on the thigh (Edwardson et al., 2017; Hidde et al., 2021; Montoye et al., 2016; Shi et al., 2019), but all found greater compliance than devices worn on the wrist or hip. Shi et al. (2019) as one of the few with children as participants found that sweating and skin irritation were the most common reasons for the removal episodes of children who wore an ActivPAL for 7 days, while the use of alcohol pads and cartoon stickers during application was positively associated with total wear time. This means that in addition to switching to wear sensors on the thigh, a user-friendly wear protocol may need to be introduced to enable continuous wear of the device. In general, we conclude that a protocol with accelerometers worn on the thigh using skin attachment methods appears to be very promising for studying 24-hour physical behavior in children and adolescents. A concluding and more in-depth discussion of the topic of 24-hour physical behavior is provided in the next subsection.

On the other hand, several sensors are now often used at the same time to take advantage of the various benefits of the individual devices (Barboza et al., 2021; Carlson et al., 2022). For example, ActivPALs on the thigh are worn simultaneously with ActiGraph accelerometers on the hip or wrist. Thus, for one, body position and sedentary behavior can be well recorded with the ActivPALs, and for another, these data can be combined and evaluated with the widely used activity evaluations of the ActiGraphs. Manufacturers of other devices such as the movisens Move-III are using additional sensors such as air pressure sensors to better estimate energy consumption when climbing stairs, for example (Gaser et al., 2022; Tost et al., 2019). With this additional sensor, the energy consumption or intensity of the activity can be better estimated, for example, if it is increased when ascending stairs.

Another future development will be that objective and subjective data can merge due to technological advances. The accompanying challenges will be to structure this merging or to connect the data sources. **Open source preprocessing methods** based on raw acceleration to obtain alternative metrics to activity counts using ActiGraph are already existing. Many different software programs, including free ones, are available to process and analyze acceleration data using different types of methods, including Euclidean Norm Minus One (ENMO) and Mean Amplitude Deviation (MAD) (Bakrania et al., 2016) as well as Activity Index and ActiGraph counts. Open source programs include, for example, the R package GGIR (Migueles, Rowlands, et al., 2019; van Hees et al., 2019) and OMGui (Open Lab of Newcastle University), as well as methods by other colleagues (Brønd et al., 2017; Brondeel et al., 2021; Neishabouri et al., 2022). GGIR is an R package that processes multiple-day raw accelerometer data for research on physical behavior and sleep. The term raw refers to data expressed as gravitational acceleration data in m/s^2, whereas previous generation accelerometers stored data in unique units such as counts for all different brands of accelerometers. The system has built-in signal processing including automatic calibration, detection of abnormally high values, non-wear time detection, and calculation of a variety of metrics to describe physical behavior. GGIR is currently one of the most widely used methods for analyzing activity monitor data from different brands, and researchers use this information to describe their recorded accelerometer data. Recently, these newer open source metrics have been used more and more frequently to analyze accelerometer data (Leppänen et al., 2022). For a long time, count-based evaluation was only possible with a proprietary algorithm because different manufacturers used different and unknown proprietary definitions to calculate them. Although methods are now known to calculate strong correlating counts produced by ActiGraphs for accelerometers from other manufacturers (Brønd & Arvidsson, 2016; Brondeel et al., 2021; Clevenger, Montoye, et al., 2022), the trend in evaluation is to use original raw data as a basis for calculating certain metrics. Faced with this

development, ActiGraph has recently published the complete algorithm to calculate the internal counts of the device, so that the raw data from other accelerometers can now officially be converted into ActiGraph counts (Neishabouri et al., 2022). Open source analyzes could improve the comparability between studies by comparing different device models and help further rationalize data analysis (Clevenger, Montoye, et al., 2022). The CHAP child classification method allows researchers to derive from triaxial ActiGraph data worn on the hip, for example, activPAL equivalent measures of sedentary time, transitions from sitting to standing, and sedentary movement patterns (Carlson et al., 2022).
On the other hand, Pfeiffer et al. (2022) recently conducted a review to identify new methods to estimate physical activity or energy expenditure from accelerometer data. They found that more than half of the published methods were never used. This makes sense since less than half of the methods they found were actually made available (e.g., as code or equations in the article). For this reason, Clevenger, Mackintosh, et al. (2022) have created a repository to improve accessibility and further develop these new methods that will help researchers choose the right accelerometer and the appropriate methods for their studies in the future.

In conclusion, the collection of physical activity data with accelerometers in MoMo did face some limitations, and there is a wide field of possibilities to develop. After two complete data collection phases, it can be concluded that it was the right decision at the time when MoMo started the accelerometer measurement to collect data on physical activity over a week using ActiGraphs on the hip. With the data now available, it is possible for studies with more specialized populations to compare their results with the MoMo data. The widespread use of ActiGraphs throughout the world also allows for good comparability internationally and the ability to integrate them into the ICAD database. This comparability is further enhanced by publishing the algorithms used to calculate the ActiGraph counts by the manufacturer (Neishabouri et al., 2022). Initial investigations have shown that this now also allows comparability with raw data from accelerometers from other manufacturers when a similar study protocol was used. In the future, in contrast, the focus should be on a 24-hour assessment, also against the background of advancing technical developments. At the same time, this automatically brings the topic of sleep and sedentary behavior into the spotlight. When selecting a device, the ActivPAL with its attachment to the thigh offers a suitable solution to collect data on physical activity, sedentary behavior, and sleep. Against the background of around-the-clock monitoring, compliance can be increased with the ActivPAL on the one hand, and on the other hand it is possible to examine the various facets of sedentary behavior in greater depth and in a more differentiated manner than has been possible to date with the ActiGraph devices. If the time difference between the survey with the activity questionnaire and the device-based recording of physical behavior is further avoided, it should be possible in the future for MoMo to provide an even more reliable picture of the physical behavior of children and adolescents in Germany. Using open source methods such as the R package GGIR (Migueles, Rowlands, et al., 2019; van Hees et al., 2019), that is witnessing an increasing number of citations from all kinds of studies, will allow future MoMo waves to efficiently generate numerous outcomes of physical behavior from the upcoming large acceleration datasets.

7.3 Issue of Comparability

How can physical behavior data be analyzed in the future?

After the previous chapter discussed the basic methods for collecting and then processing physical activity behavior, the following chapter focuses on the subsequent analysis of the data.
The increasingly rapid pace of technological progress makes it more difficult to compare the methods of recording physical activity because something about the recording is constantly changing. This has already become apparent in the recent past when comparing changes in questionnaires or accelerometry in large epidemiological studies (e.g., carrying position wrist versus hip of accelerometers in "NHANES"). The large number of different indicators of physical activity makes comparisons between studies difficult, if not impossible, despite good documentation (Bai et al., 2016; Burchartz, Anedda, et al., 2020). However, this is not practical in the long term, as documenting different and multifactorial methods would require lengthy explanations and requires a high level of technical expertise from readers (Migueles, Cadenas-Sanchez, et al., 2019). **Datapooling** and uniform reanalysis of raw data could be a solution to overcome inconsistencies in processing criteria. In the thematic fields of MoMo, these approaches have already been successfully applied to accelerometers (http://www.mrc-epid.cam.ac.uk/research/studies/icad/ Atkin et al. (2017) and Steene-Johannessen et al. (2020)) as well as physical fitness data (Eberhardt et al., 2021; Kloe et al., 2019).

Figure 7.3: Logo ICAD

www.mrc-epid.cam.ac.uk/.../icad/

Figure 7.4: Logo MO|RE data

www.motor-research-data.org

These data sets can be preprocessed uniformly, even if they originate from different studies, and then evaluated in a comprehensive and uniform manner. A recent study by Rowlands (2018) has shown that the key results of physical activity derived from data from different devices that were processed identically afterwards were also highly equivalent. This would avoid inconsistencies in the evaluation, and even smaller studies or surveys with different devices could contribute to large data sets and evaluations. Although the ICAD database currently only includes ActiGraph accelerometers worn on the hip, there are promising developments for the future that will make it possible, for example, to include data from various sensors in the same database. Crowley et al. (2019) and colleagues for example, have studied three different brands of thigh-worn accelerometers and found negligible differences, which supports the harmonization of data between cohorts when analyzed in the same way.

With the ability to collect and store even larger amounts of data, it will not be possible to examine large amounts of data in the future without artificial intelligence. However, the promising field of ecological momentary assessment will then also allow us, thanks to artificial intelligence, to directly consider contextual influences (environment, situation, social interaction, etc.) when investigating movement and activity. Then, the congruence of objective and subjective perception before, during, and after movement can be captured directly and individually.

In the future, however, data pooling may also help address another relatively new problem. As explained in Chapter 1, physical activity recommendations have only been widely applied in the last decade. However, the different guidelines make it difficult to make comparisons between recommendations, especially since the WHO, for example, has set itself the goal of updating and adapting the guidelines every 10 years. Therefore, the progression of time alone results in different values of compliance with the recommendations, as the recommendations change on a regular basis. The pooling of data can support this since even large data sets can be evaluated relatively quickly according to more recent or even past recommendations, and thus a comparison with earlier evaluations is still possible.

This apparent paradox raises the question of the **suitability of the Activity Guideline** to quantify activity levels. Is a daily activity of one hour of moderate intensity valued more highly than an activity that is performed less frequently, but for a longer time and at a higher intensity? What constitutes the health effect? Does it depend on the frequency or the total amount of activity? Until 2022 research has not yet provided a clear answer to these questions. The new WHO Guideline of 2020 (Chaput et al., 2020) is now recommending to examine the average time spent with physical activity each week. Therefore, days with activity times longer than 60 min would compensate for those with less activity (Colley et al., 2017). However, the daily stimulus is very important in children (Dwyer et al., 1983). The results of Chapter 5.4 determined whether the subject was active for at least 60 minutes or not on each day individually. Taking a look at the exact times spent with MVPA on each day will result in fewer days of MVPA of at least 60 minutes when both evaluation methods are compared (Colley et al., 2017). However, the main reason since the study was already planned and started in 2014 was that the questionnaire questions still referred to the 2010 WHO Guidelines (World Health Organization, 2010). Only now have the youth activity recommendations changed from a recommendation of at least 60 minutes per day to a recommendation of an average of 60 min per day (Chaput et al., 2020). This adaptation will require changes to the survey questions and sampling methods for future monitoring. However, changing the formulation of the question is unlikely to address the need for physical activity monitoring among children who, especially at young ages, cannot answer a complex question about average behavior during the last days, weeks, or months (Corder et al., 2009; Ekelund et al., 2011). In the future, this may require the use of proxy reports from multiple respondents, including parents and teachers, although both may also miss observing large portions of the day (Troiano et al., 2020).

The alternative of asking for the duration of the day for an entire week may be more accurate, but increases the response time to the survey. Therefore, the measurement of daily PA remains a strength of portable devices for now, and adapting questionnaires to the new WHO guideline remains a real challenge. However, we also reexamined our accelerometer data in accordance with the new WHO guidelines. It should be noted that these results cannot be compared with the results of the question in the questionnaire used in this study and are presented in Chapter 5.4. The questionnaire asked for the

number of days that the subjects have MVPA for more than 60 min, so it is not possible to calculate the average value over several days from this question.
Reanalyzing and comparing the number of subjects who meet the old versus the new guideline based on accelerometer data, we see that the percentage increases from 3% to 34% of study participants. In Figure 7.5 the percentage of participants who meet the two WHO recommendations is presented as a heat map. The heat map shows how many days the subjects have reached the respective recommendation. On the Y-axis, subjects are ranked according to the 2010 recommendation of "daily 60 min MVPA" and on the X-axis, the same subjects are ranked according to the current recommendation of "average 60 min MVPA per day". Both times, accelerometer data are used as the basis for the calculation. If we now look at the 7-day row respective column, this means that 31% of the participants who do not reach MVPA at 60 min on all days still have days of the week on which they do so much physical activity that these outweigh the remaining days according to the new guideline compared to the old one. Together with the 3% that had already reached the 2010 recommendations (cp. Figure 5.2a in Chapter 5.4), the total is 34%. Another example can be illustrated by the following. 9% of those who achieved 60 min MVPA on four days according to the 2010 guideline (compare 4 days ACC WHO 2010 in row 5) achieved a total of more than 420 min MVPA (compare 7 days ACC WHO 2020 in column 8), that is, an average of at least 60 min MVPA on each day for seven days. This drastically reduces the proportion of children and adolescents who are not active enough, which is also likely to cause some political controversy. Until these questions are answered, the overall usefulness of the Activity Guideline must be questioned.

Figure 7.5: WHO recommendation 2010 vs. 2020

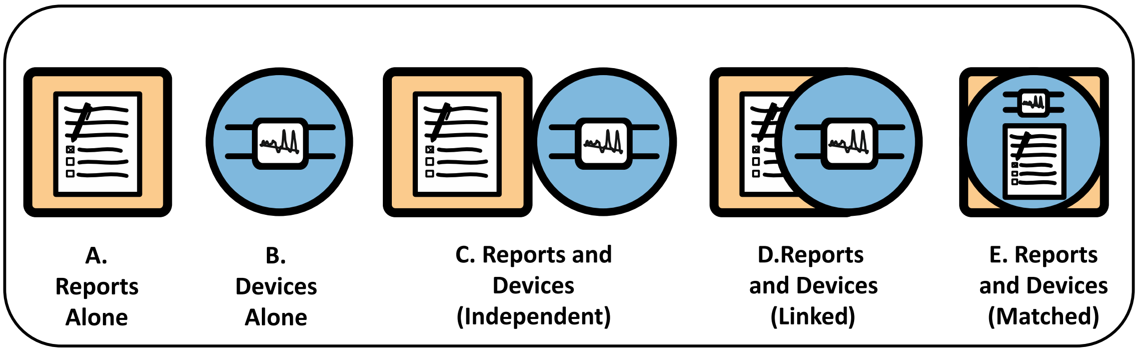
		ACC WHO 2020								**total**
	days	0	1	2	3	4	5	6	7	N
ACC WHO 2010	0	0%	2%	6%	10%	7%	2%	0%	0%	25%
	1	0%	0%	0%	3%	8%	7%	2%	0%	20%
	2	0%	0%	0%	0%	2%	6%	6%	2%	16%
	3	0%	0%	0%	0%	0%	1%	4%	7%	13%
	4	0%	0%	0%	0%	0%	0%	1%	9%	11%
	5	0%	0%	0%	0%	0%	0%	0%	8%	8%
	6	0%	0%	0%	0%	0%	0%	0%	5%	5%
	7	0%	0%	0%	0%	0%	0%	0%	3%	**3%**
total N	%	0%	2%	6%	13%	16%	15%	13%	**34%**	100%

Note: Percentage of participants reaching the WHO Guideline on the one hand, according to the 2010 recommendation of "daily 60 min MVPA" on the Y-axis, and on the other hand on the X-axis, according to the current recommendation of "an average of 60 min MVPA per day". Heat map is color-coded (white/min < gray/moderate < black/max)

Specific perspectives on the development of the assessment of physical activity are offered in the three consensus articles of the International Workshop of the Center for the Assessment of Physical Activity (CAPA) (Nigg & Woll, 2020). In addition to the article on accelerometry (Chapter 2), Reichert et al. (2020) provides detailed information on ambulatory assessment and Nigg et al. (2020) discusses the best practices of physical activity questionnaires to provide some standardization for the field. The fact that there is no general approach to measuring physical activity behavior and related determinants is frustrating for many researchers studying physical behavior. Unfortunately, there is no simple decision tree that leads to a single evaluation method. However, the three articles give some useful rules that should be followed in general.

One of these suggestions is to measure physical activity behavior in parallel with questionnaires but also using accelerometers. The purpose of this idea is to make use of the advantages of both methods in the subsequent analysis of the data. Assessment of physical activity using a self-report is a valuable approach to understanding the health effects of behavioral choices, which cannot be adequately represented by the classification of intensity of physical activity (Nigg et al., 2020). Second, the outcome evaluation of interventions targeting specific types of physical activity in specific settings (e.g. different locomotion types in the context of active commuting to school) is easier with questionnaires. However, one of the main reasons is to identify factors at the individual, socioeconomic, and environmental levels that can influence the behavior of physical activity and possibly be addressed through intervention. The use of questionnaires is often argued to be a cost-effective method of monitoring and observing physical activity and inactivity in a large population. This is certainly still true; however, this argument is gradually being mitigated by the possibility of obtaining accelerometers, which are now relatively inexpensive.
Troiano et al. (2012) and Fig 7.6 describe several different combinations of approaches to how physical activity behavior can be represented using both methods (questionnaire and device-based).

Figure 7.6: Approaches to assessing physical activity and sedentary behavior by report and devices. (adapted from Troiano et al. (2012)).



In 2012, Scenarios A or B were still used. Therefore, there was only interest in physical activity in different settings or only device-based measured physical activity was recorded. In the last 10 years, the field has tended to move towards C. In MoMo, both questionnaires and accelerometers were also used to obtain independent measurement results (see Chapter 5.3). The current challenge is now to link both approaches in a way that allows for an even more comprehensive survey. Partially, this has already been done in MoMo with the logbooks, but there the problems described above were encountered (Section 7.1). For the future, the connection of the two methods should be linked in parts or even completely as in the graphics D and E accordingly. Only then can the limitations of one method be compensated to some extent by the strengths of the other method. When closely looking at application C in MoMo, children and adolescents living in Germany show a very low adherence to the 2010 WHO Guideline (World Health Organization,

2010) (as seen in Chapter 5). These results were confirmed by both the accelerometer and the PA questionnaire. Surprisingly, the differences in meeting the Guideline between the measurement methods are much smaller for younger children than for older age groups. Although a continuous overestimation of the self-report was observed in contrast to the accelerometer in all other age groups.

As seen in Chapter 5.3, at elementary school age, approximately one in ten children achieves the activity guideline on each day measured with an accelerometer (ACC), 16 % when using the MoMo physical activity questionnaire (PAQ). With the transition to secondary school, the proportion of young people who meet this activity guideline drops to 2 % (ACC), respectively 7% with PAQ. Finally, between the ages of 14 and 17, only one in 100 (ACC) and four in 100 (PAQ) achieve the required activity guideline. However, these figures must be put into perspective in light of the development of physical activity in childhood and adolescence. The results suggest that with increasing institutionalization of the child's world, due to longer school days and increasing amount of homework, there is less and less time for daily exercise. At the same time, there is a shift in physical activity from everyday activities to sports activities in and out of clubs several times a week, which decisively changes the frequency but also the quality of physical activity (Schmidt et al., 2017).

These developmental processes can explain the apparent paradox in which adolescents play more organized sports in their free time as they age, but at the same time they fulfill the activity guideline less and less frequently. In childhood, everyday activities (e.g., playing outdoors) dominate, which have a lower degree of organization than sports activities and can therefore be practiced more frequently on a daily basis. In this way, children are more likely to meet the Activity Guideline. In contrast, adolescents' activity shifts predominantly to recreational sports (both club and non-club). This type of activity is performed less frequently on a daily basis, but with a longer duration and higher intensity of exercise units. However, again, the true value of physical behavior lies somewhere in between the results of the questionnaire and the accelerometer. A development that, at least on the side of accelerometry, will allow for an even more comprehensive and precise description of the actual behavior of physical activity in the future is **24-hour recording**. As mentioned above, in recent years more attention has been paid to the analysis of complete 24-hour physical activity behaviors (including physical activity (PA), sedentary behavior (SED), and sleep) than to the isolated examination of these behaviors (de Craemer et al., 2021).

Biswas et al. (2015) reported that extended periods of sedentary behavior, regardless of the amount of physical activity performed at other times, are associated with adverse health outcomes. However, this evidence, which at first glance seems inconclusive, is also reflected in various other studies, even if they sometimes come to divergent conclusions (Ekelund et al., 2019; Katzmarzyk et al., 2019; Koster et al., 2012; Stamatakis et al., 2019). In short, the actual dependence between sedentary behavior and physical activity is an important issue that can only be resolved if both of these behaviors are measured simultaneously using the same method and can be analyzed separately but also together in subsequent data analysis. However, physical activity and sedentary behavior are only two aspects of daily physical activity behaviors of people. As introduced in Chapter 1 and also presented again in Figure 7.7, the daily 24 hour cycle of movement and non-movement can be conceptualized as an aggregation of the three entities of physical activity, sedentary behavior, and sleep (Rosenberger et al., 2019; Tremblay et al., 2017). Evaluation and analysis of all data during the 24-hour cycle is a necessary approach for future research.

Figure 7.7: 24-hour Physical Behavior Cycle (adapted from Tremblay et al., 2017).

Note: The movements that take place in the course of a day are divided into two components: On the outside, different categories are presented based on the posture of the body. Inside, divided by energy expenditure, are the three categories of sleep, physical activity, and sedentary behavior. The proportion of space occupied by each behavior in this figure is not prescriptive of the amount of time that should be spent on those behaviors each day.

Many of the results of current research are still based on previous analysis approaches, such as regression-based thresholds to classify different intensities of activity or linear regressions to predict energy consumption. Researchers continue to use many count-based accelerometers from previous studies for reasons of comparability to previously collected data or even economic reasons because the equipment is already available, which is understandable. Therefore, it is also not surprising that PA research is often cluttered with conflicts and debates about which devices to use, where to place them and how to further process the data.
In the future, rather than forgetting past experiences, we should build on them and take more advantage of existing **raw accelerometer data**. It will be equally important to recognize the limitations of raw signal-based analyzes as with regression-based approaches, but new approaches to high-resolution data analysis offer great potential. The capabilities of these new instruments will further enhance the research of PA research.

Accelerometers capture motion in an absolute manner, specifically by deriving the recorded data over time at different intensities. However, if the measurement period is defined as finite, like the 24 hours of a day or the wearing time in MoMo, a compo-

sitional approach can also be applied to physical behavior. Chastin et al. (2015) and Dumuid et al. (2018) are of the view that information on physical behavior should be studied in relation to the entire composition. For example, as explained in Chapter 1, the physical activity behavior of a day is then made up of the three behaviors, activity, sedentary behavior, and sleep. These three behaviors are reassigned in their absolute length each day, partially replacing each other in the process, and ultimately combining to form 24 hours. Recently, a discussion has emerged about how to deal with this closeness, in which various possible approaches are being debated as potential strategies. Aadland, Kvalheim, et al. (2019) provide an overview of this, with compositional analysis on one side (Chastin et al., 2015; Dumuid et al., 2018) and isotemporal substitution models on the other (Mekary et al., 2009). As discussed in Chapter 2.3.5, the application of these statistical approaches is rather complex in comparison with standard procedures of multivariate statistical analyzes (e.g. linear regression, ANOVA). The Granada Consensus Report (Migueles et al., 2021) provides a comprehensive description, discussion, and consensus on the strengths and limitations of the different accelerometer-based descriptors of physical behaviors. Among other things, examples of the application of various methods such as "Linear regression modelling" (Henson et al., 2013; Mora-Gonzalez et al., 2019), "Isotemporal substitution model" (Grgic et al., 2018; Mekary et al., 2009), "Multivariate pattern analysis" (Aadland, Andersen, et al., 2019; Aadland, Kvalheim, et al., 2019; Kvalheim et al., 2018; Rajalahti et al., 2009; Rajalahti & Kvalheim, 2011), "Functional data analysis" (Augustin et al., 2017; Goldsmith et al., 2015; Menai et al., 2017; Ramsay, 2005; Sørensen et al., 2013) as well as first approaches of "Machine learning" (Alaa et al., 2019; Bi et al., 2019; Hua et al., 2018) are presented there.
Currently, there is no gold standard that can be used to test which of the analytical approaches presented is the best for each case. Therefore, there is no recommendation for one analytical approach or the other at this stage. Practical recommendations and the decision tree for choosing an analytical approach given by Migueles et al. (2021) are good for selecting the most appropriate approaches for a particular research question. Although the scientific debate between the two camps is ongoing, caution should be used when comparing the results of analyses of absolute (raw, normalized, and logarithmic transformed) and relative (compositional) data.

In any case, it makes sense to look at both the analyses with absolute and relative values. Multivariate analysis of absolute data patterns showed weak associations for lower intensities and strong associations with cardiometabolic health and locomotor skills for vigorous intensities, but the association patterns were in conflict with compositional data (Aadland, Andersen, et al., 2019). In Chapter 6 we also discuss the difficulties between absolute and relative values in the results. The results of Chapter 6.4 suggest that the percentage intensity distribution of SED, LPA, MPA, and VPA on the weekend is similar to that during the week, including Friday, despite the lower WT (compare Fig. 6.2). Hence, lower MVPA and SED times on the weekend cannot be attributed to a fundamental change in PA behavior. On the contrary, the PB patterns observed during the week also prevailed over the weekend. Since the shorter waking phase on the weekend indicates that less time is available overall, this also results in less time spent for the different PB intensities.

In summary, data pooling and uniform reanalysis of raw data could be a solution to overcome inconsistencies in processing criteria even if they originated from different studies and different devices. Together with the measurement of physical activity behavior in parallel with questionnaires, but also by accelerometers, one could use the **advantages**

of both methods in the subsequent analysis of the data if the connection of the two methods will be linked or even matched in great parts or even completely. In the future, this could also clarify questions as to whether, for example, a daily activity of one hour of moderate intensity is valued more highly than an activity that is performed less frequently, but for a longer time and at a higher intensity? What constitutes the health effect? Does it depend on the frequency or the total amount of activity? All against the premise that public health guidelines on physical behaviors should recognize that behaviors are related to each other, which can affect the guidelines as they have traditionally been understood. For this reason, it is important to record complete physical behavior for 24 hours for at least one week. On the one hand, this allows us to investigate how the behaviors influence each other. For example, what influence do sleeping times have on physical activity and sedentary behavior, and are different waking times on the weekend actually the only influencing factor on very similar physical behavior to school days?

7.4 Final Discussion

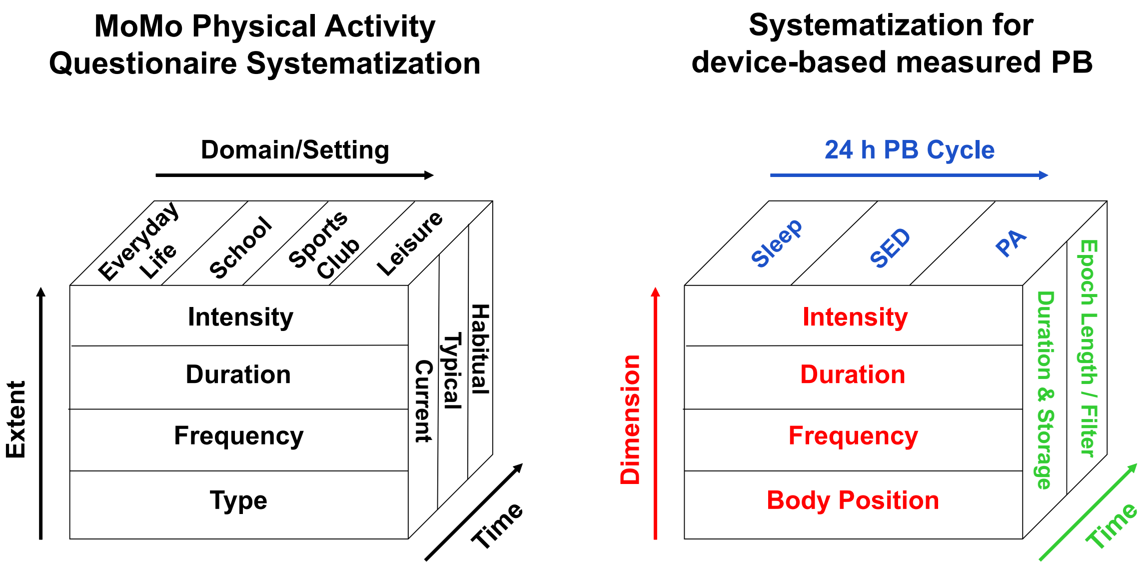
The literature contains a multitude of attempts to systematize physical activity in all its aspects (e.g. Ainsworth et al. (1993) and Bouchard et al. (2012)). The MoMo activity questionnaire was based on the systematization according to Woll (1996) shown in Figure 7.8 on the left. The questionnaire focuses on recording the central aspects of the type, intensity, frequency, and duration of typical or habitual activity. To record these aspects of physical activity as holistically as possible, the MoMo activity questionnaire is differentiated in various domains or settings.

Recording physical activity with accelerometers extends this approach to include intensity of activity, which is rather difficult to capture with questionnaires. With accelerometers, the focus is on recording the pure intensity of the activities. The energy consumption can then be determined by the duration and frequency of the activity, and the body position can also be determined by means of additional sensors or by attaching the device to the thigh. However, current research is still completely at the beginning when it comes to recognizing activity patterns and the type of activity, which is again very easy with questionnaires. This requires very complex algorithms that are not yet sophisticated enough according to the current state-of-the-art.

Similarly to the systematized cube of Woll (1996), which gets to the core of the activity questionnaire when recording the different facets of physical activity, the same systematization also helps when using device-based measured physical activity. As a researcher, sometimes the technical details of accelerometry are excessive when selecting and applying devices to quantify physical behavior. The effects of these decisions on the metrics (activity intensity, energy expenditure, body position, and activity patterns) can occur in a variety of ways. The carrying position (hip, wrist, thigh) and the device selection and the resulting decisions made regarding the recording parameters (epoch length, frequency, memory capacity, recording frequency, and filters) have a large influence on the measured activity.

Different background, such as study design (one-time and repeated measurements) and recording duration (time frame, wear time), as well as data storage and evaluation, must be taken into account when determining the parameters. Finally, the evaluation must adjust several levers (raw data, context information, non-wear time, intensity classification, compliance) depending on the target variables. For this reason and for the thesis described above, we propose a novel construct, a systematization for device-based measured physical behavior (right figure 7.8).

Figure 7.8: On the left the facets of physical activity modified version of Woll (1996) for surveying by activity questionnaire, on the right the facets for device-based measured physical behavior



But we must not stop here. Even if device-based measurement of physical behavior is the next logical step, the results are still too far from the underlying reality in their estimation.

A promising approach for the next generation of devices is the establishment of more user-friendly (validated) equipment, such as (water-proof) smart patches (Schneller et al., 2017). The first commercial products (e.g., biobeat patch or Moio.care smart patch) are already on the market. Currently, these products are still associated with high costs for individual smart patches. In addition, these devices can often only be used once or a limited number of times and must then be disposed of, which produces a lot of waste and is not as sustainable as reusable devices such as typical scientific used accelerometers. Due to technological development, sensors and device sizes are becoming smaller and therefore easier to use. Leading sports manufacturers have already taken a promising approach in this direction by inserting such microchips into sportswear (e.g., shirts such as the Hexokin Smart Shirt or shoes such as the Digitsole Smart Shoe or the Skiin Smart Underwear), which can be used more than once. But even smart clothes still have a very limited lifespan; moisture during washing and the strain on cables and sensors caused by movement limit the amount of time they can be used. Another potential future direction is the development of subdermal accelerometer computer chip implants. The potential in terms of data collection is great; however, this raises major ethical issues as subdermal microchips cannot be easily removed. This would limit human rights with respect to privacy and make them 'transparent humans' with no chance of escaping permanent (possible) observation through data gained within these chips.
Sensor data fusion will also be a key topic in the next years. You can illustrate this quite simply with the example of humans themselves. We constantly collect different kinds of sense impression and then we are able to make predictions from them and anticipate our behavior in an efficient way. For this, we need all the senses of our body. In the future, sensors will also be able to operate and react autonomously in a similar way.

When different sensors are combined into a single compact sensor application, we can gather even more information. With artificial self-learning intelligences, through deep learning algorithms, we can directly evaluate these data in the future. In the setting of MoMo, for example, sensor-based recording during a small movement parkour would be imaginable, the data of which could then be used to automatically create a movement profile, if necessary. Based on these training data, an artificial intelligence could possibly use a personalized identifier that could then recognize different types of movement even in the following week and produce a personalized profile of physical behavior and give concluding tips after the week about how to be more active. All of this will result in even more data.

Making big data available is a key to future studies. These data will no longer be used only in their own study. They will also then be shared, with appropriate agreements, with collaborative partners who can perform analyses in their own areas of expertise. Through these collaborations, new research models can be found, which in turn allow for new funding opportunities. However, science should not forget to think outside the box. Scientifically collected data is normally very valid and reliable, but unfortunately the methods and devices we use for this are not "sexy". Another way to obtain large data sets are collaborations with the big technology companies that can collect massive amounts of data with their trendy and widely used products. It is often a big challenge in research to find enough study participants. The large number of smartphone users worldwide offer easy ways to reach more participants and collect more data than ever before with unified apps. Examples of this are Apple ResearchKit, Google Study Kit, or Android ResearchStack (Sezgin, 2021). Here, companies help aggregate the data and make them available for research. It is important to mention that no one has to participate and that data is not used in any way without explicit consent. More specifically, you first have to sign up and qualify for certain research. This could be combined with the "normal" methods for finding participants. With their smartphone, which most subjects always have with them anyway, they no longer have to go to university, a test site, or another institution to take tests or complete questionnaires. Instead, smartphone's increasingly multifaceted sensors can be used to generate data from anywhere. This could create even greater cross-sectional sources of information.
In addition to finding voluntary participants, survey data are often subject to individual evaluation. This applies to participants ("How often were you active for more than 60 minutes in the last week?"), as well as to the determination of the test results. Apple showed an example in its keynote where Parkinson's sufferers were asked to walk 20 steps and then have the medical staff record how straight or curvy the person was in doing so (on a subjective scale of 0 to 5). A smartphone uses sensors that can accurately determine whether a person is walking straight or how much these deviations are. Similarly, not only 20 steps can be studied in a test, but virtually every step that the subject takes every day. This extends the observation of classic studies from one-time testing to several days, entire months, or even years, during which the smartphone (or smartwatch) permanently tracks the subject. There is promising research in human activity recognition for health research (Straczkiewicz et al., 2021) and there are several examples that have already successfully conducted studies with these toolkits (Jungnickel et al., 2022; Sezgin, 2021). Chu et al. (2022) for example examined quality of life and physical activity using the Apple ResearchKit similar to Wunsch et al. (2021) in the MoMo study. Apple took an unusual and very positive step with ResearchKit in 2015 and other providers have followed since then. The entire framework is open source. This means

that the source code is freely available. This allows the system to be tested independently and applications can also be developed for other devices and operating systems (Android, etc.). On the other hand, there are new problems to be aware of. Within a study app, incorrect information can be entered in self-disclosure that would be directly noticeable in an on-site examination (age information, body measurements, etc.). There is often no replacement for a personal conversation, but here too there may be workarounds in the future, such as video chats like those used in MoMo during the Corona pandemic. However, the biggest obstacles are still privacy, trust, data use policies, and security. Although there are significant efforts in privacy and data use policies, the implementation of such studies should be further examined and researched. This includes how these commercial toolkits can be used in healthcare from a regulatory, policy, and research perspective.

Even without significant changes in study design, some near-term advances can also be derived for the MoMo study from the results and discussions presented. Closer inspection is needed on two key topics: More in-depth evaluations and improvement of future MoMo surveys.

A closer examination of the data in MoMo (e.g. fitness levels, sports disciplines, migration status) will give a more detailed answer as to the reasons behind the results we found in our articles. Furthermore, there are plans to perform an in-depth analysis of the associations between physical activity and different health-related parameters (eg, obesity) and socioeconomic parameters (eg, education). This includes investigation of cross-links and trends between questionnaires and device-based collected activity behavior data (eg, physical activity differences between groups in both data sets). It is certainly also interesting to see how media use affects activity behavior. For example, the hypothesis of reading displacement (Beentjes & van der Voort, 1989) could be transferred to physical activity behavior. Originally, the displacement hypothesis stated that increased media use among children can displace their reading activities. However, media consumption itself has changed considerably in recent years, and screen time is no longer the same as TV time. A transfer of this hypothesis would be, for example, that increased media consumption tends to displace light activity and sleep, whereas more intensive physical activities do not lose any shares.
The MoMo study also has the potential to look at specific populations that are normally underrepresented. Studies typically do not address socioeconomic status, race, or ethnicity (Pfeiffer et al., 2022). However, since MoMo has collected these parameters, it is possible to at least check to what extent there are differences between these populations. Another possibility is to compare the representative data sets of MoMo with the data sets of other studies that only include subjects with specific characteristics. Some first collaborations already exist where it is possible to compare the results based on the study design parameters and the data collection defined in this thesis. Siaplaouras et al. (2020) is looking at physical activity among children with congenital heart defects in a nationwide survey in Germany, Götte et al. (2022) did compare the physical activity of 9-15 year old pediatric cancer survivors with the MoMo sample and the ActiMON study (TARISMA Konsortium, ActiMON Project website accessed on 14.10.2022) monitors activity in adolescents and young adults with inflammatory rheumatic musculoskeletal diseases and will also compare their data with the MoMo sample.

The possibilities for further development and the potential for improvement of the MoMo study should be relatively straightforward by now. The implications of this work for future

MoMo surveys, especially in the area of device-based physical activity assessment are 24-hour recordings and getting an even more realistic view in the physical behavior data.
Our results have highlighted the importance of extending the measurement to a 24-hour recording, which thus also includes sleeping hours. These recordings will substantially improve our understanding of the complete physical behavior cycle and will add the next step toward a complete understanding of the interaction between sleep, sedentary behavior, and physical activity. The widespread use of ActiGraph devices and the original idea of integrating study results into the ICAD database were certainly the right decision in 2017 for the last surveys. Now we have the availability of more open source methods for data processing, and the subsequent analysis and more research into solving the comparability problem is already underway. Since the entire 24-hour behavior cycle is a key area to explore, ActivPAL with its attachment to the thigh offers a suitable solution to collect data in all three categories (physical activity, sedentary behavior, and sleep).
Data pooling and big data are terms that are now becoming increasingly familiar in sports science and public health. The large and representative survey, as it has been conducted in MoMo so far, is still very time-consuming and cost-intensive despite technical progress and should continue to approach these two concepts. In the near future, it should be reconsidered whether it is possible to increase the sample size by more cost-effective methods, on the one hand, and to reduce the bias in the recruitment of test persons, on the other. Commercial research frameworks, as mentioned above, might be an idea. Combined with existing expertise in recruiting test participants, it would be possible, for example, to test entire classes or even entire schools by improving the process flow during on-site examination instead of inviting individual test participants. A digitized questionnaire as an app on the smartphone could replace time-consuming interviews, and children and adolescents could be tracked by sensors on their own smartphones instead of accelerometers around the waist or on the thigh.

Looking back at the introduction of this thesis, we have now arrived at the whole through numerous individual parts. We have been able to establish the importance of each part, each technical decision in capturing physical activity, and, as a part of the whole, the use of accelerometers to describe physical activity behavior in MoMo. We have considered measured activity not only as a whole, but have also taken a closer look at each step that is taken to record it.
The main objectives were to clarify technical decisions that should be made when using accelerometers in large epidemiological studies of activity behavior research, investigate the implications of different evaluation methods in the preprocessing of accelerometer data in general, and provide reference data to practitioners and researchers when comparing their accelerometer-measured physical activity with the MoMo study.
We clarified what aspects should be considered in planning and evaluation before the study in Chapter 2 and what technical decisions should be made when using accelerometers in large epidemiological studies of activity behavior research; we introduced a specific overview of the decisions for the methods and protocols used to assess device-based physical activity in children and adolescents with accelerometers specifically in the MoMo-Study in Chapter 3; we examined the implications of choosing various metrics when processing the recorded data in Chapter 4 and it became clear why it is so important for reproducibility that all technical decisions in the processing and evaluation of device-based measured physical activity are documented; we presented differences between self-reported and device-based measured in Chapter 5 and found that children and adolescents living in Germany show very low adherence to the WHO physical activ-

ity guidelines with both methods; and lastly in Chapter 6 we focused on a typical day of physical activity intensities measured by accelerometer and compared weekdays and weekend days. This work provides a valuable contribution to improving research practice within large epidemiological cohort studies, provides mandatory information for survey replication that focuses on other contexts, and general measurement of physical activity with accelerometers in the MoMo study. The conclusion that emerges from the entire work is therefore as follows.

7.5 Conclusion

It's a fact that children are not getting enough physical activity these days, no matter what algorithms are used to evaluate the data, the result remains basically the same. This leads to a variety of health problems, such as obesity and an increased risk of cardiovascular disease.

Right now, however, we are at a crucial transition point where we have to take additional actions in order to capture the physical behavior of children as close to reality as possible. This way, in the future, their physical behavior can be supported as individually as possible. More precise measurements will allow critical decisions to be made based on the most accurate data, which in turn will lead to improved efficiency and performance in intervention. Accurate measurement devices are therefore of utmost importance in the research of physical behavior.

What happens if we do not continue to work on the technical measurement and become even more accurate? We will continue to produce inaccurate and estimated values that do not correspond to reality. Then these data suggest a false baseline for scientists who interpret the data and develop interventions based on it. The pure methodology must therefore become better and better and at the same time less and less biased by subjective decisions in the selection of algorithms. We physical behavior scientists at the technical level have a social responsibility, especially towards our children. As the father of a young daughter, I am aware of this every day, and anyone who takes a closer look at the current generation of children will be able to see it for themselves. We need to ensure that physical activity is promoted from an early age and that the world is a better place when our children grow up. Our children should not suffer from lack of physical activity and the resulting health risks as the current generations.

We have and will be assessing large amounts of data and we cannot afford to just collect these data and not analyze it. It must not be the case that data pooling databases such as the ICAD database do not include further studies due to a lack of financial and human resources. In this case, politics must free up further funds to enable more comprehensive and precise reporting. The cost of missing out is too great. The data sets are there, they are in the right format, they are collected with the same study design and only need to be combined and analyzed. In the case of the ICAD database, this is possible not only on a national level, but also on an international level. Therefore, the WHO, the EU, and also the individual country governments together must provide further funds so that this global problem can be studied in more detail. Only talking about the pandemic of inactivity does not help, the available data must be analyzed in depth, and for this big data is the appropriate approach.

However, we cannot stop here. We need to be even more precise in capturing data and we need to portray the actual movement behavior even better. Therefore, more validation studies with short epoch lengths are needed, especially for young children, and the formerly proprietary ActiGraph signal preprocessing algorithms that are now released by the manufacturer need to be used to compare the data with studies using devices from other manufacturers. In addition, we mentioned two additional issues in Chapter 7. In particular, we expect that the issue of development on how data will be collected in the future will be based primarily on the simultaneous recording and evaluation of all aspects of physical behavior. It will be increasingly important to record the full 24 hours a day over a week or more and reduce the non-wear time to a minimum by using smaller and waterproof devices and when possibly with sensors directly attached to or subdermal implants in the body. This will allow us to further reduce the large gaps we currently still have in the actual mapping of movement behavior. With the issue of comparability, we complete the picture of perspectives and emphasize the importance of open source and data-pooling methods when analyzing physical behavior data in the future. This will include data from different studies, and studies will increasingly use combined survey approaches in which reports and devices are linked so closely that the limitations of one method can be compensated to some extent by the strengths of the other method.

Here, too, we return one last time to the beginning of this work. The international CAPA workshop, which, among other things, resulted in three consensus articles, one of which is part of this work, was the right starting signal at that time. This form of international cooperation has had a significant influence on this work and has pushed the development in this area considerably ahead. Now, it is precisely this collaboration that needs to be refocused and established on a regular basis. It is necessary to bring the international experts to one table, to other CAPA workshops. Experts from sports science, sleep research and sports informatics need to put their heads together with electrical engineers and experts for accelerometers, experts for data pooling and big data analyses, and think about how we can refine the data collection methods to such an extent that the bias in the data acquisition disappears and how we can offer a realistic picture of the movement behavior. Here, CAPA can be pivotal and become an international meet-up and catalyst in physical behavior research. Even in light of the fact that some internal restructuring will be necessary, and while it may no longer roll off the tongue as easily, the "Center for the Assessment of Physical Activity" should prepare to rename itself to CAP-B, a "Center for the Assessment of Physical Behavior".

References Discussion and Conclusion

Aadland, E., Andersen, L. B., Resaland, G. K., & Kvalheim, O. M. (2019). Interpretation of Multivariate Association Patterns between Multicollinear Physical Activity Accelerometry Data and Cardiometabolic Health in Children-A Tutorial. *Metabolites*, *9*(7). https://doi.org/10.3390/metabo9070129.

Aadland, E., Kvalheim, O. M., Anderssen, S. A., Resaland, G. K., & Andersen, L. B. (2019). Multicollinear physical activity accelerometry data and associations to cardiometabolic health: challenges, pitfalls, and potential solutions. *The international journal of behavioral nutrition and physical activity*, *16*(1), 74. https://doi.org/10.1186/s12966-019-0836-z

ActiGraph. (2018). Low frequency extension explained. Retrieved August 16, 2022, from https://actigraphcorp.my.site.com/support/s/article/Low-Frequency-Extension-Explained

Ainsworth, B. E., Haskell, W. L., Leon, A. S., Jacobs, D. R., Montoye, H. J., Sallis, J. F., & Paffenbarger, R. S. (1993). Compendium of physical activities: classification of energy costs of human physical activities. *Medicine and science in sports and exercise*, *25*(1), 71–80. https://doi.org/10.1249/00005768-199301000-00011

Alaa, A. M., Bolton, T., Di Angelantonio, E., Rudd, J. H. F., & van der Schaar, M. (2019). Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. *PLoS One*, *14*(5), e0213653. https://doi.org/10.1371/journal.pone.0213653

Ancoli-Israel, S., Cole, R., Alessi, C., Chambers, M., Moorcroft, W., & Pollak, C. P. (2003). The role of actigraphy in the study of sleep and circadian rhythms. *Sleep*, *26*(3), 342–392. https://doi.org/10.1093/sleep/26.3.342

Arundell, L., Salmon, J., Koorts, H., Contardo Ayala, A. M., & Timperio, A. (2019). Exploring when and how adolescents sit: cross-sectional analysis of activPAL-measured patterns of daily sitting time, bouts and breaks. *BMC public health*, *19*(1), 653. https://doi.org/10.1186/s12889-019-6960-5

Arvidsson, D., Fridolfsson, J., Börjesson, M., Andersen, L. B., Ekblom, Ö., Dencker, M., & Brønd, J. C. (2019). Re-examination of accelerometer data processing and calibration for the assessment of physical activity intensity. *Scandinavian Journal of Medicine & Science in Sports*, *29*(10), 1442–1452. https://doi.org/10.1111/sms.13470

Atkin, A. J., Biddle, S. J. H., Broyles, S. T., Chinapaw, M. J. M., Ekelund, U., Esliger, D. W., Hansen, B. H., Kriemler, S., Puder, J. J., Sherar, L. B., & van Sluijs, E. M. F. (2017). Harmonising data on the correlates of physical activity and sedentary behaviour in young people: Methods and lessons learnt from the international Children's Accelerometry database (ICAD). *International Journal of Behavioral Nutrition and Physical Activity*, *14*(1), 174. https://doi.org/10.1186/s12966-017-0631-7

Augustin, N. H., Mattocks, C., Faraway, J. J., Greven, S., & Ness, A. R. (2017). Modelling a response as a function of high-frequency count data: The association between physical activity and fat mass. *Statistical methods in medical research*, *26*(5), 2210–2226. https://doi.org/10.1177/0962280215595832

Bai, J., Di, C., Xiao, L., Evenson, K. R., LaCroix, A. Z., Crainiceanu, C. M., & Buchner, D. M. (2016). An Activity Index for Raw Accelerometry Data and Its Comparison with Other Activity Metrics. *PLoS One*, *11*(8), e0160644. https://doi.org/10.1371/journal.pone.0160644.

Bakrania, K., Yates, T., Rowlands, A. V., Esliger, D. W., Bunnewell, S., Sanders, J., Davies, M. J., Khunti, K., & Edwardson, C. L. (2016). Intensity Thresholds on Raw Acceleration Data: Euclidean Norm Minus One (ENMO) and Mean Amplitude Deviation (MAD) Approaches. *PLOS ONE*, *11*(10), e0164045. https://doi.org/10.1371/journal.pone.0164045

Banda, J. A., Haydel, K. F., Davila, T., Desai, M., Bryson, S., Haskell, W. L., Matheson, D., & Robinson, T. N. (2016). Effects of Varying Epoch Lengths, Wear Time Algorithms, and Activity Cut-Points on Estimates of Child Sedentary Behavior and Physical Activity from Accelerometer Data. *PLOS ONE*, *11*(3), e0150534. https://doi.org/10.1371/journal.pone.0150534

Barboza, L. L., Gandarela, L., Santana, J. G. S., Silva, E. C. M., Machado, E. S., Silva, R. J. S., Gomes, T. N., & Silva, D. R. (2021). Agreement Between Different Days of activPAL and Actigraph GT3X Measurement of Sedentary Behavior and Physical Activity During the School Hours in Elementary Children. *Journal for the Measurement of Physical Behaviour*, *4*(2), 111–117. https://doi.org/10.1123/jmpb.2020-0025

Beentjes, J. W., & van der Voort, T. H. (1989). Television and Young People's Reading Behaviour: A Review of Research. *European Journal of Communication*, *4*(1), 51–77. https://doi.org/10.1177/0267323189004001004

Bi, Q., Goodman, K. E., Kaminsky, J., & Lessler, J. (2019). What is Machine Learning? A Primer for the Epidemiologist. *American journal of epidemiology*, *188*(12), 2222–2239. https://doi.org/10.1093/aje/kwz189

Biswas, A., Oh, P. I., Faulkner, G. E., Bajaj, R. R., Silver, M. A., Mitchell, M. S., & Alter, D. A. (2015). Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. *Annals of internal medicine*, *162*(2), 123–132. https://doi.org/10.7326/M14-1651

Blackwood, J., Suzuki, R., Webster, N., Karczewski, H., Ziccardi, T., & Shah, S. (2022). Use of activPAL to Measure Physical Activity in Community-Dwelling Older Adults: A Systematic Review. *Archives of rehabilitation research and clinical translation*, *4*(2), 100190. https://doi.org/10.1016/j.arrct.2022.100190

Bouchard, C., Blair, S. N., & Haskell, W. L. (2012). *Physical Activity and Health*. Human Kinetics. https://books.google.de/books?id=tO96DwAAQBAJ

Breau, B., Coyle-Asbil, H. J., Haines, J., Ma, D. W. L., & Vallis, L. A. (2022). ActiGraph Cutpoints Impact Physical Activity and Sedentary Behavior Outcomes in Young Children. *Journal for the Measurement of Physical Behaviour*, *5*(2), 85–96. https://doi.org/10.1123/jmpb.2021-0042

Brønd, J. C., Aadland, E., Andersen, L. B., Resaland, G. K., Andersen, S. A., & Arvidsson, D. (2019). The ActiGraph counts processing and the assessment of vigorous activity. *Clinical Physiology and Functional Imaging*, *39*(4), 276–283. https://doi.org/10.1111/cpf.12571

Brønd, J. C., Andersen, L. B., & Arvidsson, D. (2017). Generating ActiGraph Counts from Raw Acceleration Recorded by an Alternative Monitor. *Medicine and science in sports and exercise*, *49*(11), 2351–2360. https://doi.org/10.1249/MSS.0000000000001344

Brønd, J. C., & Arvidsson, D. (2016). Sampling frequency affects the processing of Actigraph raw acceleration data to activity counts. *Journal of applied physiology (Bethesda, Md. : 1985)*, *120*(3), 362–369. https://doi.org/10.1152/japplphysiol.00628.2015

Brondeel, R., Kestens, Y., Rahimipour Anaraki, J., Stanley, K., Thierry, B., & Fuller, D. (2021). Converting Raw Accelerometer Data to Activity Counts Using Open-Source Code: Implementing a MATLAB Code in Python and R, and Comparing the Results to ActiLife. *Journal for the Measurement of Physical Behaviour*, *4*(3), 205–211. https://doi.org/10.1123/jmpb.2019-0063

Burchartz, A., Anedda, B., Auerswald, T., Giurgiu, M., Hill, H., Ketelhut, S., Kolb, S., Mall, C., Manz, K., Nigg, C. R., Reichert, M., Sprengeler, O., Wunsch, K., & Matthews, C. E. (2020). Assessing physical behavior through accelerometry – State of the science, best practices and future directions. *Psychology of Sport and Exercise*, *49*, 101703. https://doi.org/10.1016/j.psychsport.2020.101703

Burchartz, A., Manz, K., Anedda, B., Niessner, C., Oriwol, D., Schmidt, S. C. E., & Woll, A. (2020). Measurement of Physical Activity and Sedentary Behavior by Accelerometry Among a Nationwide Sample from the KiGGS and MoMo Study: Study Protocol. *JMIR Research Protocols*, *9*(7), e14370. https://doi.org/10.2196/14370

Bussmann, J. B. J., & van den Berg-Emons, R. J. G. (2013). To total amount of activity..... and beyond: perspectives on measuring physical behavior. *Frontiers in Psychology*, *4*, 463. https://doi.org/10.3389/fpsyg.2013.00463

Carlson, J. A., Ridgers, N. D., Nakandala, S., Zablocki, R., Tuz-Zahra, F., Bellettiere, J., Hibbing, P. R., Steel, C., Jankowska, M. M., Rosenberg, D. E., Greenwood-Hickman, M. A., Zou, J., LaCroix, A. Z., Kumar, A., & Natarajan, L. (2022). CHAP-child: an open source method for estimating sit-to-stand transitions and sedentary bout patterns from hip accelerometers among children. *International Journal of Behavioral Nutrition and Physical Activity*, *19*(1), 109. https://doi.org/10.1186/s12966-022-01349-2

Cavagna, G. A., & Franzetti, P. (1986). The determinants of the step frequency in walking in humans. *The Journal of physiology*, *373*(1), 235–242. https://physoc.onlinelibrary.wiley.com/doi/pdf/10.1113/jphysiol.1986.sp016044

Cavagna, G. A., Willems, P. A., Franzetti, P., & Detrembleur, C. (1991). The two power limits conditioning step frequency in human running. *The Journal of physiology*, *437*(1), 95–108. https://physoc.onlinelibrary.wiley.com/doi/pdf/10.1113/jphysiol.1991.sp018586

Chaput, J.-P., Willumsen, J. F., Bull, F. C., Chou, R., Ekelund, U., Firth, J., Jago, R., Ortega, F. B., & Katzmarzyk, P. T. (2020). 2020 WHO guidelines on physical activity and sedentary behaviour for children and adolescents aged 5-17 years: summary of the evidence. *The international journal of behavioral nutrition and physical activity*, *17*(1), 141. https://doi.org/10.1186/s12966-020-01037-z

Chastin, S. F. M., Palarea-Albaladejo, J., Dontje, M. L., & Skelton, D. A. (2015). Combined Effects of Time Spent in Physical Activity, Sedentary Behaviors and Sleep on Obesity and Cardio-Metabolic Health Markers: A Novel Compositional Data Analysis Approach. *PLoS One*, *10*(10), e0139984. https://doi.org/10.1371/journal.pone.0139984.

Chen, K. Y., Acra, S. A., Majchrzak, K., Donahue, C. L., Baker, L., Clemens, L., Sun, M., & Buchowski, M. S. (2003). Predicting energy expenditure of physical activity using hip- and wrist-worn accelerometers. *Diabetes technology & therapeutics*, *5*(6), 1023–1033. https://doi.org/10.1089/152091503322641088

Choi, L., Liu, Z., Matthews, C. E., & Buchowski, M. S. (2011). Validation of accelerometer wear and nonwear time classification algorithm. *Medicine and science in sports and exercise*, *43*(2), 357–364. https://doi.org/10.1249/MSS.0b013e3181ed61a3

Chu, B., O'Connor, D. M., Wan, M., Barnett, I., Shou, H., Judson, M., & Rosenbach, M. (2022). Quality of Life and Physical Activity in 629 Individuals With Sarcoidosis: Prospective, Cross-sectional Study Using Smartphones (Sarcoidosis App). *JMIR mHealth and uHealth*, *10*(8), e38331. https://doi.org/10.2196/38331

Clevenger, K. A., Mackintosh, K. A., McNarry, M. A., Pfeiffer, K. A., Montoye, A. H. K., & Brønd, J. C. (2022). Comparison of Child and Adolescent Physical Activity Levels From Open-Source Versus ActiGraph Counts. *Journal for the Measurement of Physical Behaviour*, *5*(2), 120–128. https://doi.org/10.1123/jmpb.2021-0057

Clevenger, K. A., Montoye, A. H. K., van Camp, C. A., Strath, S. J., & Pfeiffer, K. A. (2022). Methods for estimating physical activity and energy expenditure using raw accelerometry data or novel analytical approaches: a repository, framework, and reporting guidelines. *Physiological Measurement*, *43*(9). https://doi.org/10.1088/1361-6579/ac89c9

Clevenger, K. A., Pfeiffer, K. A., Mackintosh, K. A., McNarry, M. A., Brønd, J. C., Arvidsson, D., & Montoye, A. H. K. (2019). Effect of sampling rate on acceleration and counts of hip- and wrist-worn ActiGraph accelerometers in children. *Physiological Measurement*, *40*(9), 095008. https://doi.org/10.1088/1361-6579/ab444b

Colley, R. C., Carson, V., Garriguet, D., Janssen, I., Roberts, K. C., & Tremblay, M. S. (2017). *Physical activity of Canadian children and youth, 2007 to 2015*. Statistics Canada.

Colley, R. C., Garriguet, D., Janssen, I., Craig, C. L., Clarke, J., & Tremblay, M. S. (2011). Physical activity of Canadian adults: accelerometer results from the 2007 to 2009 Canadian Health Measures Survey. *Health Reports*, *22*(1), 7–14.

Corder, K., van Sluijs, E. M. F., Wright, A., Whincup, P., Wareham, N. J., & Ekelund, U. (2009). Is it possible to assess free-living physical activity and energy expenditure in young people by self-report? *The American journal of clinical nutrition*, *89*(3), 862–870. https://doi.org/10.3945/ajcn.2008.26739

Coyle-Asbil, H. J., Breau, B., Ma, D. W. L., Haines, J., & Vallis, L. A. (2020). Examining the effects of applying ActiGraph low-frequency extension feature to analyze the sleeping behaviours of preschool-aged children. *Applied physiology, nutrition, and metabolism = Physiologie appliquee, nutrition et metabolisme*, *45*(12), 1396–1399. https://doi.org/10.1139/apnm-2019-0969

Crowley, P., Skotte, J., Stamatakis, E., Hamer, M., Aadahl, M., Stevens, M. L., Rangul, V., Mork, P. J., & Holtermann, A. (2019). Comparison of physical behavior estimates from three different thigh-worn accelerometers brands: a proof-of-concept for the Prospective Physical Activity, Sitting, and Sleep consortium (ProPASS). *The international journal of behavioral nutrition and physical activity*, *16*(1), 65. https://doi.org/10.1186/s12966-019-0835-0

de Craemer, M., Decraene, M., Willems, I., Buysse, F., van Driessche, E., & Verbestel, V. (2021). Objective Measurement of 24-Hour Movement Behaviors in Preschool Children Using Wrist-Worn and Thigh-Worn Accelerometers. *International journal of environmental research and public health*, *18*(18). https://doi.org/10.3390/ijerph18189482

Diaz, K. M., Howard, V. J., Hutto, B., Colabianchi, N., Vena, J. E., Safford, M. M., Blair, S. N., & Hooker, S. P. (2017). Patterns of Sedentary Behavior and Mortality in U.S. Middle-Aged and Older Adults: A National Cohort Study. *Annals of internal medicine*, *167*(7), 465–475. https://doi.org/10.7326/M17-0212

Doherty, A., Jackson, D., Hammerla, N., Plötz, T., Olivier, P., Granat, M. H., White, T., van Hees, V. T., Trenell, M. I., Owen, C. G., Preece, S. J., Gillions, R., Sheard, S., Peakman, T., Brage, S., & Wareham, N. J. (2017). Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. *PLoS One*, *12*(2), e0169649. https://doi.org/10.1371/journal.pone.0169649

Dowd, K. P., Harrington, D. M., Bourke, A. K., Nelson, J., & Donnelly, A. E. (2012). The measurement of sedentary patterns and behaviors using the activPAL™ Professional physical activity monitor. *Physiological measurement*, *33*(11), 1887–1899. https://doi.org/10.1088/0967-3334/33/11/1887

Dowd, K. P., Harrington, D. M., & Donnelly, A. E. (2012). Criterion and concurrent validity of the activPAL™ professional physical activity monitor in adolescent females. *PLoS One*, *7*(10), e47633. https://doi.org/10.1371/journal.pone.0047633

Dumuid, D., Stanford, T. E., Martín-Fernández, J.-A., Pedišić, Ž., Maher, C. A., Lewis, L. K., Hron, K., Katzmarzyk, P. T., Chaput, J.-P., Fogelholm, M., Hu, G., Lambert, E. V., Maia, J., Sarmiento, O. L., Standage, M., Barreira, T. V., Broyles, S. T., Tudor-Locke, C. E., Tremblay, M. S., & Olds, T. (2018). Compositional data analysis for physical activity, sedentary time and sleep research. *Statistical methods in medical research*, *27*(12), 3726–3738. https://doi.org/10.1177/0962280217710835.

Dwyer, T., Coonan, W. E., Leitch, D. R., Hetzel, B. S., & Baghurst, R. A. (1983). An investigation of the effects of daily physical activity on the health of primary school students in South Australia. *International journal of epidemiology*, *12*(3), 308–313. https://doi.org/10.1093/ije/12.3.308

Eberhardt, T., Bös, K., & Niessner, C. (2021). The Fitness Barometer: A Best Practice Example for Monitoring Motor Performance With Pooled Data Collected From Practitioners. *Frontiers in public health*, *9*, 720589. https://doi.org/10.3389/fpubh.2021.720589

Edwardson, C. L., Winkler, E. A. H., Bodicoat, D. H., Yates, T., Davies, M. J., Dunstan, D. W., & Healy, G. N. (2017). Considerations when using the activPAL monitor in field-based research with adult populations. *Journal of sport and health science*, *6*(2), 162–178. https://doi.org/10.1016/j.jshs.2016.02.002

Ekelund, U., Tarp, J., Steene-Johannessen, J., Hansen, B. H., Jefferis, B., Fagerland, M. W., Whincup, P., Diaz, K. M., Hooker, S. P., Chernofsky, A., Larson, M. G., Spartano, N., Vasan, R. S., Dohrn, I.-M., Hagströmer, M., Edwardson, C., Yates, T., Shiroma, E., Anderssen, S. A., & Lee, I.-M. (2019). Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: systematic review and harmonised meta-analysis. *BMJ (Clinical research ed.)*, *366*, l4570. https://doi.org/10.1136/bmj.l4570

Ekelund, U., Tomkinson, G., & Armstrong, N. (2011). What proportion of youth are physically active? Measurement issues, levels and recent time trends. *British journal of sports medicine*, *45*(11), 859–865. https://doi.org/10.1136/bjsports-2011-090190

Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S., & McMurray, R. G. (2008). Calibration of two objective measures of physical activity for children. *Journal of sports sciences*, *26*(14), 1557–1565. https://doi.org/10.1080/02640410802334196

Freedson, P. S., & John, D. (2013). Comment on "estimating activity and sedentary behavior from an accelerometer on the hip and wrist". *Medicine and science in sports and exercise*, *45*(5), 962–963. https://doi.org/10.1249/MSS.0b013e31827f024d

Full, K. M., Kerr, J., Grandner, M. A., Malhotra, A., Moran, K., Godoble, S., Natarajan, L., & Soler, X. (2018). Validation of a physical activity accelerometer device worn on the hip and wrist against polysomnography. *Sleep health*, *4*(2), 209–216. https://doi.org/10.1016/j.sleh.2017.12.007

Gaser, D., Peters, C., Götte, M., Oberhoffer-Fritz, R., Feuchtinger, T., Schmid, I., von Luettichau, I., & Kesting, S. (2022). Analysis of self-reported activities of daily living, motor performance and physical activity among children and adolescents with cancer: Baseline data from a randomised controlled trial assessed shortly after diagnosis of leukaemia or non-Hodgkin lymphoma. *European journal of cancer care*, *31*(2), e13559. https://doi.org/10.1111/ecc.13559

German National Cohort Consortium. (2014). The German National Cohort: aims, study design and organization. *European journal of epidemiology*, *29*(5), 371–382. https://doi.org/10.1007/s10654-014-9890-7

Giurgiu, M., Kolb, S., Nigg, C., Burchartz, A., Timm, I., Becker, M., Rulf, E., Doster, A.-K., Koch, E. D., Bussmann, J. B. J., Nigg, C. R., Ebner-Priemer, U. W., & Woll, A. (2022). Assessment of 24-hour physical behaviour in children and adolescents via wearables: a systematic review of free-living validation studies. *BMJ Open Sport & Exercise Medicine*, *8*(2), e001267. https://doi.org/10.1136/bmjsem-2021-001267

Goldsmith, J., Zipunnikov, V., & Schrack, J. (2015). Generalized multilevel function-on-scalar regression and principal component analysis. *Biometrics*, *71*(2), 344–353. https://doi.org/10.1111/biom.12278

Götte, M., Basteck, S., Beller, R., Gauß, G., Schmidt, S., Burchartz, A., Kolb, S., Grydeland, M., & Reinhardt, D. (2022). Physical activity in 9–15 year-old pediatric cancer survivors compared to a nationwide sample. *Journal of Cancer Research and Clinical Oncology*. https://doi.org/10.1007/s00432-022-04392-5

Grant, P. M., Ryan, C. G., Tigbe, W. W., & Granat, M. H. (2006). The validation of a novel activity monitor in the measurement of posture and motion during everyday activities. *British journal of sports medicine*, *40*(12), 992–997. https://doi.org/10.1136/bjsm.2006.030262

Grgic, J., Dumuid, D., Bengoechea, E. G., Shrestha, N., Bauman, A., Olds, T., & Pedisic, Z. (2018). Health outcomes associated with reallocations of time between sleep, sedentary behaviour, and physical activity: a systematic scoping review of isotemporal substitution studies. *International Journal of Behavioral Nutrition and Physical Activity*, *15*(1), 69. https://doi.org/10.1186/s12966-018-0691-3

Henson, J., Yates, T., Biddle, S. J. H., Edwardson, C. L., Khunti, K., Wilmot, E. G., Gray, L. J., Gorely, T., Nimmo, M. A., & Davies, M. J. (2013). Associations of objectively measured sedentary behaviour and physical activity with markers of cardiometabolic health. *Diabetologia*, *56*(5), 1012–1020. https://doi.org/10.1007/s00125-013-2845-9

Hidde, M. C., Bandera, V., Lyden, K., Braun, B., & Leach, H. J. (2021). Wrist-worn Versus Thigh-worn Accelerometers: Compliance And Participant Perspectives On Accelerometer Placement. *Medicine & Science in Sports & Exercise*, *53*(8S), 233. https://doi.org/10.1249/01.mss.0000761736.86114.8f

Hua, A., Quicksall, Z., Di, C., Motl, R., LaCroix, A. Z., Schatz, B., & Buchner, D. M. (2018). Accelerometer-based predictive models of fall risk in older women: a pilot study. *NPJ digital medicine*, *1*, 25. https://doi.org/10.1038/s41746-018-0033-5

Hurter, L., Fairclough, S. J., Knowles, Z. R., Porcellato, L. A., Cooper-Ryan, A. M., & Boddy, L. M. (2018). Establishing Raw Acceleration Thresholds to Classify Sedentary and Stationary Behaviour in Children. *Children (Basel, Switzerland)*, *5*(12). https://doi.org/10.3390/children5120172

Jungnickel, T., von Jan, U., & Albrecht, U.-V. (2022). Implementation of Mobile Psychological Testing on Smart Devices: Evaluation of a ResearchKit-Based Design Approach for the Implicit Association Test. *Frontiers in digital health*, *4*, 785591. https://doi.org/10.3389/fdgth.2022.785591

Katzmarzyk, P. T., Powell, K. E., Jakicic, J. M., Troiano, R. P., Piercy, K., & Tennant, B. (2019). Sedentary Behavior and Health: Update from the 2018 Physical Activity Guidelines Advisory Committee. *Medicine and science in sports and exercise*, *51*(6), 1227–1241. https://doi.org/10.1249/MSS.0000000000001935

Kim, Y., Barry, V. W., & Kang, M. (2015). Validation of the ActiGraph GT3X and activPAL Accelerometers for the Assessment of Sedentary Behavior. *Measurement in Physical Education and Exercise Science*, *19*(3), 125–137. https://doi.org/10.1080/1091367X.2015.1054390

Kloe, M., Niessner, C., Woll, A., & Bös, K. (2019). Open Data im sportwissenschaftlichen Anwendungsfeld motorischer Tests. *German Journal of Exercise and Sport Research*, *49*(4), 503–513. https://doi.org/10.1007/s12662-019-00620-2

Koster, A., Caserotti, P., Patel, K. V., Matthews, C. E., Berrigan, D., van Domelen, D. R., Brychta, R. J., Chen, K. Y., & Harris, T. B. (2012). Association of sedentary time with mortality independent of moderate to vigorous physical activity. *PLoS One*, *7*(6), e37696. https://doi.org/10.1371/journal.pone.0037696

Kozey-Keadle, S., Libertine, A., Lyden, K., Staudenmayer, J., & Freedson, P. S. (2011). Validation of wearable monitors for assessing sedentary behavior. *Medicine and science in sports and exercise*, *43*(8), 1561–1567. https://doi.org/10.1249/MSS.0b013e31820ce174

Kvalheim, O. M., Arneberg, R., Grung, B., & Rajalahti, T. (2018). Determination of optimum number of components in partial least squares regression from distributions of the root-mean-squared error obtained by Monte Carlo resampling. *Journal of Chemometrics*, *32*(4), e2993. https://doi.org/10.1002/cem.2993

Leppänen, M. H., Migueles, J. H., Abdollahi, A. M., Engberg, E., Ortega, F. B., & Roos, E. (2022). Comparing estimates of physical activity in children across different cut-points and the associations with weight status. *Scandinavian Journal* *of Medicine & Science in Sports*, *32*(6), 971–983. https://doi.org/10.1111/sms.14147

Lobelo, F., Rohm Young, D., Sallis, R., Garber, M. D., Billinger, S. A., Duperly, J., Hutber, A., Pate, R. R., Thomas, R. J., Widlansky, M. E., McConnell, M. V., & Joy, E. A. (2018). Routine Assessment and Promotion of Physical Activity in Healthcare Settings: A Scientific Statement From the American Heart Association. *Circulation*, *137*(18), e495–e522. https://doi.org/10.1161/CIR.0000000000000559

Logan, G. R. M., Duncan, S., Harris, N. K., Hinckson, E. A., & Schofield, G. (2016). Adolescent physical activity levels: discrepancies with accelerometer data analysis. *Journal of sports sciences*, *34*(21), 2047–2053. https://doi.org/10.1080/02640414.2016.1150599

Lynch, B. A., Kaufman, T. K., Rajjo, T. I., Mohammed, K., Kumar, S., Murad, M. H., Gentile, N. E., Koepp, G. A., McCrady-Spitzer, S. K., & Levine, J. A. (2019). Accuracy of Accelerometers for Measuring Physical Activity and Levels of Sedentary Behavior in Children: A Systematic Review. *Journal of primary care & community health*, *10*, 2150132719874252. https://doi.org/10.1177/2150132719874252

Matthews, C. E., Chen, K. Y., Freedson, P. S., Buchowski, M. S., Beech, B. M., Pate, R. R., & Troiano, R. P. (2008). Amount of time spent in sedentary behaviors in the United States, 2003-2004. *American journal of epidemiology*, *167*(7), 875–881. https://doi.org/10.1093/aje/kwm390

Matthews, C. E., Keadle, S. K., Troiano, R. P., Kahle, L., Koster, A., Brychta, R. J., van Domelen, D., Caserotti, P., Chen, K. Y., Harris, T. B., & Berrigan, D. (2016). Accelerometer-measured dose-response for physical activity, sedentary time, and mortality in US adults. *The American journal of clinical nutrition*, *104*(5), 1424–1432. https://doi.org/10.3945/ajcn.116.135129

Mekary, R. A., Willett, W. C., Hu, F. B., & Ding, E. L. (2009). Isotemporal substitution paradigm for physical activity epidemiology and weight change. *American journal of epidemiology*, *170*(4), 519–527. https://doi.org/10.1093/aje/kwp163.

Menai, M., van Hees, V. T., Elbaz, A., Kivimaki, M., Singh-Manoux, A., & Sabia, S. (2017). Accelerometer assessed moderate-to-vigorous physical activity and successful ageing: results from the Whitehall II study. *Scientific Reports*, *8*, 45772. https://doi.org/10.1038/srep45772

Migueles, J. H., Rowlands, A. V., Huber, F., Sabia, S., & van Hees, V. T. (2019). GGIR: A Research Community–Driven Open Source R Package for Generating Physical Activity and Sleep Outcomes From Multi-Day Raw Accelerometer Data. *Journal for the Measurement of Physical Behaviour*, *2*(3), 188–196. https://doi.org/10.1123/jmpb.2018-0063

Migueles, J. H., Aadland, E., Andersen, L. B., Brønd, J. C., Chastin, S. F. M., Hansen, B. H., Konstabel, K., Kvalheim, O. M., McGregor, D. E., Rowlands, A. V., Sabia, S., van Hees, V. T., Walmsley, R., Ortega, F. B., & External Review Group. (2021). GRANADA consensus on analytical approaches to assess associations with accelerometer-determined physical behaviours (physical activity, sedentary behaviour and sleep) in epidemiological studies. *British journal of sports medicine*. https://doi.org/10.1136/bjsports-2020-103604

Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Delisle Nyström, C., Mora-Gonzalez, J., Löf, M., Labayen, I., Ruiz, J. R., & Ortega, F. B. (2017). Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports medicine*

(Auckland, N.Z.), *47*(9), 1821–1845. https://doi.org/10.1007/s40279-017-0716-0

Migueles, J. H., Cadenas-Sanchez, C., Tudor-Locke, C. E., Löf, M., Esteban-Cornejo, I., Molina-Garcia, P., Mora-Gonzalez, J., Rodriguez-Ayllon, M., Garcia-Marmol, E., Ekelund, U., & Ortega, F. B. (2019). Comparability of published cut-points for the assessment of physical activity: Implications for data harmonization. *Scandinavian Journal of Medicine & Science in Sports*, *29*(4), 566–574. https://doi.org/10.1111/sms.13356

Montoye, A. H. K., Pivarnik, J. M., Mudd, L. M., Biswas, S., & Pfeiffer, K. A. (2016). Validation and Comparison of Accelerometers Worn on the Hip, Thigh, and Wrists for Measuring Physical Activity and Sedentary Behavior. *AIMS public health*, *3*(2), 298–312. https://doi.org/10.3934/publichealth.2016.2.298

Mora-Gonzalez, J., Esteban-Cornejo, I., Cadenas-Sanchez, C., Migueles, J. H., Molina-Garcia, P., Rodriguez-Ayllon, M., Henriksson, P., Pontifex, M. B., Catena, A., & Ortega, F. B. (2019). Physical Fitness, Physical Activity, and the Executive Function in Children with Overweight and Obesity. *The Journal of pediatrics*, *208*, 50–56.e1. https://doi.org/10.1016/j.jpeds.2018.12.028

Neishabouri, A., Nguyen, J., Samuelsson, J., Guthrie, T., Biggs, M., Wyatt, J., Cross, D., Karas, M., Migueles, J. H., Khan, S., & Guo, C. C. (2022). Quantification of acceleration as activity counts in ActiGraph wearable. *Scientific Reports*, *12*(1), 11958. https://doi.org/10.1038/s41598-022-16003-x

Nigg, C. R., Fuchs, R., Gerber, M., Jekauc, D., Koch, T., Krell-Roesch, J., Lippke, S., Mnich, C., Novak, B., Ju, Q., Sattler, M. C., Schmidt, S. C. E., van Poppel, M. N. M., Reimers, A. K., Wagner, P., Woods, C., & Woll, A. (2020). Assessing physical activity through questionnaires – A consensus of best practices and future directions. *Psychology of Sport and Exercise*, *50*, 101715. https://doi.org/10.1016/j.psychsport.2020.101715

Nigg, C. R., & Woll, A. (2020). Best practices and future research directions: Consensus from the 2 International Workshop of the Center for the Assessment of Physical Activity (CAPA). *Psychology of Sport and Exercise*, *50*, 101734. https://doi.org/10.1016/j.psychsport.2020.101734

Pereira, J. R., Sousa-Sá, E., Zhang, Z., Cliff, D. P., & Santos, R. (2020). Concurrent validity of the ActiGraph GT3X+ and activPAL for assessing sedentary behaviour in 2-3-year-old children under free-living conditions. *Journal of Science and Medicine in Sport*, *23*(2), 151–156. https://doi.org/10.1016/j.jsams.2019.08.009

Pfeiffer, K. A., Clevenger, K. A., Kaplan, A., van Camp, C. A., Strath, S. J., & Montoye, A. H. K. (2022). Accessibility and use of novel methods for predicting physical activity and energy expenditure using accelerometry: a scoping review. *Physiological Measurement*, *43*(9). https://doi.org/10.1088/1361-6579/ac89ca

Rajalahti, T., Arneberg, R., Berven, F. S., Myhr, K.-M., Ulvik, R. J., & Kvalheim, O. M. (2009). Biomarker discovery in mass spectral profiles by means of selectivity ratio plot. *Chemometrics and Intelligent Laboratory Systems*, *95*(1), 35–48. https://doi.org/10.1016/j.chemolab.2008.08.004

Rajalahti, T., & Kvalheim, O. M. (2011). Multivariate data analysis in pharmaceutics: a tutorial review. *International journal of pharmaceutics*, *417*(1-2), 280–290. https://doi.org/10.1016/j.ijpharm.2011.02.019

Ramsay, J. O. (2005). *Functional Data Analysis* (2nd ed.). Springer. https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=328488

Reichert, M., Giurgiu, M., Koch, E. D., Wieland, L. M., Lautenbach, S., Neubauer, A. B., von Haaren-Mack, B., Schilling, R., Timm, I., Notthoff, N., Marzi, I., Hill, H., Brüssler, S., Eckert, T., Fiedler, J., Burchartz, A., Anedda, B., Wunsch, K., Gerber, M., ... Liao, Y. (2020). Ambulatory Assessment for Physical Activity Research: State of the Science, Best Practices and Future Directions. *Psychology of Sport and Exercise*, *50*. https://doi.org/10.1016/j.psychsport.2020.101742

Romanzini, M., Petroski, E. L., Ohara, D., Dourado, A. C., & Reichert, F. F. (2014). Calibration of ActiGraph GT3X, Actical and RT3 accelerometers in adolescents. *European journal of sport science*, *14*(1), 91–99. https://doi.org/10.1080/17461391.2012.732614

Rosenberger, M. E., Fulton, J. E., Buman, M. P., Troiano, R. P., Grandner, M. A., Buchner, D. M., & Haskell, W. L. (2019). The 24-Hour Activity Cycle: A New Paradigm for Physical Activity. *Medicine and science in sports and exercise*, *51*(3), 454–464. https://doi.org/10.1249/MSS.0000000000001811

Rosenberger, M. E., Haskell, W. L., Albinali, F., Mota, S., Nawyn, J., & Intille, S. (2013). Estimating activity and sedentary behavior from an accelerometer on the hip or wrist. *Medicine and science in sports and exercise*, *45*(5), 964–975. https://doi.org/10.1249/MSS.0b013e31827f0d9c

Routen, A. C., Upton, D., Edwards, M. G., & Peters, D. M. (2012). Discrepancies in accelerometer-measured physical activity in children due to cut-point non-equivalence and placement site. *Journal of sports sciences*, *30*(12), 1303–1310. https://doi.org/10.1080/02640414.2012.709266

Rowlands, A. V. (2018). Moving Forward With Accelerometer-Assessed Physical Activity: Two Strategies to Ensure Meaningful, Interpretable, and Comparable Measures. *Pediatric Exercise Science*, *30*(4), 450–456. https://doi.org/10.1123/pes.2018-0201

Rowlands, A. V. (2007). Accelerometer Assessment of Physical Activity in Children: An Update. *Pediatric Exercise Science*, *19*(3), 252–266. https://doi.org/10.1123/pes.19.3.252

Rütten, A., & Pfeifer, K. (Eds.). (2017). *Nationale Empfehlungen für Bewegung und Bewegungsförderung* (Aufl. 1.2.06.17, Vol. Sonderheft 3). Bundeszentrale für gesundheitliche Aufklärung (BZgA).

Schmidt, S. C. E., Henn, A., Albrecht, C., & Woll, A. (2017). Physical Activity of German Children and Adolescents 2003-2012: The MoMo-Study. *International journal of environmental research and public health*, *14*(11). https://doi.org/10.3390/ijerph14111375

Schneller, M. B., Bentsen, P., Nielsen, G., Brønd, J. C., Ried-Larsen, M., Mygind, E., & Schipperijn, J. (2017). Measuring Children's Physical Activity: Compliance Using Skin-Taped Accelerometers. *Medicine and science in sports and exercise*, *49*(6), 1261–1269. https://doi.org/10.1249/mss.0000000000001222

Sezgin, E. (2021). Can We Use Commercial Mobile Apps Instead of Research Mobile Apps in Healthcare Research? *Frontiers in public health*, *9*, 685439. https://doi.org/10.3389/fpubh.2021.685439

Sherar, L. B., Griew, P., Esliger, D. W., Cooper, A. R., Ekelund, U., Judge, K., & Riddoch, C. J. (2011). International children's accelerometry database (ICAD): Design and methods. *BMC public health*, *11*, 485. https://doi.org/10.1186/1471-2458-11-485

Shi, Y., Huang, W. Y., Yu, J. J., Sheridan, S., Sit, C. H.-P., & Wong, S. H.-S. (2019). Compliance and Practical Utility of Continuous Wearing of activPAL™ in Adolescents. *Pediatric Exercise Science*, *31*(3), 363–369. https://doi.org/10.1123/pes.2018-0170

Siaplaouras, J., Niessner, C., Helm, P. C., Jahn, A., Flemming, M., Urschitz, M. S., Sticker, E., Abdul-Khaliq, H., Bauer, U. M., & Apitz, C. (2020). Physical Activity Among Children With Congenital Heart Defects in Germany: A Nationwide Survey. *Frontiers in pediatrics*, *8*, 170. https://doi.org/10.3389/fped.2020.00170

Smith, C., Galland, B. C., Taylor, R. W., & Meredith-Jones, K. A. (2019). ActiGraph GT3X+ and Actical Wrist and Hip Worn Accelerometers for Sleep and Wake Indices in Young Children Using an Automated Algorithm: Validation With Polysomnography. *Frontiers in psychiatry*, *10*, 958. https://doi.org/10.3389/fpsyt.2019.00958

Smits, E. J., Salomoni, S., Costa, N., Rodríguez-Romero, B., & Hodges, P. W. (2022). How reliable is measurement of posture during sleep: real-world measurement of body posture and movement during sleep using accelerometers. *Physiological measurement*, *43*(1). https://doi.org/10.1088/1361-6579/ac482f

Sørensen, H., Goldsmith, J., & Sangalli, L. M. (2013). An introduction with medical applications to functional data analysis. *Statistics in medicine*, *32*(30), 5222–5240. https://doi.org/10.1002/sim.5989

Stamatakis, E., Gale, J., Bauman, A., Ekelund, U., Hamer, M., & Ding, D. (2019). Sitting Time, Physical Activity, and Risk of Mortality in Adults. *Journal of the American College of Cardiology*, *73*(16), 2062–2072. https://doi.org/10.1016/j.jacc.2019.02.031

Steene-Johannessen, J., Hansen, B. H., Dalene, K. E., Kolle, E., Northstone, K., Møller, N. C., Grøntved, A., Wedderkopp, N., Kriemler, S., Page, A. S., Puder, J. J., Reilly, J. J., Sardinha, L. B., van Sluijs, E. M. F., Andersen, L. B., van der Ploeg, H. P., Ahrens, W., Flexeder, C., Standl, M., . . . Ekelund, U. (2020). Variations in accelerometry measured physical activity and sedentary time across Europe - harmonized analyses of 47,497 children and adolescents. *International Journal of Behavioral Nutrition and Physical Activity*, *17*(1), 38. https://doi.org/10.1186/s12966-020-00930-x

Straczkiewicz, M., James, P., & Onnela, J.-P. (2021). A systematic review of smartphone-based human activity recognition methods for health research. *NPJ digital medicine*, *4*(1), 148. https://doi.org/10.1038/s41746-021-00514-4

Swartz, A. M., Strath, S. J., Bassett, D. R., O'Brien, W. L., King, G. A., & Ainsworth, B. E. (2000). Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Medicine and science in sports and exercise*, *32*(9 Suppl), S450–6. https://doi.org/10.1097/00005768-200009001-00003

Tost, H., Reichert, M., Braun, U., Reinhard, I., Peters, R., Lautenbach, S., Hoell, A., Schwarz, E., Ebner-Priemer, U., Zipf, A., & Meyer-Lindenberg, A. (2019). Neural correlates of individual differences in affective benefit of real-life urban green space exposure. *Nature neuroscience*, *22*(9), 1389–1393. https://doi.org/10.1038/s41593-019-0451-y

Tremblay, M. S., Aubert, S., Barnes, J. D., Saunders, T. J., Carson, V., Latimer-Cheung, A. E., Chastin, S. F. M., Altenburg, T. M., & Chinapaw, M. J. M. (2017). Sedentary Behavior Research Network (SBRN) - Terminology Consensus Project process and outcome. *The international journal of behavioral nutrition and physical activity*, *14*(1), 75. https://doi.org/10.1186/s12966-017-0525-8

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & McDowell, M. (2008). Physical activity in the United States measured by accelerometer.

Medicine and science in sports and exercise, *40*(1), 181–188. https://doi.org/10.1249/mss.0b013e31815a51b3

Troiano, R. P., McClain, J. J., Brychta, R. J., & Chen, K. Y. (2014). Evolution of accelerometer methods for physical activity research. *British journal of sports medicine*, *48*(13), 1019–1023. https://doi.org/10.1136/bjsports-2014-093546

Troiano, R. P., Pettee Gabriel, K. K., Welk, G. J., Owen, N., & Sternfeld, B. (2012). Reported physical activity and sedentary behavior: why do you ask? *Journal of physical activity & health*, *9 Suppl 1*, S68–75. https://doi.org/10.1123/jpah.9.s1.s68

Troiano, R. P., Stamatakis, E., & Bull, F. C. (2020). How can global physical activity surveillance adapt to evolving physical activity guidelines? Needs, challenges and future directions. *British journal of sports medicine*, *54*(24), 1468–1473. https://doi.org/10.1136/bjsports-2020-102621

Trost, S. G. (2007). State of the Art Reviews: Measurement of Physical Activity in Children and Adolescents. *American Journal of Lifestyle Medicine*, *1*(4), 299–314. https://doi.org/10.1177/1559827607301686

Tudor-Locke, C. E., Barreira, T. V., & Schuna, J. M. (2015). Comparison of step outputs for waist and wrist accelerometer attachment sites. *Medicine and science in sports and exercise*, *47*(4), 839–842. https://doi.org/10.1249/MSS.0000000000000476

U.S. Department of Health and Human Services. (2018). Physical Activity Guidelines for Americans, 2nd edition. https://health.gov/paguidelines/second-edition/pdf/Physical_Activity_Guidelines_2nd_edition.pdf

van Hees, V. T., Fang, Z., Mirkes, E. M., Heywood, J., Zhao, J. H., Joan, C. P., Sabia, S., & Migueles, J. H. (2019). GGIR. https://doi.org/10.5281/ZENODO.1051064

Welch, W. A., Bassett, D. R., Thompson, D. L., Freedson, P. S., Staudenmayer, J. W., John, D., Steeves, J. A., Conger, S. A., Ceaser, T., Howe, C. A., Sasaki, J. E., & Fitzhugh, E. C. (2013). Classification accuracy of the wrist-worn gravity estimator of normal everyday activity accelerometer. *Medicine and science in sports and exercise*, *45*(10), 2012–2019. https://doi.org/10.1249/MSS.0b013e3182965249

Woll, A. (1996). *Gesundheitsförderung in der Gemeinde : eine empirische Untersuchung zum Zusammenhang von sportlicher Aktivität, Fitneß und Gesundheit bei Personen im mittleren und späteren Erwachsenenalter* (Dissertation). Johann Wolfgang Goethe Universität Frankfurt am Main. Neu-Isenburg, Institut für Sportwissenschaften.

Woll, A., Albrecht, C., & Worth, A. (2017). Motorik-Module (MoMo) – the KiGGS wave 2 module to survey motor performance and physical activity. *Journal of Health Monitoring*, *2*(S3), 66–73. https://doi.org/10.17886/RKI-GBE-2017-110

World Health Organization. (2010). *Global recommendations on physical activity for health*.

Wunsch, K., Nigg, C. R., Weyland, S., Jekauc, D., Niessner, C., Burchartz, A., Schmidt, S., Meyrose, A.-K., Manz, K., Baumgarten, F., & Woll, A. (2021). The relationship of self-reported and device-based measures of physical activity and health-related quality of life in adolescents. *Health and Quality of Life Outcomes*, *19*(1), 67. https://doi.org/10.1186/s12955-021-01682-3

A. Appendix

Here you find documents and notes referred to in the previous chapters.

A.1 Comparison WHO Guidelines 2010 vs. 2020

The article in Chapter 5 determined whether the subject was active for at least 60 min or not on each day individually. To look at the exact times spent with MVPA on every single day will result in fewer days of at least 60 min MVPA when both evaluation methods are compared (Colley et al., 2017). The main reason, however, since the study was already planned and started in 2014, the questions in the questionnaire still referred to the 2010 WHO Guidelines (World Health Organization, 2010). In 2020 the recommendations on youth activity have changed from a recommendation of at least 60 min per day to a recommendation of an average of 60 min per day (Chaput et al., 2020). We have looked at the accelerometer data with the background of the new WHO guideline as well. It should be noted, however, that these results cannot be compared with the results of the question used in this study about the number of days on which the subjects have MVPA for more than 60 min. However, if one compares the number of subjects meeting the old versus the new guideline based on the accelerometer data in Figure A.1, we see that the percentage increases from 3% to a full 34% of the study participants. This means that 31% of the participants who do not reach 60 min MVPA on all days still have days in the week on which they do so much physical activity that these outweigh the remaining days under the new guideline compared to the old one. This drastically reduces the proportion of children and adolescents who are too inactive, which is also likely to cause some political controversy in the future.

Figure A.1: Comparison of participants reaching the WHO Guidelines of 2010 (60 min MVPA every day) and 2020 (60 min MVPA per day on average)

WHO 2010 x 2020

		ACC Days - WHO 2020								Overall
	days	0	1	2	3	4	5	6	7	n
ACC Days - WHO 2010	0	1	38	129	217	147	35	0	0	567
	1	0	0	9	67	170	147	43	2	438
	2	0	0	0	2	48	129	134	38	351
	3	0	0	0	0	2	32	93	155	282
	4	0	0	0	0	0	1	24	210	235
	5	0	0	0	0	0	0	0	180	180
	6	0	0	0	0	0	0	0	114	114
	7	0	0	0	0	0	0	0	69	**69**
Overall	n	1	38	138	286	367	344	294	**768**	**2236**

WHO 2010 x 2020 (%)

		ACC Days - WHO 2020								Overall
	days	0	1	2	3	4	5	6	7	n %
ACC Days - WHO 2010	0	0%	2%	6%	10%	7%	2%	0%	0%	25%
	1	0%	0%	0%	3%	8%	7%	2%	0%	20%
	2	0%	0%	0%	0%	2%	6%	6%	2%	16%
	3	0%	0%	0%	0%	0%	1%	4%	7%	13%
	4	0%	0%	0%	0%	0%	0%	1%	9%	11%
	5	0%	0%	0%	0%	0%	0%	0%	8%	8%
	6	0%	0%	0%	0%	0%	0%	0%	5%	5%
	7	0%	0%	0%	0%	0%	0%	0%	3%	**3%**
Overall	n %	0%	2%	6%	13%	16%	15%	13%	**34%**	100%

A.2 Supplement Material Chapter 4

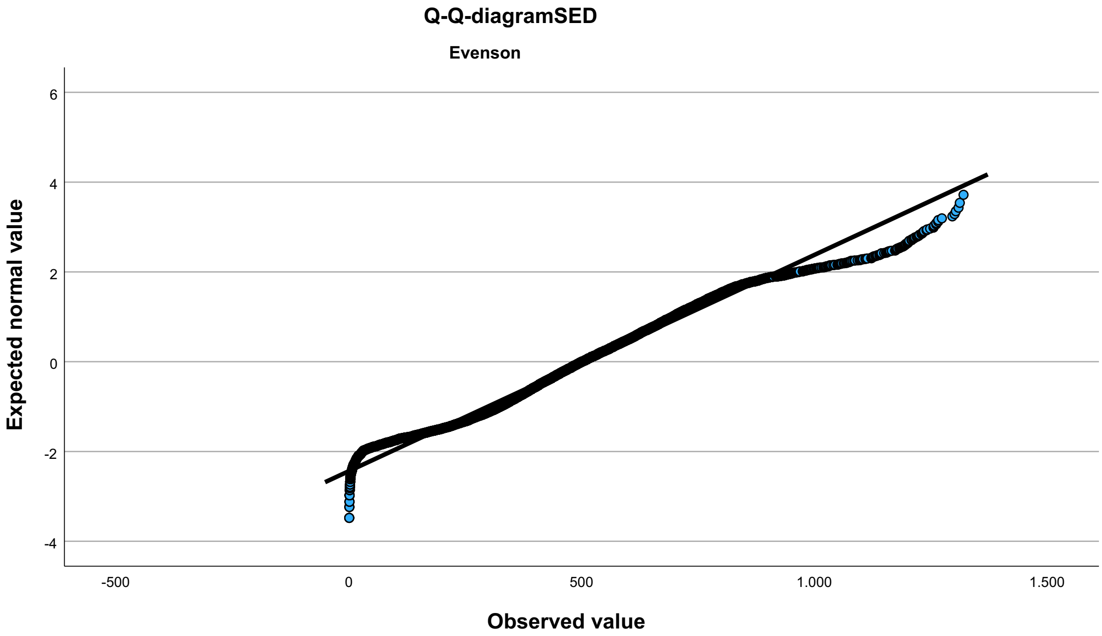
The article in Chapter 4 has some Supplement Material and can be found here.

Table A.1: Histogram of the wear time for each day per hour separated between the two non-wearing algorithms

	Choi et al. 2011		Troiano et al. 2008	
h	**N**	**%**	**N**	**%**
0	6196	5.9%	3066	2.9%
1	2654	2.5%	5798	5.5%
2	1050	1.0%	1066	1.0%
3	830	0.8%	892	0.8%
4	920	0.9%	1048	1.0%
5	1030	1.0%	988	0.9%
6	1302	1.2%	1436	1.4%
7	1208	1.1%	1340	1.3%
8	1910	1.8%	1942	1.8%
9	2470	2.3%	2534	2.4%
10	3910	3.7%	4036	3.8%
11	5520	5.2%	6182	5.9%
12	8970	8.5%	9494	9.0%
13	13370	12.7%	13816	13.1%
14	17090	16.2%	17282	16.4%
15	14460	13.7%	14108	13.4%
16	9940	9.4%	9406	8.9%
17	5620	5.3%	5238	5.0%
18	2780	2.6%	2230	2.1%
19	940	0.9%	828	0.8%
20	350	0.3%	542	0.5%
21	470	0.4%	508	0.5%
22	370	0.4%	504	0.5%
23	510	0.5%	680	0.6%
24	1700	1.6%	606	0.6%

Note: Note: *h* wear time hours per day, *N* Number of datasets, *%* percentage of datasets

Figure A.2: Q-Q Plots for SED with Evenson and Romanzini Algorithms



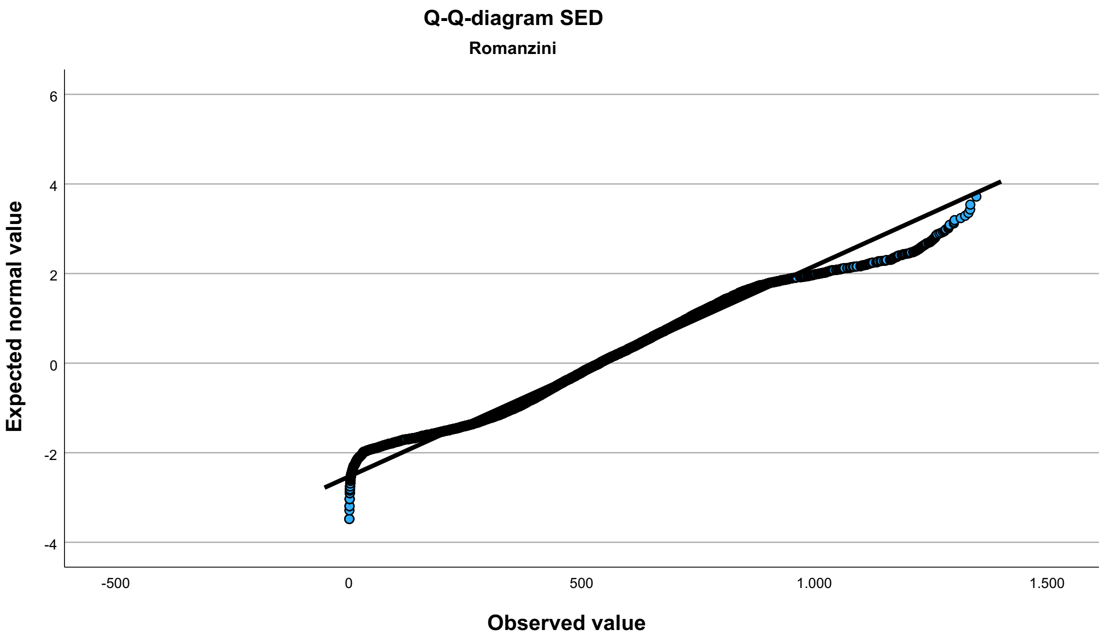
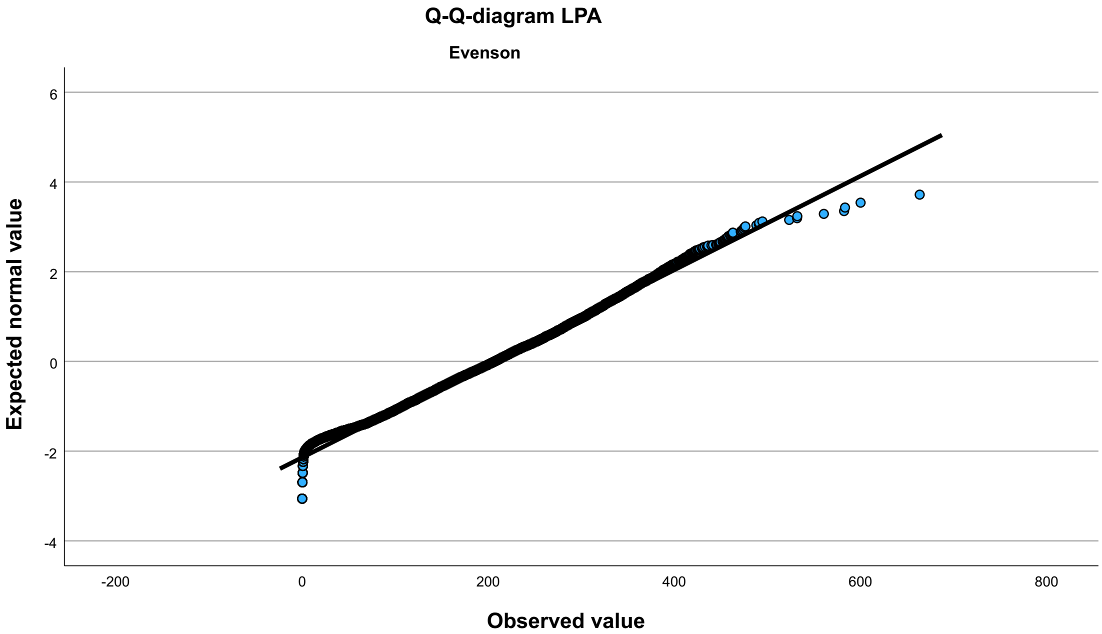


Figure A.3: Q-Q Plots for LPA with Evenson and Romanzini Algorithms



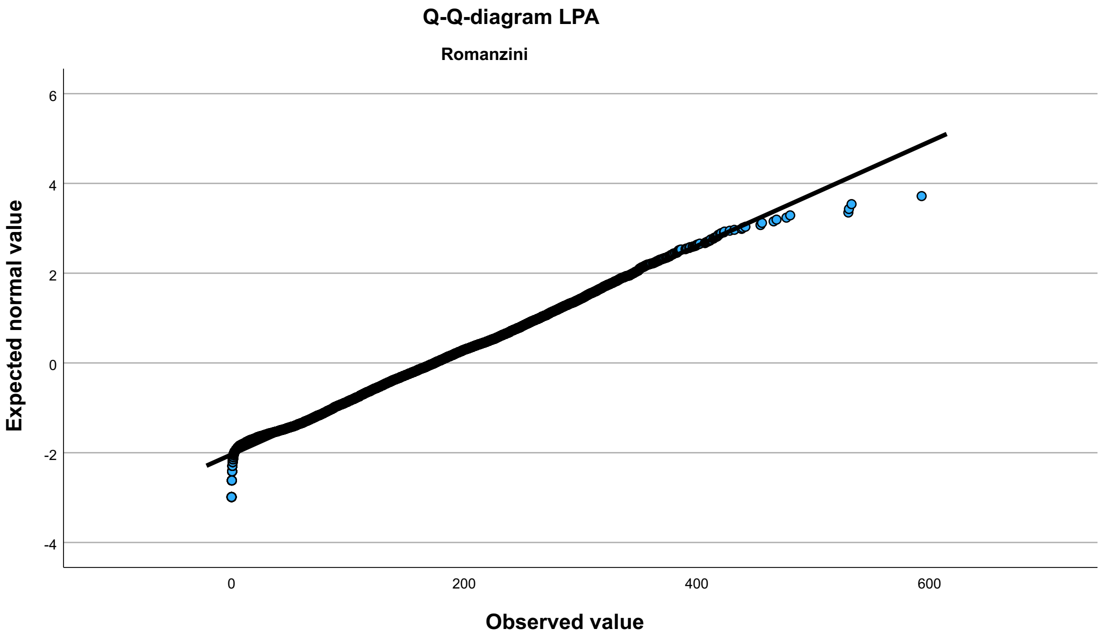
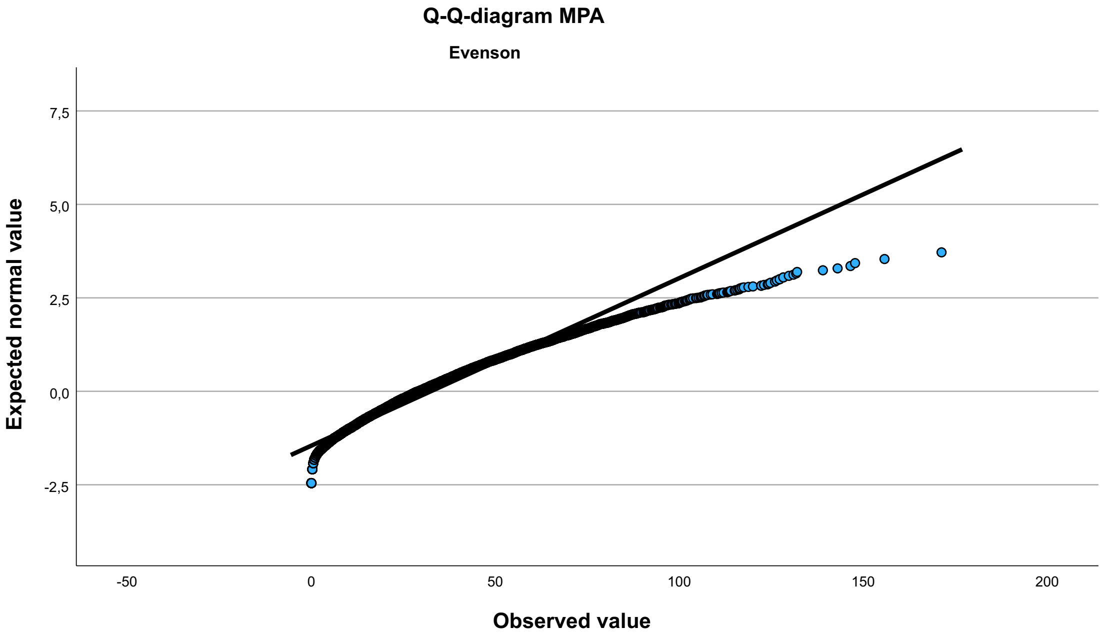


Figure A.4: Q-Q Plots for MPA with Evenson and Romanzini Algorithms



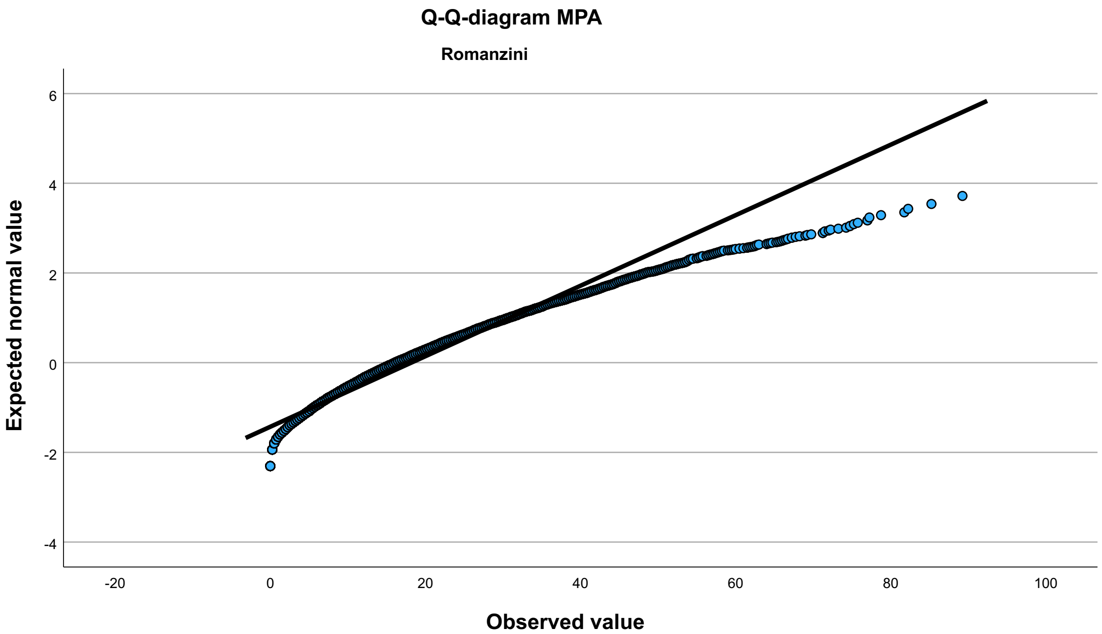
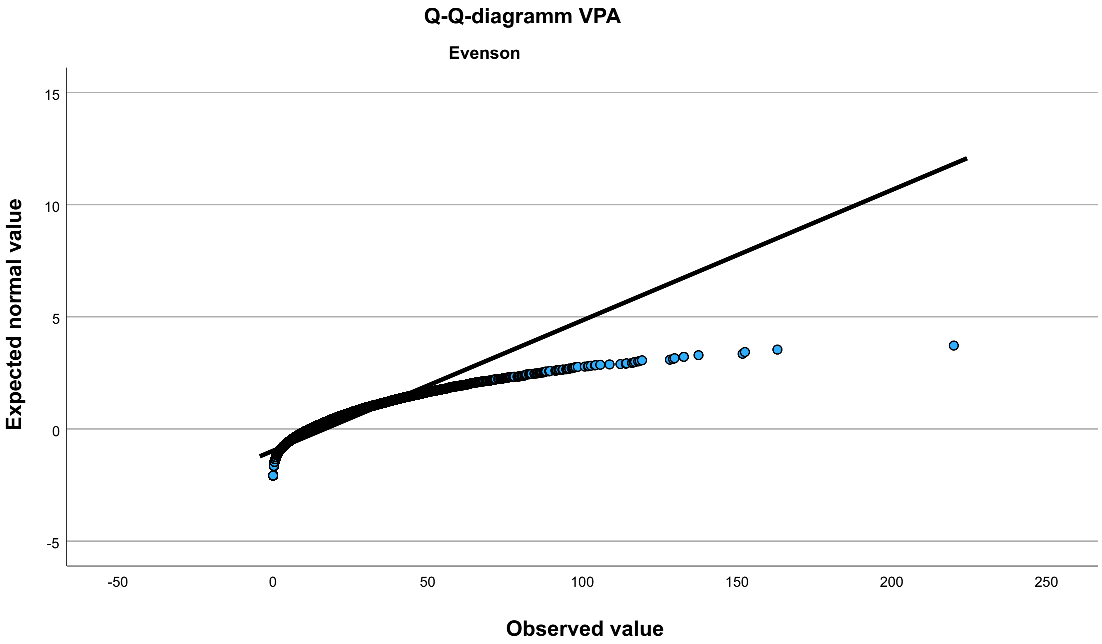
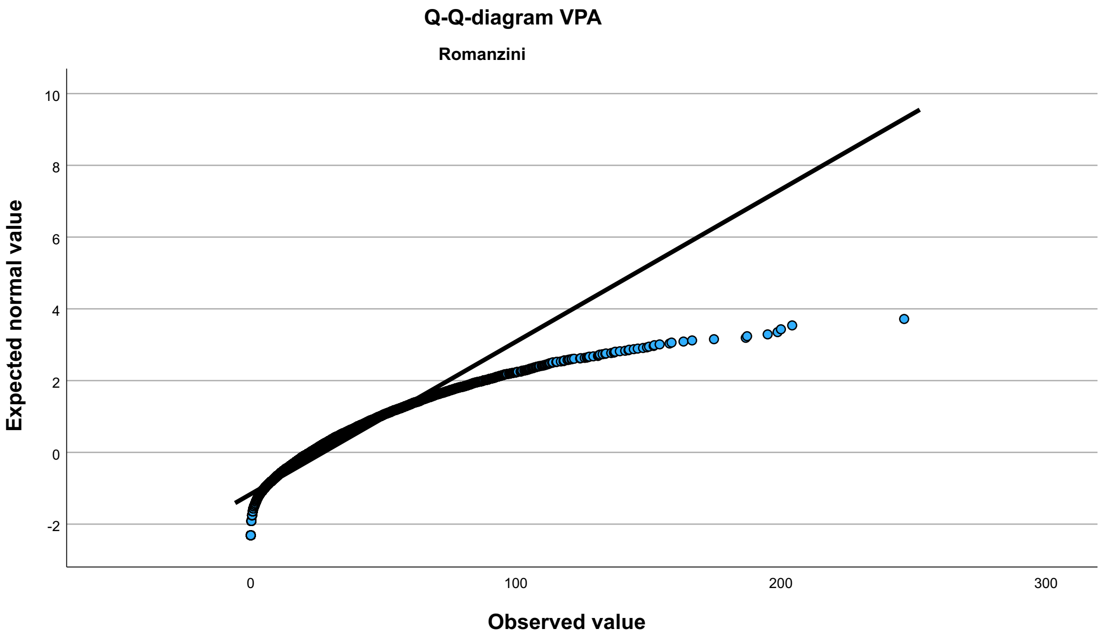


Figure A.5: Q-Q Plots for VPA with Evenson and Romanzini Algorithms





A.3 Poster Presentations

Burchartz, A., Anedda B., Oriwol, D., Niessner, C., & Woll, A. (2018). Impact of different evaluation methods on the dropout rate of objectively detected physical activity on age specific subgroups: Results from the Motorik-Modul Study (MoMo). In N. Anoke, ... K. Wijndaele (Eds.), *7th International Society for Physical Activity and Health Congress (p. 216). Human Kinetics.*, see Figure A.6

Burchartz, A., Anedda, B., Oriwol, D., Kolb, S., Schmidt, S. C. E., & Woll, A. (2019). Erreichen Kinder und Jugendliche in Deutschland die WHO-Richtlinien für körperliche Aktivität: - erhoben durch Akzelerometer und Fragebogen. In A. Arampatzis, S. Braun, K. Schmitt, & B. Wolfarth (Eds.), *Schriften der Deutschen Vereinigung für Sportwissenschaft: Vol. 282. Sport im öffentlichen Raum: 24 dvs-Hochschultag • Berlin 18 – 20. September 2019 • Abstracts (1st ed., p. 324)*. Feldhaus. Won first place in the natural sciences category, see Figure A.7

Figure A.6: Poster Presentation at the 7th International Society for Physical Activity and Health Congress (ISPAH) 2018 in London

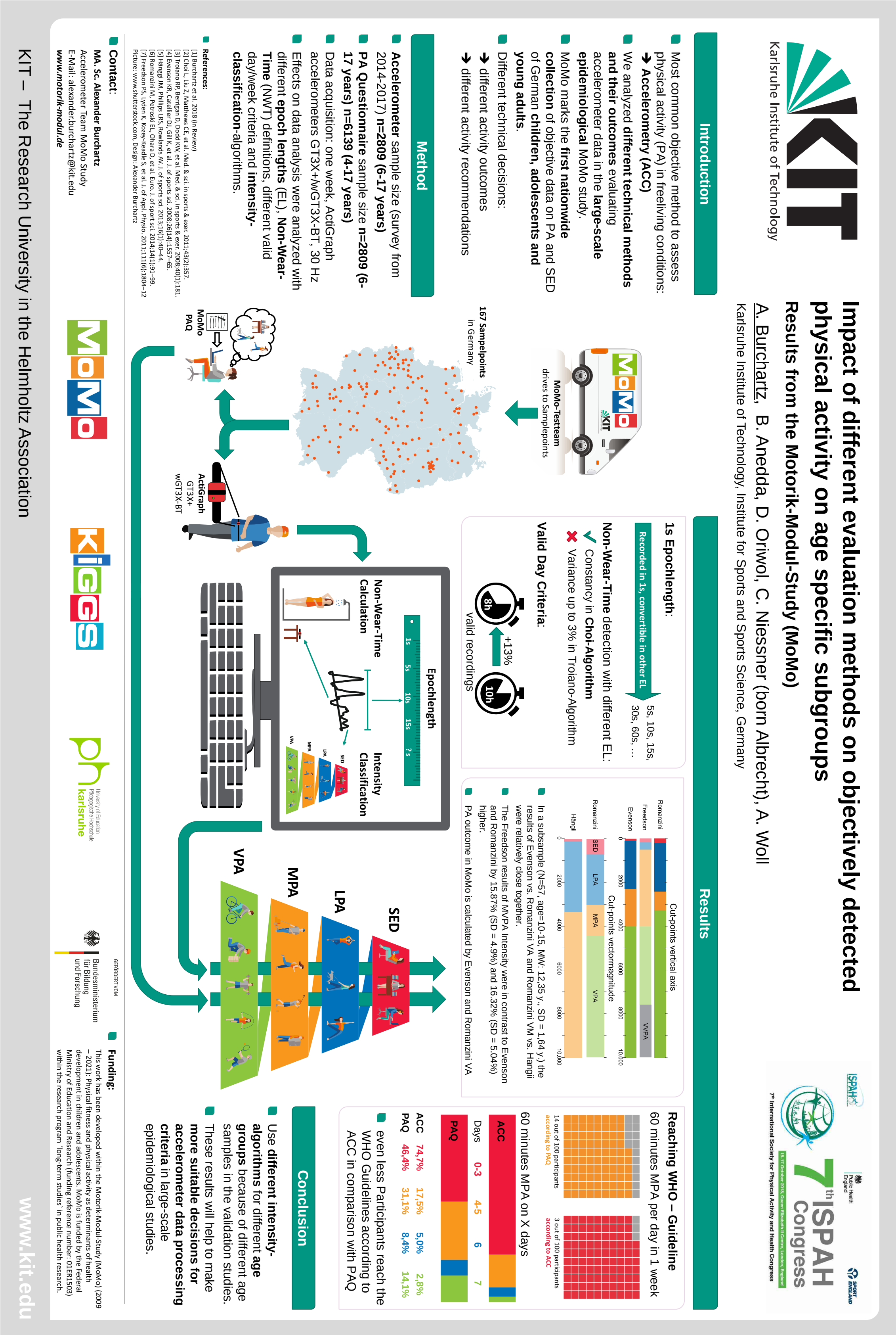
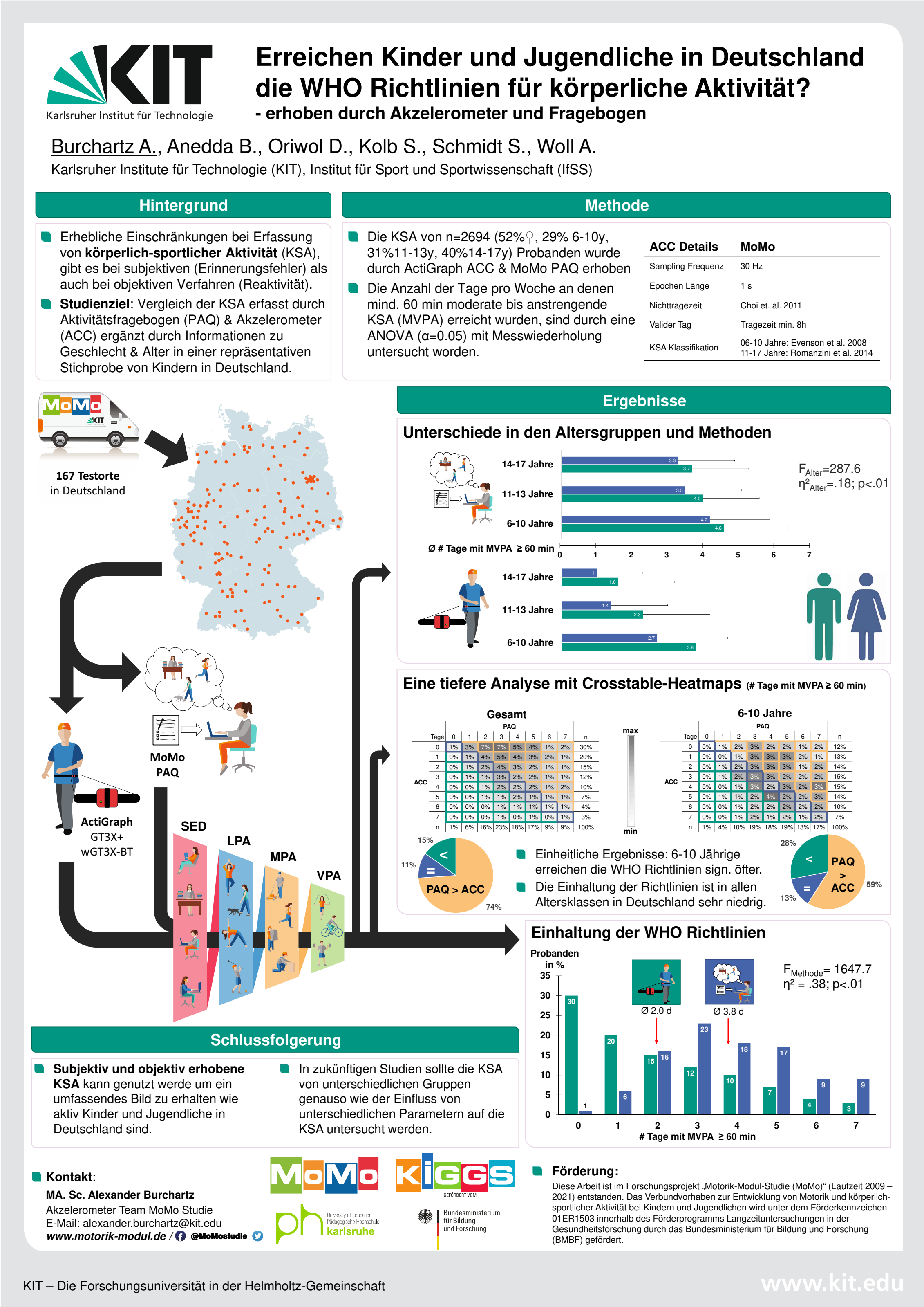


Figure A.7: Poster Presentation at dvs-Hochschultag 2019 in Berlin. Won first place in the natural sciences category.



B. Publications

Research Articles as Author or Co-Author

33	2023	**Burchartz, A.**; Kolb, S.; Klos, Leon.; Schmidt, S. CE; Haaren-Mack, B. von; Niessner, C.; & Woll, A. (2023): How specific combinations of epoch length, non-wear time and cut-points influence physical activity - Processing accelerometer data from children and adolescents in the nationwide MoMo study. *German Journal of Exercise and Sport Research.*Advance online publication. DOI: 10.1007/s12662-023-00892-9
32	2023	von Haaren-Mack, B., **Burchartz, A.**, Opper, E., & Woll, A. (in Review). Diagnostik zur Erfassung körperlichen Verhaltens: Ansätze und Anwendung. In A. Thiel, S. Tittlbach, G. Sudeck, P. Wagner, & A. Woll (Eds.), *Handbuch Bewegungsbezogene Gesundheitsförderung: Beiträge zur Lehre und Forschung im Sport.* Hofmann-Verlag.
31	2023	Klos, L., **Burchartz, A.**, Niessner, C., Reimers, A. K., Thron, M., Woll, A., & Wäsche, H. (2023). Active school transport routines during school transitions: Socio-structural predictors of changes from childhood into early adulthood Corresponding. *Health and Place.* 81,103005. DOI: 10.1016/j. healthplace.2023.103005
30	2023	Nigg, C., Fiedler, J., Burchartz, A., Reichert, M., Niessner, C., Woll, A., & Schipperijn, J. (in Review). Associations between green space and youth's physical activity in urban and rural areas: results of a German cohort study. *Landscape and Urban Planning.*
29	2023	Manz, K., **Burchartz, A.**, Niessner, C., Kolb, S., Schienkiewitz, A., & Mensink, G. B. M. (accepted). The Challenge of Incomplete Data in Accelerometer Studies: Characteristics of Nonparticipation and Noncompliance in a Nationwide Sample of Adolescents and Young Adults in Germany. *Journal of Physcial Activity & Health.* 1–13. DOI: 10.1123/jpah.2022-0443

28 2022 Reichert, M., Nigg, C., Brüßler, S., **Burchartz, A.**, Jekauc, D., Limberger, M., Krell-Rösch, J., von Haaren-Mack, B., Ebner-Priemer, U., Niesser, C., Schipperijn, J., & Woll, A. (in Review). Urban living can benefit physical activity: Germany's youth city dwellers engage in more moderate to vigorous physical activity compared to their rural counterparts. *Environment International*.

27 2022 Fritsch, J., Weyland, S., Feil, K., **Burchartz, A.**, Schmidt, S. C. E., Woll, A., Strauch, U., Wienke, B., & Jekauc, D. (2022). A study on the psychometric properties of the short version of the Physical Activity Enjoyment Scale in an adult population. *International Journal of Environmental Research and Public Health.*, 19(22), 15294. 10.3390/ijerph192215294

26 2022 Kolb, S., **Burchartz, A.**, Krause, L., Klos, L., Schmidt, S. C. E., Woll, A., & Niessner, C. (2022). Physical Activity and Recurrent Pain in Children and Adolescents in Germany: Results from the MoMo-Study. *Children* (Basel, Switzerland), 9(11), 1645. 10.3390/children9111645

25 2022 **Burchartz, A.**, Oriwol, D., Kolb, S., Schmidt, S. C. E., Haaren-Mack, B. von, Niessner, C., & Woll, A. (2022). Impact of weekdays versus weekend days on accelerometer measured physical behavior among children and adolescents: Results from the MoMo-Study. *German Journal of Exercise and Sport Research.* Advance online publication. 10.1007/s12662-022-00811-4

24 2022 Giurgiu, M., Kolb, S., Nigg, C., **Burchartz, A.**, Timm, I., Becker, M., Rulf, E., Doster, A.-K., Koch, E., Bussmann, J. B. J., Nigg, C. R., Ebner-Priemer, U. W., & Woll, A. (2022). Assessment of 24-hour physical behaviour in children and adolescents via wearables: A systematic review of free-living validation studies. *BMJ Open Sport & Exercise Medicine, 8*(2), e001267. 10.1136/bmjsem-2021-001267

23 2022 Götte, M., Basteck, S., Beller, R., Gauß, G., Schmidt, S. C. E., **Burchartz, A.**, Kolb, S., Grydeland, M., & Reinhardt, D. (2022). Physical activity in 9–15 year-old pediatric cancer survivors compared to a nationwide sample. *Journal of Cancer Research and Clinical Oncology.* Advance online publication. 10.1007/s00432-022-04392-5

22 2022 Nigg, C., **Burchartz, A.**, Reichert, M., Woll, A., & Niessner, C. (2022). Children and adolescents do not compensate for physical activity but do compensate for sedentary behavior. *German Journal of Exercise and Sport Research.* Advance online publication. 10.1007/s12662-022-00808-z

21 2022 Nigg, C., Niessner, C., **Burchartz, A.**, Woll, A., & Schipperijn, J. (2022). The geospatial and conceptual configuration of the natural environment impacts the association with health outcomes and behavior in children and adolescents. *International Journal of Health Geographics*, 21(1). 10.1186/s12942-022-00309-0

20 2021 **Burchartz, A.**, Oriwol, D., Kolb, S., Schmidt, S. C. E., Wunsch, K., Manz, K., Niessner, C., & Woll, A. (2021). Comparison of self-reported & device-based, measured physical activity among children in Germany. *BMC Public Health, 21*(1). 10.1186/s12889-021-11114-y

19	2021	Chen, C., Weyland, S., Fritsch, J., Woll, A., Niessner, C., **Burchartz, A.**, Schmidt, S. C. E., & Jekauc, D. (2021). A Short Version of the Physical Activity Enjoyment Scale: Development and Psychometric Properties. *International Journal of Environmental Research and Public Health, 18*(21). 10.3390/ijerph182111035
18	2021	Fiedler, J., Eckert, T., **Burchartz, A.**, Woll, A., & Wunsch, K. (2021). Comparison of Self-Reported and Device-Based Measured Physical Activity Using Measures of Stability, Reliability, and Validity in Adults and Children. *Sensors, 21*(8). 10.3390/s21082672
17	2021	Kolb, S., **Burchartz, A.**, Oriwol, D., Schmidt, S. C. E., Woll, A., & Niessner, C. (2021). Indicators to Assess Physical Health of Children and Adolescents in Activity Research—A Scoping Review. *International Journal of Environmental Research and Public Health, 18*(20), 10711. 10.3390/ijerph182010711
16	2021	Migueles, J. H., Aadland, E., Andersen, L. B., Brønd, J. C., Chastin, S. F., Hansen, B. H., Konstabel, K., Kvalheim, O. M., McGregor, D. E., Rowlands, A. V., Sabia, S., van Hees, V. T., Walmsley, R., Ortega, F. B., & **External Review Group** (2021). Granada consensus on analytical approaches to assess associations with accelerometer-determined physical behaviours (physical activity, sedentary behaviour and sleep) in epidemiological studies. *British Journal of Sports Medicine.* Advance online publication. 10.1136/bjsports-2020-103604
15	2021	Nigg, C., Oriwol, D., Wunsch, K., **Burchartz, A.**, Kolb, S., Worth, A., Woll, A., & Niessner, C. (2021). Population density predicts youth's physical activity changes during Covid-19 – Results from the MoMo-Study. *Health & Place, 70*, 102619. 10.1016/j. healthplace.2021.102619
14	2021	Nigg, C. R., Wunsch, K., **Burchartz, A.**, & Woll, A. (2021, February 22). Improving the Future of Physical Activity Self-reports: Commentary on „Physical activity self-reports: past or future?". *British Journal of Sports Medicine.* bjsports-2020-103595.responses
13	2021	Schmidt, S. C. E., Anedda, B., **Burchartz, A.**, Eichsteller, A., Kolb, S., Nigg, C., Niessner, C., Oriwol, D., Worth, A., & Woll, A. (2021). Author Correction: Physical activity and screen time of children and adolescents before and during the COVID-19 lockdown in Germany: A natural experiment. *Scientific Reports, 11*(1), 24329. 10.1038/s41598-021-03905-5
12	2021	Schmidt, S. C. E., **Burchartz, A.**, Kolb, S., Niessner, C., Oriwol, D., Hanssen-Doose, A., Worth, A., & Woll, A. Zur Situation der körperlich-sportlichen Aktivität von Kindern und Jugendlichen während der COVID-19 Pandemie in Deutschland : Die Motorik-Modul Studie (MoMo). *KIT Scientific Working Papers.* Advance online publication. 10.5445/IR/1000133697

11 2021 Schmidt, S. C. E., **Burchartz, A.**, Kolb, S., Niessner, C., Oriwol, D., & Woll, A. (2021). Influence of socioeconomic variables on physical activity and screen time of children and adolescents during the COVID-19 lockdown in Germany: the MoMo study. *German Journal of Exercise and Sport Research.* Advance online publication. 10.1007/s12662-021-00783-x

10 2021 Woll, A., Klos, L., **Burchartz, A.**, Hanssen-Doose, A., Niessner, C., Oriwol, D., Schmidt, S. C. E., Bös, K., & Worth, A. (2021). Cohort Profile Update: The Motorik-Modul (MoMo) Longitudinal Study-physical fitness and physical activity as determinants of health development in German children and adolescents. *International Journal of Epidemiology.* Advance online publication. 10.1093/ije/dyaa281

9 2021 Wunsch, K., Nigg, C., Niessner, C., Schmidt, S. C. E., Oriwol, D., Hanssen-Doose, A., **Burchartz, A.**, Eichsteller, A., Kolb, S., Worth, A., & Woll, A. (2021). Reply to Kersting et al. Comment on "Wunsch et al. The Impact of COVID-19 on the Interrelation of Physical Activity, Screen Time and Health-Related Quality of Life in Children and Adolescents in Germany: Results of the Motorik-Modul Study. Children 2021, 8, 98". *Children (Basel, Switzerland), 8*(7). 10.3390/children8070533

8 2021 Wunsch, K., Nigg, C., Niessner, C., Schmidt, S. C. E., Oriwol, D., Hanssen-Doose, A., **Burchartz, A.**, Eichsteller, A., Kolb, S., Worth, A., & Woll, A. (2021). The impact of COVID-19 on the interrelation of physical activity, screen time and health-related quality of life in children and adolescents in Germany: Results of the Motorik-Modul Study. *Children, 8*(2), 98. 10.3390/children8020098

7 2021 Wunsch, K., Nigg, C. R., Weyland, S., Jekauc, D., Niessner, C., **Burchartz, A.**, Schmidt, S. C. E., Meyrose, A.-K., Manz, K., Baumgarten, F., & Woll, A. (2021). The relationship of self-reported and device-based measures of physical activity and health-related quality of life in adolescents. *Health and Quality of Life Outcomes, 19*(1), 67. 10.1186/s12955-021-01682-3

6 2020 Schmidt, S. C. E., Anedda, B., **Burchartz, A.**, Eichsteller, A., Kolb, S., Nigg, C., Niessner, C., Oriwol, D., Worth, A., & Woll, A. (2020). Physical activity and screen time of children and adolescents before and during the COVID-19 lockdown in Germany: a natural experiment. *Scientific Reports, 10*(1), 21780. 10.1038/s41598-020-78438-4

5 2020 Reichert, M., Giurgiu, M., Koch, E., Wieland, L. M., Lautenbach, S., Neubauer, A. B., Haaren-Mack, B. von, Schilling, R., Timm, I., Notthoff, N., Marzi, I., Hill, H., Brüßler, S., Eckert, T., Fiedler, J., **Burchartz, A.**, Anedda, B., Wunsch, K., Gerber, M., . . . Liao, Y. (2020). Ambulatory Assessment for Physical Activity Research: State of the Science, Best Practices and Future Directions. *Psychology of Sport and Exercise, 50.* 10.1016/j.psychsport.2020.101742

4	2020	**Burchartz, A.**, Anedda, B., Auerswald, T., Giurgiu, M., Hill, H., Ketelhut, S., Kolb, S., Mall, C., Manz, K., Nigg, C. R., Reichert, M., Sprengeler, O., Wunsch, K., & Matthews, C. E. (2020). Assessing physical behavior through accelerometry – State of the science, best practices and future directions. *Psychology of Sport and Exercise, 49*, 101703. 10.1016/j.psychsport.2020.101703
3	2020	**Burchartz, A.**, Manz, K., Anedda, B., Niessner, C., Oriwol, D., Schmidt, S. C. E., & Woll, A. (2020). Measurement of Physical Activity and Sedentary Behavior by Accelerometry Among a Nationwide Sample from the KiGGS and MoMo Study: Study Protocol. *JMIR Research Protocols, 9*(7), e14370. 10.2196/14370
2	2020	Schmidt, S. C. E., Anedda, B., **Burchartz, A.**, Oriwol, D., Kolb, S., Wäsche, H., Niessner, C., & Woll, A. (2020). The physical activity of children and adolescents in Germany 2003-2017: The MoMo-study. *PloS One, 15*(7), e0236117. 10.1371/journal.pone.0236117
1	2019	Woll, A., Oriwol, D., Anedda, B., **Burchartz, A.**, Hanssen-Doose, A., Kopp, M., Niessner, C., Schmidt, S. C. E., Bös, K., & Worth, A. (2019). Körperliche Aktivität, motorische Leistungsfähigkeit und Gesundheit in Deutschland: Ergebnisse aus der Motorik-Modul-Längsschnittstudie (MoMo). *KIT Scientific Working Papers.* Advance online publication. 10.5445/IR/1000095369