Mobile Health interventions to enhance physical activity.

Overview, methodological considerations, and just-in-time adaptive interventions.

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

Janis Fiedler

KIT-Dekan: Prof. Dr. Michael Schefczyk

1. Gutachter: Prof. Dr. Alexander Woll

2. Gutachter: Prof. Dr. Ulrich Ebner-Priemer

Tag der mündlichen Prüfung: 13.10.2022

Acknowledgments

I thank Alexander Woll and Kathrin Wunsch for their outstanding guidance throughout my doctoral thesis. I am grateful for all the things I was able to learn during your supervision and our work together.

Furthermore, I would like to thank my colleagues at the Institute of Sports and Sport Science of the Karlsruhe Institute of Technology for the fruitful discussions and their valuable assistance in a variety of matters. Tobias Eckert deserves a special notion for the close and constructive collaboration within the SMART*FAMILY* study.

I would also like to thank all coauthors for their contributions, helpful discussions, and excellent feedback regarding our publications. Here, I want to thank Caroline Seiferth in particular for the countless hours with we spent delving into data processing and statistical approaches for multilevel modeling in R.

Moreover, I am grateful for the immense help from our research assistants and students regarding data assessments and processing. Without you, this project would have taken forever.

Finally, this endeavor would not have been possible without my family, friends, and especially my beloved partner Marleen. Thank you for your unconditional love and support.

Summary

Physical activity has far-reaching health benefits and contributes to the prevention of noncommunicable diseases like cardiovascular disease, cancer, and diabetes. Today's level of physical activity; however, is below the recommendations of e.g. the World Health Organization for all age groups. This amount of physical inactivity (i.e. not meeting physical activity guidelines) contributes to the rising cases of noncommunicable diseases and is responsible for over 7% of all-cause deaths along with a huge economic toll on the society. Recently, the COVID-19 crisis aggravated matters as many opportunities to be physically active were limited and sports clubs were temporarily closed. Today, effective interventions with a large reach are required to facilitate health behavior change towards more physical activity in the population. Here, even minor changes towards a more physically active lifestyle e.g. going for a daily ten-minute walk or interrupting prolonged physical inactivity can accumulate valuable health benefits over time.

There are a variety of evidence-based interventions for different settings which range from individual or group-based face-to-face interventions to digital interventions. While the former is well established in today's physical activity promotion, especially for rehabilitation, the latter is especially promising to promote physical activity on a broad scale due to the availability, fast-evolving technological progress, and ease of use of digital devices in modern society. Digital interventions for health behavior change can be delivered on desktop personal computers (e.g. via DVD), over the internet (e.g. on websites), or on mobile devices (e.g. via text message or mobile application). As nearly every household worldwide has access to and experience with at least one of those devices, the potential reach and cost-efficiency of such interventions are promising. Here, the use of information and communication technologies for health, in general, is defined as electronic health while every health practice supported by mobile devices is defined as mobile health. Recently, technological advances lead to the development of smaller, more convenient, and accurate devices to continuously measure physical activity (e.g. energy expenditure, step count, and classification of physical exertion), physiological (e.g. heart rate, blood sugar, and cortisol), and report psychological (e.g. valence, energetic arousal, and calmness) parameters. This opens up new perspectives using multilevel modeling in longitudinal designs to distinguish between within- and between-person effects and allows for a higher grade of individualization of interventions. One intervention type which greatly benefits from these continuous measurements and the technological advances is just-in-time adaptive interventions. These interventions aim to deliver interventional content (e.g. motivation to be

physically active) during the most promising time for the desired health behavior (i.e. physical activity) or during the most vulnerable time for unhealthy behavior (i.e. inactivity) and aim to maximize the usefulness of the intervention while minimizing participant burden. To do so, they rely on high-resolution data to depict opportune moments to deliver the intervention content. Recent progress with machine learning processes also benefits just-in-time adaptive interventions by offering sophisticated decision-making algorithms which can be guided by participants' behavior and preferences.

Previous studies on electronic and mobile interventions found heterogenic results for the effectiveness of digital health interventions for physical activity promotion. Here, evidence- and theory-based interventions which are guided by behavior change techniques (e.g. goal-setting or demonstration of behavior) were associated with higher intervention effectiveness. Furthermore, including the social context (e.g. peers, school, work, or family) in the interventions can be beneficial but it is important to distinguish between e.g. collaborative vs competitive settings based on participants' preferences. Finally, a high degree of individualization delivered by e.g. just-in-time adaptive interventions can enhance the effectiveness of mobile health interventions. However, the importance of the different interventional and contextual facets along with additional influences on the evaluation of the effectiveness remains unclear in the fast-developing field of electronic and mobile health behavior change interventions for children, adolescents, and adults.

To help close the gap between technological advances and the state of the research in electronic and mobile health interventions for physical activity promotion, this thesis aimed to 1) provide an overview of the effectiveness of electronic and mobile health interventions regarding physical activity promotion and 2) delve into important considerations and research gaps depicted by the overview (i.e. the choice of a measurement tool for physical activity and just-in-time adaptive interventions).

In our first paper, we conducted an umbrella review to summarize the evidence on the overall effectiveness of electronic and mobile health interventions along with the association of the key facets of theoretical foundation, behavior change techniques, social context, and just-in-time adaptive interventions with effectiveness. Derived from the eleven included reviews (182 original studies) we found significant benefits in favor of the intervention group (vs. control or over time) in the majority of interventions (59%). Here, the use of theoretical foundations and behavior change techniques were associated with higher effectiveness, the social context was often reported but not evaluated and just-in-time adaptive interventions were not included in any of the studies. One frequently reported shortcoming was the difficulty do compare self-reported and device-based measured results between studies. These findings suggest the potential effectiveness of digital interventions which is very likely facilitated by the key facets. Moreover, these findings helped us to determine promising but understudied facets of intervention effectiveness (i.e. just-in-time adaptive interventions) and depict frequently reported methodological issues (i.e. comparability of different measurement tools) which we could address within our thesis.

In our second paper, we explored the reliability, comparability, and stability of self-reported (i.e. questionnaire and physical activity diary) vs. device-based measured physical activity (i.e. analyzed using 10-second and 60-second epochs) in adults and children. We included two independent measurement weeks from 32 adults and 32 children in the control group of the SMART*FAMILY* trial to investigate if the differences between measurement tools were systematic over time. Here, participants wore an accelerometer on the right hip during daily life and completed a daily physical activity diary for seven consecutive days. Additionally, the international physical activity questionnaire was completed by participants at the end of each week. Results indicated non-systematic differences between the measurement tools (up to fourfold). Higher associations between the measurement tools were found for moderate than for vigorous physical activity and the results differed between children and adults. These results confirm the importance of carefully considering the measurement tool to be suitable for the research question and target group and the very limited comparability between different measurement tools. Additionally, the differences within accelerometer-derived results (10-second epochs vs. 60-second epochs) point to the need for comprehensive reporting for each measurement tool to compare and replicate the results.

In our third paper, we summarized previous frameworks of just-in-time adaptive interventions and pointed out opportunities and challenges within this research field. We combined recommendations of three previous frameworks and refined that just-in-time adaptive interventions should 1) correspond to real-time needs; 2) adapt to input data; 3) be system-triggered. This can be enhanced by 4) be goal-oriented; and 5) be customized to user preferences. By doing so, just-in-time adaptive interventions can achieve a high degree of individualization which is closely fitted to each individual. The main challenge hereby remains the opportune moment identification (i.e. the exact moment when participants are either likely to engage in unhealthy behavior or when they face opportunities to perform healthy behaviors) to timely deliver intervention content. This can be explored using ambulatory assessments and assessing the context of the behavior. The decision-making process can be enhanced by machine learning algorithms. These results guided the reporting and design of the examinations included in our fourth and fifth papers.

In our fourth paper, we evaluated the importance of engaging with a just-in-time adaptive intervention triggered after a period of physical inactivity. For this secondary data analysis, 47 adults and 33 children were included in the analysis who wore an accelerometer on the right hip and used our SMART*FAMILY*2.0 application during the three-week intervention period of the SMART*FAMILY*2.0 trial. Here, we analyzed 907 just-in-time adaptive intervention triggers and compared step and metabolic equivalent count in the hour after occasions when participants answered the trigger (i.e. responded to the question regarding their previous physical inactivity) within 60 minutes ("engaged" condition) with the hour after occasions when they did not answer the trigger within 60 minutes ("not engaged" condition) in the mobile application. Results indicated significantly higher metabolic equivalent and step count for the "engaged" condition within-persons. This shows that if a person engaged with a trigger within 60 minutes, he or she showed significantly higher physical activity in the following hour compared to when the same person did not engage with the trigger. This expands previous research about participants' engagement with the intervention and the importance of an opportune moment identification to enhance this engagement.

In our fifth paper, we explored the association of sleep quality and core affect with physical activity during a mobile health intervention period. Based on the same intervention period reported in the fourth paper, but with different inclusion criteria for the data (e.g. minimum wear time of the accelerometer for 8 hours per day instead of 80% of the hour of interest), daily accumulated self-rated mental state was compared to step count and minutes of moderate-to-vigorous physical activity for 49 adults and 40 children in a secondary data analysis. Overall, 996 measurement days of the participants were included in this analysis. Our results showed that higher reported valence and energetic arousal values were associated with more physical activity, while higher reported calmness values were associated with less physical activity within-persons on the same day. No distinct association was found between sleep quality and physical activity. Our results confirm previous ambulatory assessment studies and we suggest that within-person associations of core affect should be considered when designing physical activity interventions for both children and adults. Additionally, core affect might be a promising consideration for opportune moment identifications in just-in-time adaptive interventions

to evaluate the feasibility and causality of targeting changes in e.g. valence to improve subsequent and daily physical activity of participants using micro-randomized trials.

Based on the current state of knowledge, our results above address important research gaps depicted by our overview in the field of digital interventions for physical activity promotion. One example is the understudied area of just-in-time adaptive interventions for which we provided a framework, evaluated the effect of engaging with such interventions on subsequent physical activity, and explored core affect and sleep quality as facilitators of physical activity behavior. With these findings in mind, we discussed important considerations to progress future mobile health studies for physical activity promotion in general, and just-in-time adaptive interventions in particular at the end of this work. Finally, we aimed to transfer this knowledge into a proposal for designing a just-in-time adaptive intervention in the special group of participants at risk for or with knee osteoporosis who could specifically benefit from this highly individualized approach.

Zusammenfassung

Die weitreichenden gesundheitlichen Vorteile von körperlicher Aktivität, insbesondere die Effekte zur Prävention nichtübertragbarer Krankheiten wie Herz-Kreislauf-Erkrankungen, Krebs und Diabetes, sind heutzutage gut belegt. Allerdings liegt das heutige Maß an körperlicher Aktivität in der Bevölkerung deutlich unter den von beispielsweise der Weltgesundheitsorganisation empfohlenen Werten für alle Altersgruppen. Dieses Ausmaß an körperlicher Inaktivität (d. h. das Nichteinhalten der Richtlinien für körperliche Aktivität) trägt heute beträchtlich zu den steigenden Fällen nichtübertragbarer Krankheiten bei, ist für über 7% aller Todesfälle verantwortlich und führt zu enormen wirtschaftlichen Kosten für die Gesellschaft. Zuletzt verschärfte die COVID-19-Krise die Situation, da viele Möglichkeiten sich körperlich zu betätigen eingeschränkt und Sportvereine vorübergehend geschlossen wurden. Daher sind wirksame Interventionen mit großer Reichweite erforderlich, um eine Änderung des Gesundheitsverhaltens hin zu mehr körperlicher Aktivität in der Bevölkerung zu ermöglichen. Selbst geringfügige Änderungen für einen körperlich aktiveren Lebensstil wie ein täglicher zehnminütiger Spaziergang oder die Unterbrechung längerer Phasen von körperlicher Inaktivität können hier über längere Zeit wertvolle gesundheitliche Vorteile bringen.

Gegenwärtig existiert eine Vielzahl evidenzbasierter Interventionen für unterschiedliche Settings, die von individuellen und gruppenbasierten Face-to-Face-Interventionen bis zu digitalen Interventionen reichen. Erstere (Face-to-Face-Interventionen) sind dabei in der derzeitigen Bewegungsförderung, insbesondere in der Rehabilitation, gut etabliert, Letzteres (digitale Interventionen) scheinen allerdings aufgrund der Verfügbarkeit des sich schnell entwickelnden technologischen Fortschritts und der Benutzerfreundlichkeit dieser Geräte besonders vielversprechend, um körperliche Aktivität in der modernen Gesellschaft als präventive Maßnahme zu fördern. Digitale Interventionen können auf Desktop-PCs (z. B. über DVD), über das Internet (z. B. auf Websites) oder auf mobilen Geräten (z. B. per SMS oder mobiler Anwendung) bereitgestellt werden. Hierbei wird die Nutzung von Informations- und Kommunikationstechnologie für die Gesundheit im Allgemeinen als „Electronic-Health“ definiert, während jede durch mobile Geräte unterstützte Gesundheitspraxis als „Mobile Health“ definiert wird. Da fast jeder Haushalt heutzutage Zugang zu und Erfahrung mit mindestens einem dieser Geräte hat, sind die potenzielle Reichweite und Kosteneffizienz solcher Interventionen verheißungsvoll. Die technologischen Fortschritte in letzter Zeit führten dabei zur Entwicklung kleinerer, praktischerer und genauerer Geräte zur kontinuierlichen Messung von körperlicher Aktivität, wie

beispielsweise Energieverbrauch, Schrittzahl und Klassifizierung der körperlichen Anstrengung, physiologischer Parameter, wie zum Beispiel Herzfrequenz, Blutzucker und Cortisol, sowie selbstberichteter psychologischer Parameter. Dazu zählen beispielsweise Stimmung, energetische Erregung und Ruhe. Diese Fortschritte eröffnen neue Perspektiven durch Längsschnittdesigns zur Unterscheidung von Effekten innerhalb und zwischen Personen. Dadurch können Interventionen gut individualisiert werden. Eine spezifische Art der Interventionen, die stark von diesen kontinuierlichen Messungen und den technologischen Fortschritten profitiert, sind so genannte „Just-In-Time-Adaptive-Interventionen“. Diese Interventionen zielen darauf ab, die Intervention während der vielversprechendsten Zeit für das gewünschte Verhalten (z. B. körperliche Aktivität) oder während der anfälligsten Zeit für unerwünschtes Verhalten (z. B. Inaktivität) bereitzustellen. Sie beabsichtigen damit den Nutzen der Intervention zu maximieren und gleichzeitig die Belastung der Teilnehmer:innen zu minimieren. Um günstige Momente für die Bereitstellung der Interventionsinhalte zu erfassen, stützen sich die Entscheidungsprozesse dabei auf hochauflösende Längsschnittdaten. Just-In-Time-Adaptive-Interventionen profitieren auch von den jüngsten Fortschritten bei maschinellen Lernprozessen durch ausgeklügelte Entscheidungsfindungsalgorithmen, die vom Verhalten und den Vorlieben der Teilnehmer gesteuert werden können.

In früheren Studien zu mobilen Interventionen waren die Ergebnisse für die Wirksamkeit digitaler Interventionen zur Steigerung von körperlicher Aktivität sehr heterogen. Hier war die Implementierung von evidenz- und theoriebasierten Interventionen, die von Techniken zur Verhaltensänderung, wie beispielsweise einer Zielsetzung oder Verhaltensdemonstration geleitet wurden, mit einer höheren Interventionseffektivität verbunden. Darüber hinaus kann es von Vorteil sein, den sozialen Kontext (z. B. Familie, Arbeitsplatz oder Freunde) in die Interventionen mit einzubeziehen, wobei es wichtig ist, zwischen den Präferenzen der Teilnehmer:innen zu unterscheiden (z. B. kollaboratives vs. kompetitives Setting). Als vierter wichtiger Aspekt kann ein hohes Maß an Individualisierung durch Just-In-Time-Adaptive-Interventionen die Wirksamkeit mobiler Gesundheitsinterventionen verbessern. Die meisten früheren Studien und insbesondere die weit verbreiteten kommerziellen Angebote wurden bisher unzureichend auf ihre Wirksamkeit geprüft. Darüber hinaus sind die Bedeutung der unterschiedlichen Facetten sowie die zusätzlichen Einflüsse auf die Bewertung der Wirksamkeit in dem sich schnell entwickelnden Bereich der elektronischen und mobilen Gesundheitsinterventionen für Kinder, Jugendliche und Erwachsene noch unklar.

Hier wird deutlich, dass sich zwischen technologischem Fortschritt und dem Stand der Forschung zu elektronischen und mobilen Gesundheitsinterventionen zur Bewegungsförderung eine Lücke befindet. Mit dieser Arbeit wollen wir einen Beitrag dazu leisten, diese Lücke zu schließen. Dazu geben wir einen Überblick über die Wirksamkeit elektronischer und mobiler Gesundheitsinterventionen zur Bewegungsförderung und beleuchten darüber hinaus wichtige Überlegungen und Forschungslücken, die in der Übersicht aufgezeigt wurden. Hierzu gehören unter anderem die Wahl eines Messinstruments für körperliche Aktivität und Just-In-Time-Adaptive-Interventionen.

In unserem ersten Artikel führten wir ein Umbrella-Review durch, welches die Evidenz zur Wirksamkeit elektronischer und mobiler Gesundheitsinterventionen zusammenfasst. Im zweiten Schritt lag unser Fokus dabei auf dem Einfluss der Schlüsselaspekte: theoretische Grundlagen, Techniken zur Verhaltensänderung, sozialer Kontext sowie Just-in-Time-Adaptiven-Interventionen auf die Effektivität der Interventionen. Abgeleitet aus den elf darin eingeschlossenen Reviews, die sich wiederum aus 182 Originalstudien zusammensetzten, zeigten sich bei der Mehrheit der Interventionen (59 %) signifikante Vorteile zugunsten der Interventionsgruppe (vs. Kontrolle oder im zeitlichen Verlauf). Hierbei wurde eine theoretische Fundierung und Techniken zur Verhaltensänderung mit einer höheren Wirksamkeit assoziiert. Die Autor:innen der eingeschlossenen Reviews berichteten dabei zwar oft vom sozialen Kontext der Interventionen, dieser wurde jedoch nicht bewertet und Just-in-Time-Adaptive-Interventionen wurden in keiner der Studien berücksichtigt. Des Weiteren haben die Autor:innen der Reviews den Vergleich von selbstberichteten und gerätebasierten Messergebnissen als häufiges Problem für die Synthese der Ergebnisse berichtet. Diese Ergebnisse deuten auf die potenzielle Wirksamkeit digitaler Interventionen hin, die sehr wahrscheinlich durch die oben genannten Schlüsselaspekte beeinflusst werden. Darüber hinaus helfen uns diese Erkenntnisse, vielversprechende aber wenig untersuchte Faktoren der Interventionswirksamkeit zu bestimmen (Just-in-Time-Adaptive-Interventionen) und zusätzlich häufig berichtete methodische Probleme aufzuzeigen (Vergleichbarkeit verschiedener Messmethoden), auf die im Rahmen dieser Arbeit eingegangen wird.

In unserem zweiten Artikel untersuchten wir die Zuverlässigkeit, die Vergleichbarkeit und die Stabilität von selbstberichteter körperlicher Aktivität durch Fragebogen und Tagebuch im Vergleich zu gerätebasiert gemessener körperlicher Aktivität mittels Akzelerometrie, analysiert anhand von 10-Sekunden- und 60-Sekunden-Epochen, bei Erwachsenen und Kindern. Hierfür wurden die zwei unabhängigen Messwochen der Kontrollgruppe der SMART*FAMILY*-

Studie mit je 32 Erwachsenen und Kindern verwendet, um zu untersuchen, ob die Unterschiede zwischen den Messinstrumenten über den Lauf der Zeit systematisch sind. Dabei trugen die Teilnehmer:innen während ihres normalen Alltags einen Akzelerometer an ihrer rechten Hüfte, füllten ein tägliches Aktivitätstagebuch aus und vervollständigten den internationalen Fragebogen für körperliche Aktivität am Ende jeder Woche. Die Ergebnisse zeigten nicht-systematische Unterschiede zwischen den Messinstrumenten, die sich bis zu einem Unterschied der vierfachen Höhe unterschieden. Hierbei wurden höhere Korrelationen für moderate als für anstrengende körperliche Aktivität zwischen den Messinstrumenten gefunden und auch zwischen Kindern und Erwachsenen zeigten sich unterschiedliche Ergebnisse. Dies bestätigt, wie wichtig es ist, das Messinstrument sorgfältig auf seine Eignung für den jeweiligen Kontext sowie die Fragestellung abzuwägen und die sehr begrenzte Vergleichbarkeit zwischen verschiedenen Messinstrumenten zu berücksichtigen. Darüber hinaus weisen die Unterschiede innerhalb der vom Akzelerometer abgeleiteten Ergebnisse (10-Sekunden-Epochen vs. 60-Sekunden-Epochen) auf die Bedeutung einer klaren Berichterstattung für jedes Messwerkzeug hin.

In unserem dritten Artikel haben wir frühere Frameworks für Just-in-Time-Adaptive-Interventionen zusammengefasst und Möglichkeiten sowie Herausforderungen in diesem Forschungsfeld aufgezeigt. Abgeleitet von den Empfehlungen der drei früheren Frameworks zeigten wir, dass adaptive Just-in-Time-Interventionen 1) den Echtzeit-Bedürfnissen entsprechen sollten, 2) an erfasste Daten angepasst sein müssen und 3) vom System ausgelöst werden. Verbesserungen können durch 4) zielorientierte Interventionen und 5) Anpassungen an Benutzerpräferenzen erreicht werden. Wenn dies berücksichtigt wird, kann durch diese Interventionen ein hohes Maß an Individualisierung erreicht werden, das präzise auf das Individuum zugeschnitten ist. Die größte Herausforderung bleibt dabei die Identifizierung des geeigneten Moments, also genau dann, wenn die Teilnehmer:innen entweder wahrscheinlich ungesundes Verhalten beginnen oder sie Gelegenheit haben, gesundes Verhalten zu zeigen, um rechtzeitig Interventionsinhalte bereitzustellen. Dies kann mit ambulanten Assessments untersucht und der Entscheidungsprozess mit maschinellen Lernalgorithmen verbessert werden. Diese Ergebnisse leiteten die Berichterstattung und das Design der Erhebungen, die in der vierten und fünften Publikation dieser Arbeit enthalten sind.

In unserem vierten Artikel haben wir die Wichtigkeit der Interaktion mit einer Just-in-Time-Adaptive-Interventionen bewertet, die nach einer Zeit der körperlichen Inaktivität ausgelöst wird. Für diese sekundäre Datenanalyse wurden 47 Erwachsene und 33 Kinder während des dreiwöchigen Interventionszeitraums der SMART*FAMILY*2.0-Studie miteinbezogen, die einen Akzelerometer an ihrer rechten Hüfte trugen und unsere SMART*FAMILY*2.0-Applikation nutzten. Hier wurden 907 Trigger der Just-in-Time-Adaptive-Interventionen analysiert und mit der Anzahl der Schritte und metabolischen Äquivalente in der Stunde nach Ereignissen, in denen die Teilnehmer den Trigger in der mobilen Anwendung innerhalb von 60 Minuten beantworteten („engaged“ Kondition), mit der Stunde nach Ereignissen, in denen sie nicht innerhalb von 60 Minuten antworteten („not engaged“ Kondition) verglichen. Die Ergebnisse weisen auf eine signifikant höhere Anzahl an metabolischen Äquivalenten und eine signifikant höhere Schrittzahl für die „engaged“ Kondition innerhalb von Personen hin. Dies bedeutet, dass, wenn eine Person auf den Trigger reagierte, diese Person in der Stunde nach dem Trigger aktiver war, als wenn er oder sie nicht auf den Trigger reagierte. Dadurch wird die bisherige Forschung durch die Bedeutung der Interaktion der Teilnehmer mit der Intervention erweitert und es wird bestärkt, wie wichtig es ist, jeweils günstige Momente dafür zu identifizieren, um diese Interaktionen zu erhöhen.

In unserem fünften Artikel untersuchten wir den Zusammenhang von Schlafqualität und Core-Affect (Stimmung, energetische Erregung und Ruhe) mit körperlicher Aktivität während einer mobilen Gesundheitsintervention. Basierend auf der gleichen Interventionsperiode wie oben angegeben, aber mit unterschiedlichen Einschlusskriterien für die Daten (z. B. Mindesttragezeit des Akzelerometers für 8 Stunden pro Tag anstelle von 80 % der relevanten Stunde), wurde der Zusammenhang des auf Tagesebene akkumulierten selbstbewerteten mentalen Zustandes mit der Schrittzahl und den Minuten mittlerer bis intensiver körperlicher Aktivität bei 49 Erwachsenen und 40 Kindern verglichen. Insgesamt wurden 996 Tage in diese sekundäre Datenanalyse einbezogen. Unsere Ergebnisse zeigten, dass höhere Werte für Stimmung und energetische Erregung mit mehr körperlicher Aktivität assoziiert waren, während höhere Werte für Ruhe mit weniger körperlicher Aktivität innerhalb von Personen am selben Tag in Verbindung gebracht wurden. Zwischen Schlafqualität und körperlicher Aktivität zeigte sich kein eindeutiger Zusammenhang. Somit bestätigen unsere Ergebnisse frühere Studien mit ambulantem Assessment und legen nahe, dass bei der Gestaltung von Interventionen sowohl für Kinder als auch für Erwachsene Zusammenhänge von Core-Affect mit körperlicher Aktivität berücksichtigt werden sollten. Darüber hinaus könnte Core-Affect eine vielversprechende Möglichkeit zur Identifizierung günstiger Momente für Just-in-Time-Adaptive-Interventionen darstellen. Zudem sollten zukünftige Studien die Machbarkeit und Kausalität von Änderungen in beispielsweise der Stimmung von Proband:innen zur Verbesserung der täglichen körperlichen Aktivität der Teilnehmer:innen überprüfen.

Basierend auf dem aktuellen Wissensstand und abgeleitet aus unserem Umbrella-Review behandeln unsere obigen Ergebnisse wichtige aufgezeigte Forschungslücken im Bereich digitaler Interventionen zur Bewegungsförderung. Ein Beispiel ist hierbei der bisher wenig untersuchte Bereich der Just-in-Time-Adaptiven-Interventionen, für welche wir ein Framework bereitstellen, die Wirkung der Interaktion mit diesen Interventionen überprüfen und den Zusammenhang von körperlicher Aktivität sowohl mit Core-Affect als auch mit Schlafqualität untersuchen. Vor dem Hintergrund dieser Ergebnisse werden am Ende dieser Arbeit wichtige Überlegungen diskutiert, um zukünftige mobile Gesundheitsstudien zur Förderung der körperlichen Aktivität im Allgemeinen und Just-in-Time-Adaptive-Interventionen im Besonderen voranzutreiben. Schließlich überführen wir dieses Wissen in einen Vorschlag für die Konzeption einer Just-in-Time-Adaptive-Intervention in der speziellen Gruppe von Teilnehmer:innen mit einem Risiko für oder mit Knie-Osteoporose, die von diesem hochindividualisierten Ansatz besonders profitieren könnten.

Content

List of Figures

List of Tables

Chapter 2

Chapter 3

Chapter 5

Chapter 6

Preface

Parts of this work have been published or have been submitted for publication.

Chapter 2: Fiedler, J., Eckert, T., Wunsch, K., & Woll, A. (2020). Key facets to build up eHealth and mHealth interventions to enhance physical activity, sedentary behavior and nutrition in healthy subjects - an umbrella review. *BMC public health*, *20*(1), 1605. https://doi.org/10.1186/s12889-020-09700-7

Chapter 3: 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). https://doi.org/10.3390/s21082672

Chapter 4: Wunsch, K.*, Eckert, T.*, Fiedler, J.*, & Woll, A. (2022). Just-in-time adaptive interventions in mobile physical activity interventions - A synthesis of frameworks and future directions. *The European Health Psychologist*, *22*(4), 834–842. https://doi.org/10.5445/IR/1000143174

Chapter 5: Fiedler, J.*, Seiferth, C.*, Eckert, T., Woll, A., & Wunsch, K. (submitted). A just-in-time adaptive intervention to enhance physical activity in the SMART*FAMILY*2.0 trial. *Sport, Exercise, and Performance Psychology.*

Chapter 6: Fiedler, J., Caroline Seiferth, Eckert, T., Woll, A., & Wunsch, K. (2022). Sleep quality, valence, energetic arousal, and calmness as predictors of device-based measured physical activity during a three-week mHealth intervention: An ecological momentary assessment study within the SMART*FAMILY* trial. *German Journal of Exercise and Sport Research.* https://doi.org/10.1007/s12662-022-00809-y

**Authors contributed equally to this work*

Chapter 1 General Introduction

Physical activity is an important contributor to human health with a variety of benefits (Warburton et al., 2006; Warburton & Bredin, 2017). Being physically active impacts all-cause mortality, cardiovascular function, muscular fitness, metabolic health, body constitution, mental health, and cognitive function among others (Bamman et al., 2014; Neufer et al., 2015). By enhancing these important physiological and psychological aspects, physical activity depicts one effective key strategy to prevent noncommunicable diseases like cardiovascular diseases, diabetes, and obesity through all stages of life (Beaglehole et al., 2011; Bull et al., 2020). Despite that knowledge, noncommunicable diseases are on the rise for many years, with 3.2 million deaths attributable to physical inactivity (Forouzanfar et al., 2016), and recent research added that physical inactivity is responsible for 7.2% of all-cause deaths (Katzmarzyk et al., 2022). Furthermore, the association of physical inactivity with non-communicable diseases ranged from 1.6% for hypertension to 8.1% for dementia (Katzmarzyk et al., 2022). These health deteriorating effects create a yearly financial toll of 67.5 billion US $ on the health system (Ding et al., 2016). Recently, the COVID-19 crisis aggravated matters as noncommunicable diseases constitute important risk factors for severe illness (Bello & Useh, 2021; Pan et al., 2021), and opportunity for being physically active was limited by lockdowns and contact restrictions (Stockwell et al., 2021; Wunsch, Kienberger & Niessner, 2022). While the COVID-19 crisis did and does have an impact on physical activity, levels of physical activity have also been insufficient for health benefits throughout all age groups for some time (Blair, 2009; Guthold et al., 2018, 2020; Woll et al., 2011). Furthermore, the modern lifestyle continues to promote inactivity (e.g. desk work or watching television) which increased over the past years (Church et al., 2011; González et al., 2017; Tremblay et al., 2011).

Today, important international and national guidelines concerning beneficial amounts of physical activity (international e.g. Bull et al., 2020, and national e.g. Pfeifer & Rütten, 2017) are broadly available for practitioners, policymakers, and researchers but have not been sufficient to sustainably change health behavior and reduce the financial and health burden from noncommunicable diseases worldwide (Katzmarzyk et al., 2022; World Health Organization, 2014). It is important to note that the guidelines of e.g. 150 minutes moderate or 75 minutes vigorous or a combination of both for adults and more than 60 minutes moderate to vigorous physical activity on average per day for children (Bull et al., 2020) are important recommendations but they should not be interpreted as miraculous thresholds. Extensive research showed that every move (Ekelund et al., 2019) and every step (Paluch et al., 2022) counts. Even small changes in physical activity over an extended amount of time can accumulate significant and relevant benefits e.g. for cardiovascular health (Martinez-Gomez et al., 2022). Therefore, effective, broadly available, and accepted long-term interventions to enhance physical activity are needed to sustainably improve health behavior, and to limit the development of noncommunicable diseases in the future (World Health Organization, 2018).

While there are a variety of evidence-based intervention methods and settings for physical activity promotion (Greaves et al., 2011; Heath et al., 2012; Kahn et al., 2002), electronic and mobile health interventions, in particular, are promising and upcoming opportunities to enhance physical activity on a large scale (Michie et al., 2017; Vandelanotte et al., 2016). Electronic health refers to all interventions that include “the use of information and communication technologies for health” (World Health Organization, 2020) while mobile health interventions refer to “medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices” (World Health Organization, 2011). With 6.3 billion smartphone users worldwide in 2021 (Statista, 2021), the potential coverage of mobile health tools coupled with intuitive and autonomous controls of the device is promising. This is especially true for the digital native generations, who use mobile health applications regularly (Naszay et al., 2018). To enhance the effectiveness of mobile health interventions, evidence- and theory-based intervention features should be implemented into the intervention (Direito et al., 2014; Mercer et al., 2016; Schoeppe et al., 2016; Yang et al., 2015) which are often guided by behavior change techniques (Michie et al., 2011; Michie et al., 2013). Furthermore, recent technological developments led to more convenient and accurate devices with the ability to continuously measure physical activity (Burchartz et al., 2020) and related parameters like heartrate, blood sugar, and core affect (Reichert et al., 2020). These parameters can be combined and used to adapt physical activity interventions precisely to the participants’ needs, preferences, and availability. Here, previous research showed that the amount of individualization can be an important factor for effective mobile health interventions (Baumann et al., 2022). One upcoming and highly personalized intervention type are so-called just-in-time adaptive interventions. These interventions aim to deliver the intervention during the most promising time for the desired behavior (i.e. physical activity) or during the most vulnerable time for unhealthy behavior (i.e. inactivity) and aim to maximize the usefulness while minimizing participant burden (Hardeman et al., 2019).

As promising as mobile health interventions seem to be, various open questions remain surrounding the inclusion and evaluation of such interventions in physical activity promotion (Schoeppe et al., 2016; Vandelanotte et al., 2016). Here, previous research suggests that including theoretical background and using behavior change techniques (Prestwich et al., 2014; Webb et al., 2010), implementing the intervention in a social context (Morrison et al., 2012; Umberson et al., 2010), and enhancing mobile health interventions with just-in-time adaptive interventions (Hardeman et al., 2019; Nahum-Shani et al., 2018) are key facets to enhance physical activity. However, the importance of the different facets along with additional influences on the evaluation of the effectiveness remains unclear in the fast-developing field of electronic and mobile health behavior change interventions for children, adolescents, and adults.

Based on this research gap, the aim of this thesis was to provide an overview of electronic and mobile health interventions for physical activity promotion. Derived from these results, we delved deeper into the important consideration of the choice of the measurement tool and the understudied aspect of just-in-time adaptive interventions (see Figure 1).

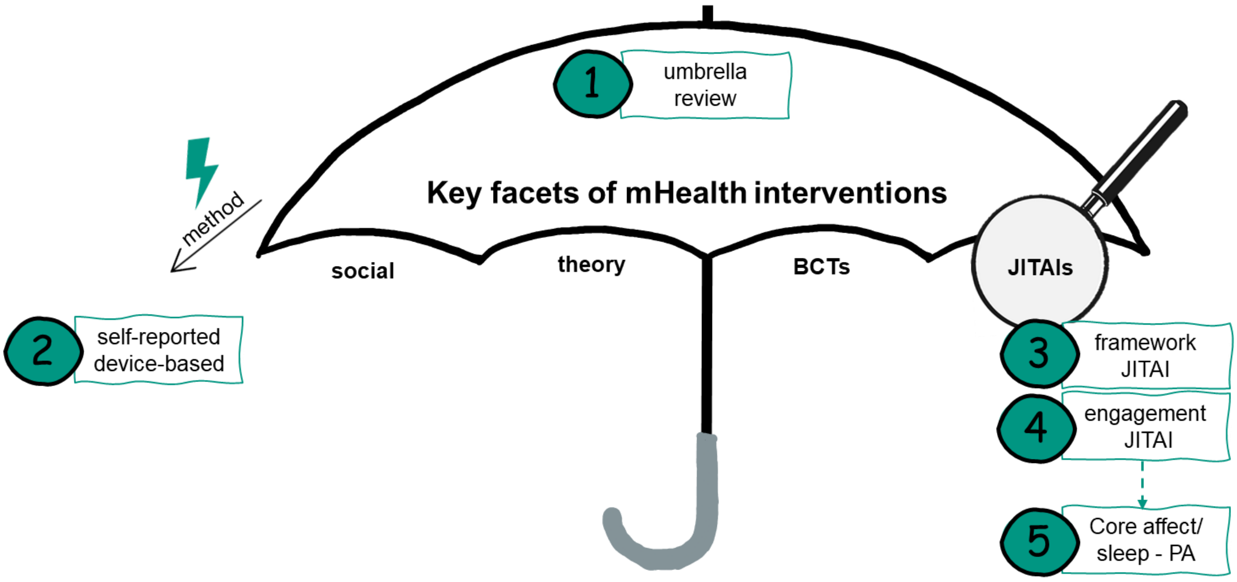

Figure 1. Overview of the five publications (numbers 1-5) included in the doctoral thesis. Research gaps found by the umbrella review (1) indicated the focus of the following publications regarding methodology (2) and just-in-time adaptive interventions (3-5). Abbreviations: BCTs: behavior change techniques, just-in-time adaptive interventions: just-in-time adaptive interventions, PA: physical activity.

Data used for publications two, four, and five origin from the two trials of the SMART*FAMILY* study (Wunsch et al., 2020). This study is a cluster-randomized, theory-based (self-determination theory (Ryan & Deci, 2000)) mobile health intervention with several behavior change techniques (e.g. goal-setting, provide feedback on performance, and plan social change (Michie et al., 2011; Michie et al., 2013)) to enhance physical activity and healthy eating in the social context of families. Data from the first trial is used for publication two, and data from the second (SMART*FAMILY2.0*) for publications four and five.

As a first step, we conducted an umbrella review to evaluate what is needed to progress the field of digital behavior change. We provided an overview of the overall effectiveness of electronic and mobile health interventions for behavior change under consideration of the a priory defined key facets theoretical foundation, behavior change techniques, social context, and just-in-time adaptive interventions in healthy children and adults (published: Fiedler et al., 2020). As physical activity, physical inactivity, and healthy eating are often included together in behavior change interventions due to their interactive effects on health (Michie et al., 2011) we focused on all three health behaviors in this umbrella review.

Question 1

How effective are digital health interventions and which key facets are related to effectiveness?

Key facets to build up eHealth and mobile health interventions to enhance physical activity, sedentary behavior, and nutrition in healthy subjects – an umbrella review.

In our first work (Fiedler et al., 2020), we found the majority (59%) of digital interventions to be effective over time and/or vs control in all age groups. The strongest evidence concerning the impact of the included key facets on effectiveness was found for theoretical foundation and behavior change techniques. The social context was often included in the studies (e.g. by school-based or workplace-based interventions) but the influence has not been analyzed while no study included just-in-time adaptive interventions.

Even though the overall results were promising, a more detailed inspection is needed to interpret these results. First, the studies ranged from 1997 to 2018 which is a very long time for the fast-evolving field of digital health and physical activity assessment (Burchartz et al., 2020). Second, the included meta-analyses for physical activity and inactivity did not find a significant pooled effect in favor of the interventions (Direito et al., 2017). Third, a variety of measurement tools have been used in the studies, ranging from device-based measured physical activity to non-validated self-report tools for some healthy eating studies. Fourth, sustainability of the significant effects over time was low and important metrics like engagement with the digital intervention over time were often not sufficiently reported which is a common issue in digital health studies (Mclaughlin et al., 2021; Vandelanotte et al., 2007). These important limitations underline the necessity to evaluate the differences between significant (and relevant) and non-significant findings within the field to develop strong recommendations for future research. Therefore, we summarized future research recommendations from all included reviews. Derived from these results, our recommendations are to uniformly and fully report all relevant aspects of the studies (e.g. intervention, methods, and outcomes), to evaluate the role of social context and relations as well as just-in-time adaptive interventions, to include power analyses in all original research publications, and to assess the cost-effectiveness of digital interventions. By adhering to these recommendations, future studies can yield a far clearer picture of the effectiveness of different aspects of digital health interventions.

One of the most frequently mentioned challenges was the comparison of self-reported and device-based measured physical activity which impacts the comparability across studies with different measurement tools or analyzes (e.g. epoch lengths of accelerometry (Edwardson & Gorely, 2010)) by a large margin. This complicates e.g. the classification of someone as sufficiently active based on physical activity guidelines (Bull et al., 2020), impedes the comparison of studies with different outcomes (Skender et al., 2016), and hampers the use of just-in-time adaptive interventions which rely on transparent physical activity estimations for their algorithms (Gonul et al., 2019).

Derived from this result, our second publication focused on the exploration of reliability, comparability, and stability of self-reported (i.e. questionnaire and physical activity diary) vs. device-based measured physical activity (i.e. analyzed using 10-second and 60-second epochs) in adults and children (published: Fiedler et al., 2021).

Question 2

How do self-reported and device-based measured physical activity compare in the SMART*FAMILY* study?

Comparison of self-reported and device-based measured physical activity using measures of stability, reliability, and validity in adults and children.

In our second work (Fiedler et al., 2021), the physical activity data of two independent measurement weeks in the SMART*FAMILY* study (Wunsch et al., 2020) has been compared between questionnaires, diary, and accelerometry. Descriptive results showed the highest physical activity values for questionnaires followed by accelerometry and the diary. We found two- to four-fold differences between the measurement tools which has a large impact on the interpretation (e.g. 810 vs. 282 minutes of moderate physical activity/week). Only the results of the accelerometry were found to be reliable, comparable (between 10-second and 60-second epochs), and stable between the two weeks, even though the absolute values also showed meaningful differences.

Our overall results are in line with previous studies which found that reliability within the measurement tools was higher than comparability between the tools (Bull et al., 2009; Dyrstad et al., 2014; Hagstromer et al., 2010; Herrmann et al., 2013; Hukkanen et al., 2018; Skender et al., 2016; Slootmaker et al., 2009), and that epoch lengths have an important impact for the interpretation of physical activity outcomes (Edwardson & Gorely, 2010; Orme et al., 2014). Contrasting previous research, we found limited reliability for the diary and the international physical activity questionnaire concerning adults' vigorous physical activity. Importantly, the inclusion of a stability measure (i.e. comparability between two measurement instruments over time) indicated stable results between the two epoch lengths and for adults' moderate physical activity concerning all measurement tools, but very limited stability for all other comparisons. These results add to the evidence that adults have a much more stable physical activity pattern compared to children (Livingstone et al., 2003). A major limitation of this examination is that accelerometry is no gold standard to assess physical activity (Keadle et al., 2019) and the results refer to comparability and not validity. Therefore, no or only a very limited conclusion about the accuracy of the measurements compared to the actual physical activity of participants can be drawn.

Overall, it remains a tedious task to decide which physical activity measure to use and how to compare studies with different outcomes. Here, including stability measurements in future evaluation studies could strengthen the interpretation of the comparability between different measurement tools. When conducting a physical activity study, it remains important to spend time choosing an instrument and suitable data processing method fitting to the purpose of the study design, study population, and outcome of interest (Burchartz et al., 2020). Furthermore, studies with varying measurement tools should only be compared and interpreted with great caution, especially if the methodology is not reported in greater detail. Ideally, open science practices should be used to limit the black box regarding data processing and analyses to accelerate digital health promotion research (Kwasnicka et al., 2022). Finally, we will conduct a replication study with new data from the SMART*FAMILY*2.0 trial to enhance the credibility of the results of this examination in the future (in progress).

Another important but understudied consideration for digital health behavior change which we derived from our umbrella review (Fiedler et al., 2020) are just-in-time adaptive interventions for healthy participants. While these interventions allow for a highly individualized provision of behavior change support (Conroy et al., 2020; Gonul et al., 2019; Nahum-Shani et al., 2018; Schembre et al., 2018), the evidence for their effectiveness is still limited (Hardeman et al., 2019; Miller, 2019). To provide an overview of this fourth key facet, we combined previous frameworks of just-in-time adaptive interventions (Gonul et al., 2019; Hardeman et al., 2019; Nahum-Shani et al., 2018) and discussed opportunities and challenges to provide an overview and future directions for this field in our third publication (published: Wunsch*, Eckert*, Fiedler* et al., 2022).

Question 3

Which opportunities and challenges of just-in-time adaptive interventions need to be considered for physical activity promotion?

Just-in-time adaptive interventions in mobile physical activity interventions – A synthesis of frameworks and future directions.

Based on the combination of previous frameworks, we defined (Wunsch*, Eckert*, Fiedler* et al., 2022) that just-in-time adaptive interventions should:

1) correspond to real-time needs (at the moment when a participant can benefit from support);

2) adapt to input data (sensor input like minutes of inactivity or user-input like availability);

3) be system-triggered (automatic trigger without the involvement of the participant);

and can be enhanced by

4) be goal-oriented (feedback on and suggestions for goal-achievement);

5) be customized to user preferences (user can e.g. choose time frames when triggers are muted).

Opportunities for just-in-time adaptive interventions are manifold. The goal that each participant can be provided with the exact amount of support he or she needs in a certain situation is outstanding and enhances the ecological validity of the intervention (Heron & Smyth, 2010). This can be achieved by combining a variety of sensors, self-reports, and e.g. access to the calendar with machine learning algorithms (Conroy et al., 2020; Gonul et al., 2019; Nahum-Shani et al., 2018; Reichert et al., 2020). However, due to the novelty of this topic, current original research studies are mainly focused on feasibility and have limited evidence for effectiveness (Hardeman et al., 2019; Rabbi et al., 2015; Thomas & Bond, 2015). The main challenge for these interventions remains to tailor the decision points (the time when the just-in-time adaptive intervention is sent to the participant) and decision rules (determining the timing, frequency, and duration of the just-in-time adaptive intervention at each decision point framed by tailoring variables like sensor or user input) precisely to the users' needs. This opportune moment identification is crucial for a high acceptance of participants, enhanced engagement with, and therefore potential effectiveness of the intervention (Gonul et al., 2019).

To evaluate the importance of participants' timely engagement with a just-in-time adaptive intervention to break inactive phases, we conducted a secondary data analysis of the intervention period in the SMART*FAMILY*2.0 study in our fourth publication (Fiedler*, Seiferth* et al., submitted).

Question 4: How important is the engagement with a just-in-time intervention to increase physical activity after inactive phases.

A just-in-time adaptive intervention to enhance physical activity in the SMARTFAMILY2.0 trial.

In our fourth and pre-registered (osf.io/u9ca2/) work (Fiedler*, Seiferth* et al., submitted), we found that if participants engaged with the just-in-time adaptive intervention, their step, and metabolic equivalent count were higher in the hour (and 90/120 minutes) after the trigger was answered compared to if they did not engage with the intervention (study design see Figure 2).

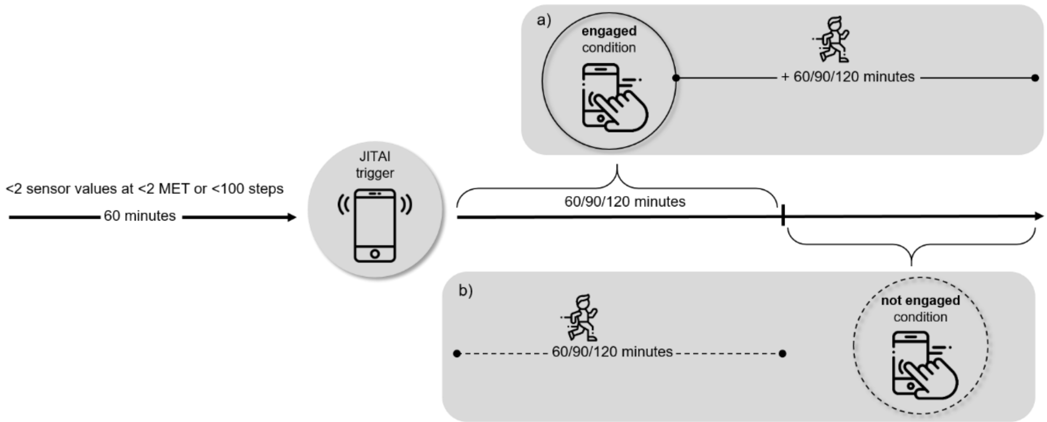

Figure 2. Illustration of 60/90/120-min time windows summarizing physical activity data (step and metabolic equivalent counts) when (a) the just-in-time adaptive intervention trigger was answered within the subsequent 60/90/120 minutes ("engaged" condition) or (b) when the just-in-time adaptive intervention trigger was not answered within this time window ("not engaged" condition). Abbreviations: *JITAI* just-in-time adaptive intervention, *MET* metabolic equivalents.

This finding extends evidence from previous studies which examined the impact of just-in-time adaptive interventions on accumulated physical activity outcomes on a daily or weekly level in small scale feasibility trials (Bond et al., 2014; Finkelstein et al., 2015; Pellegrini et al., 2015; Rabbi et al., 2015). By considering the timely association between sending the trigger and the behavioral response in the daily life of participants, our examination assesses the phenomena in the time period, and natural environment when they are supposed to occur (Dunton, 2017; Stone & Shiffman, 1994; Trull & Ebner-Priemer, 2013). This allowed us to draw better ecologically valid conclusions compared to daily or even weekly values regarding participants' engagement with the trigger. As the decision rule of this just-in-time adaptive intervention was only guided by the physical activity level of the participant in the previous hour, we also explored possible reasons for physical activity uptake descriptively. Here, our descriptive results indicated e.g. that if participants reported that they felt unwell in the previous hour, physical activity remained low in the next hour compared to other reasons. Future studies should further explore known factors which are associated with physical activity uptake or barriers thereof (Dunton, 2017), like contextual factors (Giurgiu et al., 2020) to enhance the individual tailoring of the interventions.

An additional important association with physical activity is the core affective state of participants (Liao et al., 2015; Trull & Ebner-Priemer, 2013). According to the three-dimensional model, core affect includes three intercorrelated affective dimensions which are expressed as bipolar items: valence (pleasure–displeasure), energetic arousal (wakefulness–tiredness), and calmness (relaxation–tension) (Russell, 2003; Schimmack & Grob, 2000). Here, previous research found valence and energetic arousal to be positively associated with physical activity while calmness was negatively associated with physical activity in adults and children (Forster et al., 2021; Koch et al., 2018). Additionally, sleep quality might be associated with physical activity uptake (Wang & Boros, 2021). As both sleep quality and core affect can be assessed using ecological momentary assessments (Stone & Shiffman, 1994; Trull & Ebner-Priemer, 2013), just-in-time adaptive interventions could be supported by the inclusion of these variables in the future. Therefore, we explored the association of sleep quality and core affective state of participants with daily steps and moderate to vigorous physical activity during the intervention period of the SMARTF*AMILY*2.0 study in our fifth publication (published: Fiedler et al., 2022)

Question 5: How are core affect and sleep quality related to physical activity during a mobile health intervention period?

In our fifth work (Fiedler et al., 2022), we found that above average ratings of a participant's daily valence and energetic arousal were associated with higher daily step count and more minutes of moderate to vigorous physical activity on the same day. Above average ratings of calmness meanwhile were associated with lower daily step count and minutes of moderate to vigorous physical activity. No clear association was found for sleep quality.

These results within our intervention study confirm previous ecological momentary assessment studies using core affect as time-lagged predictors (Cushing et al., 2017; Dunton et al., 2014; Liao et al., 2017; Niermann et al., 2016; Reichert et al., 2016; Schwerdtfeger et al., 2010), and on a daily level (Do et al., 2021). This underlines the important association of core affect with physical activity measures and the potential to include core affect in mobile health intervention designs. Concerning the association of sleep quality with physical activity, previous studies found inconsistent results (Antczak et al., 2020; Semplonius & Willoughby, 2018). The main challenge to examine this association are the various constructs used e.g. sleep quality, sleep duration, or sleep efficiency (Wang & Boros, 2021). Here, future studies might benefit from device-based measured sleep quality which is increasingly accessible by more convenient devices (Mendonca et al., 2019).

Our results highlight the importance to consider core affect in intervention designs due to its association with physical activity. Here, physical activity studies might benefit by adapting the intervention content to participants' core affect or by targeting core affect as a proxy to enhance physical activity if the relationship is based on causation. This could be achieved by e.g. enhancing the valence of the participant and thereby enhancing the physical activity during the next hour or even day. However, there will still be a long way as these promising approaches are still in the early stages of development and many causal associations are yet unclear.

References

Antczak, D., Lonsdale, C., Lee, J., Hilland, T., Duncan, M. J., Del Pozo Cruz, B., Hulteen, R. M., Parker, P. D., & Sanders, T. (2020). Physical activity and sleep are inconsistently related in healthy children: A systematic review and meta-analysis. *Sleep medicine reviews*, *51*, 101278. https://doi.org/10.1016/j.smrv.2020.101278

Bamman, M. M., Cooper, D. M., Booth, F. W., Chin, E. R., Neufer, P. D., Trappe, S., Lightfoot, J. T., Kraus, W. E., & Joyner, M. J. (2014). Exercise biology and medicine: innovative research to improve global health. *Mayo Clinic proceedings*, *89*(2), 148–153. https://doi.org/10.1016/j.mayocp.2013.11.013

Baumann, H., Fiedler, J., Wunsch, K., Woll, A., & Wollesen, B. (2022). mHealth Interventions to Reduce Physical Inactivity and Sedentary Behavior in Children and Adolescents: Systematic Review and Meta-analysis of Randomized Controlled Trials. *JMIR mHealth and uHealth*, *10*(5), e35920. https://doi.org/10.2196/35920

Beaglehole, R., Bonita, R., Horton, R., Adams, C., Alleyne, G., Asaria, P., Baugh, V., Bekedam, H., Billo, N., Casswell, S., Cecchini, M., Colagiuri, R., Colagiuri, S., Collins, T., Ebrahim, S., Engelgau, M., Galea, G., Gaziano, T., Geneau, R., . . . Watt, J. (2011). Priority actions for the non-communicable disease crisis. *The Lancet*, *377*(9775), 1438–1447. https://doi.org/10.1016/S0140-6736(11)60393-0

Bello, B., & Useh, U. (2021). COVID-19: Are Non-Communicable Diseases Risk Factors for Its Severity? *American journal of health promotion*, *35*(5), 720–729. https://doi.org/10.1177/0890117121990518

Blair, S. N. (2009). Physical inactivity: the biggest public health problem of the 21st century. *British journal of sports medicine*, *41*(1), 1–2.

Bond, D. S., Thomas, J. G., Raynor, H. A., Moon, J., Sieling, J., Trautvetter, J., Leblond, T., & Wing, R. R. (2014). B-MOBILE—a smartphone-based intervention to reduce sedentary time in overweight/obese individuals: a within-subjects experimental trial. *PloS one*, *9*(6), e100821. https://doi.org/10.1371/journal.pone.0100821

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C., Chaput, J.-P., Chastin, S., Chou, R., Dempsey, P. C., DiPietro, L., Ekelund, U., Firth, J., Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., . . . Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British journal of sports medicine*, *54*(24), 1451–1462. https://doi.org/10.1136/bjsports-2020-102955

Bull, F. C., Maslin, T. S., & Armstrong, T. (2009). Global physical activity questionnaire (GPAQ): nine country reliability and validity study. *Journal of physical activity & health*, *6*(6), 790–804. https://doi.org/10.1123/jpah.6.6.790

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

Church, T. S., Thomas, D. M., Tudor-Locke, C., Katzmarzyk, P. T., Earnest, C. P., Rodarte, R. Q., Martin, C. K., Blair, S. N., & Bouchard, C. (2011). Trends over 5 decades in U.S. occupation-related physical activity and their associations with obesity. *PloS one*, *6*(5), e19657. https://doi.org/10.1371/journal.pone.0019657

Conroy, D. E., Lagoa, C. M., Hekler, E. B., & Rivera, D. E. (2020). Engineering Person-Specific Behavioral Interventions to Promote Physical Activity. *Exercise and sport sciences reviews*, *48*(4), 170–179. https://doi.org/10.1249/JES.0000000000000232

Cushing, C. C., Mitchell, T. B., Bejarano, C. M., Walters, R. W., Crick, C. J., & Noser, A. E. (2017). Bidirectional Associations Between Psychological States and Physical Activity in Adolescents: A mHealth Pilot Study. *Journal of pediatric psychology*, *42*(5), 559–568. https://doi.org/10.1093/jpepsy/jsw099

Ding, D., Lawson, K. D., Kolbe-Alexander, T. L., Finkelstein, E. A., Katzmarzyk, P. T., van Mechelen, W., & Pratt, M. (2016). The economic burden of physical inactivity: a global analysis of major non-communicable diseases. *The Lancet*, *388*(10051), 1311–1324. https://doi.org/10.1016/S0140-6736(16)30383-X

Direito, A., Carraca, E., Rawstorn, J., Whittaker, R., & Maddison, R. (2017). mHealth Technologies to Influence Physical Activity and Sedentary Behaviors: Behavior Change Techniques, Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Annals of Behavioral Medicine*, *51*(2), 226–239. https://doi.org/10.1007/s12160-016-9846-0

Direito, A., Dale, L. P., Shields, E., Dobson, R., Whittaker, R., & Maddison, R. (2014). Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? *BMC public health*, *14*, 646. https://doi.org/10.1186/1471-2458-14-646

Do, B., Wang, S. D., Courtney, J. B., & Dunton, G. F. (2021). Examining the day-level impact of physical activity on affect during the early months of the COVID-19 pandemic: An ecological momentary assessment study. *Psychology of Sport and Exercise*, *56*. https://doi.org/10.1016/j.psychsport.2021.102010

Dunton, G. F. (2017). Ecological Momentary Assessment in Physical Activity Research. *Exercise and sport sciences reviews*, *45*(1), 48–54. https://doi.org/10.1249/JES.0000000000000092

Dunton, G. F., Huh, J., Leventhal, A. M., Riggs, N., Hedeker, D., Spruijt-Metz, D., & Pentz, M. A. (2014). Momentary assessment of affect, physical feeling states, and physical activity in children. *Health psychology: official journal of the Division of Health Psychology, American Psychological Association*, *33*(3), 255–263. https://doi.org/10.1037/a0032640

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

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., 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

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 (Basel, Switzerland)*, *21*(8). https://doi.org/10.3390/s21082672

Fiedler, J., Eckert, T., Wunsch, K., & Woll, A. (2020). Key facets to build up eHealth and mHealth interventions to enhance physical activity, sedentary behavior and nutrition in healthy subjects - an umbrella review. *BMC public health*, *20*(1), 1605.

Fiedler, J., Seifert, C., Eckert, T., Woll, A., & Wunsch, K. (submitted). A just-in-time adaptive intervention to enhance physical activity in the SMARTFAMILY2.0 trial. *Sport, Exercise, and Performance Psychology*.

Fiedler, J., Seiferth, C., Eckert, T., Woll, A., & Wunsch, K. (2022). Sleep quality, valence, energetic arousal, and calmness as predictors of device-based measured physical activity during a three-week mHealth intervention. *German Journal of Exercise and Sport Research.* Vorab-Onlinepublikation. https://doi.org/10.1007/s12662-022-00809-y

Finkelstein, J., McKenzie, B., Xuan, L., Jeffrey, W., & Pamela, O. (Hrsg.) (2015). *Mobile App to Reduce Inactivity in Sedentary Overweight Women*. MedInfo 2015: eHealth-enabled.

Forouzanfar, M. H., Afshin, A., Alexander, L. T., Anderson, H. R., Bhutta, Z. A., Biryukov, S., Brauer, M., Burnett, R., Cercy, K., Charlson, F. J., Cohen, A. J., Dandona, L., Estep, K., Ferrari, A. J., Frostad, J. J., Fullman, N., Gething, P. W., Godwin, W. W., Griswold, M., . . . Murray, C. J. L. (2016). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, *388*(10053), 1659–1724. https://doi.org/10.1016/S0140-6736(16)31679-8

Forster, A. K., Richards, E. A., Foli, K. J., McGowan, B., Hass, Z., Becker, M., & Miller, A. (2021). Influence of Affect on Physical Activity: An Integrative Review. *Clinical nursing research*, *30*(7), 934–949. https://doi.org/10.1177/1054773820968039

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

Gonul, S., Namli, T., Huisman, S., Laleci Erturkmen, G. B., Toroslu, I. H., & Cosar, A. (2019). An expandable approach for design and personalization of digital, just-in-time adaptive interventions. *Journal of the American Medical Informatics Association: JAMIA*, *26*(3), 198–210. https://doi.org/10.1093/jamia/ocy160

González, K., Fuentes, J., & Márquez, J. L. (2017). Physical Inactivity, Sedentary Behavior and Chronic Diseases. *Korean journal of family medicine*, *38*(3), 111–115. https://doi.org/10.4082/kjfm.2017.38.3.111

Greaves, C. J., Sheppard, K. E., Abraham, C., Hardeman, W., Roden, M., Evans, P. H., & Schwarz, P. (2011). Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. *BMC public health*, *11*(1), 1–12. https://doi.org/10.1186/1471-2458-11-119

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

Hagstromer, M., Ainsworth, B. E., Oja, P., & Sjostrom, M. (2010). Comparison of a subjective and an objective measure of physical activity in a population sample. *Journal of physical activity & health*, *7*(4), 541–550. https://doi.org/10.1123/jpah.7.4.541

Hardeman, W., Houghton, J., Lane, K., Jones, A., & Naughton, F. (2019). A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. *The international journal of behavioral nutrition and physical activity*, *16*(1), 31. https://doi.org/10.1186/s12966-019-0792-7

Heath, G. W., Parra, D. C., Sarmiento, O. L., Andersen, L. B., Owen, N., Goenka, S., Montes, F., & Brownson, R. C. (2012). Evidence-based intervention in physical activity: lessons from around the world. *The Lancet*, *380*(9838), 272–281. https://doi.org/10.1016/S0140-6736(12)60816-2

Heron, K. E., & Smyth, J. M. (2010). Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. *British journal of health psychology*, *15*(Pt 1), 1–39. https://doi.org/10.1348/135910709X466063

Herrmann, S. D., Heumann, K. J., Der Ananian, C. A., & Ainsworth, B. E. (2013). Validity and Reliability of the Global Physical Activity Questionnaire (GPAQ). *Measurement in Physical Education and Exercise Science*, *17*(3), 221–235. https://doi.org/10.1080/1091367X.2013.805139

Hukkanen, H., Husu, P., Sievänen, H., Tokola, K., Vähä-Ypyä, H., Valkeinen, H., Mäki-Opas, T., Suni, J. H., & Vasankari, T. (2018). Aerobic physical activity assessed with accelerometer, diary, questionnaire, and interview in a Finnish population sample. *Scandinavian journal of medicine & science in sports*, *28*(10), 2196–2206. https://doi.org/10.1111/sms.13244

Kahn, E. B., Ramsey, L. T., Brownson, R. C., Heath, G. W., Howze, E. H., Powell, K. E., Stone, E. J., Rajab, M. W., Corso, P., & Task Force on Community Preventive Services (2002). The Effectiveness of Interventions to Increase Physical Activity: A Systematic Review. *American journal of preventive medicine*, *22*(4), 73–107.

Katzmarzyk, P. T., Friedenreich, C., Shiroma, E. J., & Lee, I.-M. (2022). Physical inactivity and non-communicable disease burden in low-income, middle-income and high-income countries. *British journal of sports medicine*, *56*(2), 101–106. https://doi.org/10.1136/bjsports-2020-103640

Keadle, S. K., Lyden, K. A., Strath, S. J., Staudenmayer, J. W., & Freedson, P. S. (2019). A Framework to Evaluate Devices That Assess Physical Behavior. *Exercise and sport sciences reviews*, *47*(4), 206–214. https://doi.org/10.1249/JES.0000000000000206

Koch, E. D, Tost, H., Braun, U., Gan, G., Giurgiu, M., Reinhard, I., Zipf, A., Meyer-Lindenberg, A., Ebner-Priemer, U. W., & Reichert, M. (2018). Mood Dimensions Show Distinct Within-Subject Associations With Non-exercise Activity in Adolescents: An Ambulatory Assessment Study. *Frontiers in psychology*, *9*, 268. https://doi.org/10.3389/fpsyg.2018.00268

Kwasnicka, D., Keller, J., Perski, O., Potthoff, S., Hoor, G. A. ten, Ainsworth, B., Crutzen, R., Dohle, S., van Dongen, A., Heino, M., Henrich, J. F., Knox, L., König, L. M., Maltinsky, W., McCallum, C., Nalukwago, J., Neter, E., Nurmi, J., Spitschan, M., . . . Sanderman, R. (2022). White Paper: Open Digital Health - accelerating transparent and scalable health promotion and treatment. *Health psychology review*, 1–17. https://doi.org/10.1080/17437199.2022.2046482

Liao, Y., Chou, C.-P., Huh, J., Leventhal, A., & Dunton, G. F. (2017). Examining acute bi-directional relationships between affect, physical feeling states, and physical activity in free-living situations using electronic ecological momentary assessment. *Journal of behavioral medicine*, *40*(3), 445–457. https://doi.org/10.1007/s10865-016-9808-9

Liao, Y., Shonkoff, E. T., & Dunton, G. F. (2015). The Acute Relationships Between Affect, Physical Feeling States, and Physical Activity in Daily Life: A Review of Current Evidence. *Frontiers in psychology*, *6*, 1975. https://doi.org/10.3389/fpsyg.2015.01975

Livingstone, M. B. E., Robson, P. J., Wallace, J. M. W., & McKinley, M. C. (2003). How active are we? Levels of routine physical activity in children and adults. *The Proceedings of the Nutrition Society*, *62*(3), 681–701. https://doi.org/10.1079/PNS2003291

Martinez-Gomez, D., Cabanas-Sanchez, V., Yu, T., Rodriguez-Artalejo, F., Ding, D., Lee, I.-M., & Ekelund, U. (2022). Long-term leisure-time physical activity and risk of all-cause and cardiovascular mortality: dose-response associations in a prospective cohort study of 210 327 Taiwanese adults. *British Journal of Sports Medicine*. Vorab-Onlinepublikation. https://doi.org/10.1136/bjsports-2021-104961

Mclaughlin, M., Delaney, T., Hall, A., Byaruhanga, J., Mackie, P., Grady, A., Reilly, K., Campbell, E., Sutherland, R., Wiggers, J., & Wolfenden, L. (2021). Associations Between Digital Health Intervention Engagement, Physical Activity, and Sedentary Behavior: Systematic Review and Meta-analysis. *Journal of medical Internet research*, *23*(2), e23180. https://doi.org/10.2196/23180

Mendonca, F., Mostafa, S. S., Morgado-Dias, F., Ravelo-Garcia, A. G., & Penzel, T. (2019). A Review of Approaches for Sleep Quality Analysis. *IEEE Access*, *7*, 24527–24546. https://doi.org/10.1109/ACCESS.2019.2900345

Mercer, K., Li, M., Giangregorio, L., Burns, C., & Grindrod, K. (2016). Behavior Change Techniques Present in Wearable Activity Trackers: A Critical Analysis. *JMIR mHealth and uHealth*, *4*(2), e40. https://doi.org/10.2196/mhealth.4461

Michie, S., Ashford, S., Sniehotta, F. F., Dombrowski, S. U., Bishop, A., & French, D. P. (2011). A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. *Psychology & health*, *26*(11), 1479–1498. https://doi.org/10.1080/08870446.2010.540664

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M. P., Cane, J., & Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Annals of behavioral medicine: a publication of the Society of Behavioral Medicine*, *46*(1), 81–95. https://doi.org/10.1007/s12160-013-9486-6

Michie, S., Yardley, L., West, R., Patrick, K., & Greaves, F. (2017). Developing and Evaluating Digital Interventions to Promote Behavior Change in Health and Health Care: Recommendations Resulting From an International Workshop. *Journal of medical Internet research*, *19*(6), e232. https://doi.org/10.2196/jmir.7126

Miller, C. K. (2019). Adaptive Intervention Designs to Promote Behavioral Change in Adults: What Is the Evidence? *Current diabetes reports*, *19*(2), 7. https://doi.org/10.1007/s11892-019-1127-4

Morrison, L. G., Yardley, L., Powell, J., & Michie, S. (2012). What design features are used in effective e-health interventions? A review using techniques from Critical Interpretive Synthesis. *Telemedicine journal and e-health: the official journal of the American Telemedicine Association*, *18*(2), 137–144. https://doi.org/10.1089/tmj.2011.0062

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Annals of Behavioral Medicine*, *52*(6), 446–462. https://doi.org/10.1007/s12160-016-9830-8

Naszay, M., Stockinger, A., Jungwirth, D., & Haluza, D. (2018). Digital age and the Public eHealth perspective: Prevailing health app use among Austrian Internet users. *Informatics for health & social care*, *43*(4), 390–400. https://doi.org/10.1080/17538157.2017.1399131

Neufer, P. D., Bamman, M. M., Muoio, D. M., Bouchard, C., Cooper, D. M., Goodpaster, B. H., Booth, F. W., Kohrt, W. M., Gerszten, R. E., Mattson, M. P., Hepple, R. T., Kraus, W. E., Reid, M. B., Bodine, S. C., Jakicic, J. M., Fleg, J. L., Williams, J. P., Joseph, L., Evans, M., . . . Laughlin, M. R. (2015). Understanding the Cellular and Molecular Mechanisms of Physical Activity-Induced Health Benefits. *Cell metabolism*, *22*(1), 4–11. https://doi.org/10.1016/j.cmet.2015.05.011

Niermann, C., Herrmann, C., Haaren, B. von, van Kann, D., & Woll, A. (2016). Affect and Subsequent Physical Activity: An Ambulatory Assessment Study Examining the Affect-Activity Association in a Real-Life Context. *Frontiers in psychology*, *7*, 677. https://doi.org/10.3389/fpsyg.2016.00677

Orme, M., Wijndaele, K., Sharp, S. J., Westgate, K., Ekelund, U., & Brage, S. (2014). Combined influence of epoch length, cut-point and bout duration on accelerometry-derived physical activity.

The international journal of behavioral nutrition and physical activity, *11*(1), 34. https://doi.org/10.1186/1479-5868-11-34

Paluch, A. E., Bajpai, S., Bassett, D. R., Carnethon, M. R., Ekelund, U., Evenson, K. R., Galuska, D. A., Jefferis, B. J., Kraus, W. E., Lee, I.-M., Matthews, C. E., Omura, J. D., Patel, A. V., Pieper, C. F., Rees-Punia, E., Dallmeier, D., Klenk, J., Whincup, P. H., Dooley, E. E., . . . Fulton, J. E. (2022). Daily steps and all-cause mortality: a meta-analysis of 15 international cohorts. *The Lancet Public Health*, *7*(3), e219-e228. https://doi.org/10.1016/S2468-2667(21)00302-9

Pan, X.-F., Yang, J., Wen, Y., Li, N., Chen, S., & an Pan (2021). Non-Communicable Diseases During the COVID-19 Pandemic and Beyond. *Engineering (Beijing, China)*, *7*(7), 899–902. https://doi.org/10.1016/j.eng.2021.02.013

Pellegrini, C. A., Hoffman, S. A., Daly, E. R., Murillo, M., Iakovlev, G., & Spring, B. (2015). Acceptability of smartphone technology to interrupt sedentary time in adults with diabetes. *Translational behavioral medicine*, *5*(3), 307–314. https://doi.org/10.1007/s13142-015-0314-3

Pfeifer, K., & Rütten, A. (2017). Nationale Empfehlungen für Bewegung und Bewegungsförderung [National Recommendations for Physical Activity and Physical Activity Promotion]. *Gesundheitswesen (Bundesverband der Arzte des Offentlichen Gesundheitsdienstes (Germany))*, *79*(S 01), S2-S3. https://doi.org/10.1055/s-0042-123346

Prestwich, A., Sniehotta, F. F., Whittington, C., Dombrowski, S. U., Rogers, L., & Michie, S. (2014). Does theory influence the effectiveness of health behavior interventions? Meta-analysis. *Health psychology: official journal of the Division of Health Psychology, American Psychological Association*, *33*(5), 465–474. https://doi.org/10.1037/a0032853

Rabbi, M., Aung, M. H., Zhang, M., & Choudhury, T. (2015). MyBehavior. In K. Mase, M. Langheinrich, D. Gatica-Perez, H. Gellersen, T. Choudhury & K. Yatani (Hrsg.), *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15* (S. 707-718). ACM Press. https://doi.org/10.1145/2750858.2805840

Reichert, M., Giurgiu, M., Koch, E. D, 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*, 101742. https://doi.org/10.1016/j.psychsport.2020.101742

Reichert, M., Tost, H., Reinhard, I., Zipf, A., Salize, H.-J., Meyer-Lindenberg, A., & Ebner-Priemer, U. W. (2016). Within-Subject Associations between Mood Dimensions and Non-exercise Activity: An Ambulatory Assessment Approach Using Repeated Real-Time and Objective Data. *Frontiers in psychology*, *7*, 918. https://doi.org/10.3389/fpsyg.2016.00918

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological review*, *110*(1), 145–172. https://doi.org/10.1037/0033-295x.110.1.145

Ryan, R., & Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, *55*(1), 68.

Schembre, S. M., Liao, Y., Robertson, M. C., Dunton, G. F., Kerr, J., Haffey, M. E., Burnett, T., Basen-Engquist, K., & Hicklen, R. S. (2018). Just-in-Time Feedback in Diet and Physical Activity Interventions: Systematic Review and Practical Design Framework. *Journal of medical Internet research*, *20*(3), e106. https://doi.org/10.2196/jmir.8701

Schimmack, U., & Grob, A. (2000). Dimensional models of core affect: a quantitative comparison by means of structural equation modeling. *European Journal of Personality*, *14*(4), 325–345. https://doi.org/10.1002/1099-0984(200007/08)14:4<325::AID-PER380>3.3.CO;2-9

Schoeppe, S., Alley, S., van Lippevelde, W., Bray, N. A., Williams, S. L., Duncan, M. J., & Vandelanotte, C. (2016). Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. *The international journal of behavioral nutrition and physical activity*, *13*. https://doi.org/10.1186/s12966-016-0454-y

Schwerdtfeger, A., Eberhardt, R., Chmitorz, A., & Schaller, E. (2010). Momentary Affect Predicts Bodily Movement in Daily Life: An Ambulatory Monitoring Study. *Journal of Sport and Exercise Psychology*, *32*(5), 674–693.

Semplonius, T., & Willoughby, T. (2018). Long-Term Links between Physical Activity and Sleep Quality. *Medicine and science in sports and exercise*, *50*(12), 2418–2424. https://doi.org/10.1249/MSS.0000000000001706

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

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

Statista. (2021). *Number of smartphone users from 2016 to 2021*. Abruf unter https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

Stockwell, S., Trott, M., Tully, M., Shin, J., Barnett, Y., Butler, L., McDermott, D., Schuch, F., & Smith, L. (2021). Changes in physical activity and sedentary behaviours from before to during the

COVID-19 pandemic lockdown: a systematic review. *BMJ open sport & exercise medicine*, *7*(1), e000960. https://doi.org/10.1136/bmjsem-2020-000960

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine*, *16*(3), 199–202. https://doi.org/10.1093/abm/16.3.199

Thomas, J. G., & Bond, D. S. (2015). Behavioral response to a just-in-time adaptive intervention (JITAI) to reduce sedentary behavior in obese adults: Implications for JITAI optimization. *Health psychology: official journal of the Division of Health Psychology, American Psychological Association*, *34S*, 1261–1267. https://doi.org/10.1037/hea0000304

Tremblay, Mark S., LeBlanc, A. G., Kho, Michelle E., Saunders, Travis J., Larouche, R., Colley, Rachel, C., Goldfield, G., & Gorber, S. C. (2011). Systematic review of sedentary behaviour and health indicators in school-aged children and youth. *The international journal of behavioral nutrition and physical activity*, *8*(98). https://doi.org/10.1186/1479-5868-8-98

Trull, T. J., & Ebner-Priemer, U. W. (2013). Ambulatory assessment. *Annual Review of Clinical Psychology*, *9*, 151–176. https://doi.org/10.1146/annurev-clinpsy050212-185510

Umberson, D., Crosnoe, R., & Reczek, C. (2010). Social Relationships and Health Behavior Across Life Course. *Annual review of sociology*, *36*, 139–157. https://doi.org/10.1146/annurev-soc-070308-120011

Vandelanotte, C., Müller, A. M., Short, C. E., Hingle, M., Nathan, N., Williams, S. L., Lopez, M. L., Parekh, S., & Maher, C. A. (2016). Past, Present, and Future of eHealth and mHealth Research to Improve Physical Activity and Dietary Behaviors. *Journal of Nutrition Education and Behavior*, *48*(3), 219-228. https://doi.org/10.1016/j.jneb.2015.12.006

Vandelanotte, C., Spathonis, K. M., Eakin, E. G., & Owen, N. (2007). Website-delivered physical activity interventions a review of the literature. *American journal of preventive medicine*, *33*(1), 54–64. https://doi.org/10.1016/j.amepre.2007.02.041

Wang, F., & Boros, S. (2021). The effect of physical activity on sleep quality: a systematic review. *European Journal of Physiotherapy*, *23*(1), 11–18. https://doi.org/10.1080/21679169.2019.1623314

Warburton, D. E. R., & Bredin, S. S. D. (2017). Health benefits of physical activity: a systematic review of current systematic reviews. *Current opinion in cardiology*, *32*(5), 541–556. https://doi.org/10.1097/HCO.0000000000000437

Warburton, D. E. R., Nicol, C. W., & Bredin, S. S. D. (2006). Health benefits of physical activity: the evidence. *CMAJ: Canadian Medical Association journal = journal de l'Association medicale canadienne*, *174*(6), 801–809. https://doi.org/10.1503/cmaj.051351

Webb, T. L., Joseph, J., Yardley, L., & Michie, S. (2010). Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *Journal of medical Internet research*, *12*(1), e4. https://doi.org/10.2196/jmir.1376

Woll, A., Kurth, B.-M., Opper, E., Worth, A., & Bös, K. (2011). The 'Motorik-Modul' (MoMo): physical fitness and physical activity in German children and adolescents. *European journal of pediatrics*, *170*(9), 1129–1142. https://doi.org/10.1007/s00431-010-1391-4

World Health Organization. (2011). *mHealth: Second Global Survey on eHealth*. World Health Organization. Abruf unter http://gbv.eblib.com/patron/FullRecord.aspx?p=851142

World Health Organization. (2014). *Global Status Report on Noncommunicable Diseases 2014*. World Health Organization. Abruf unter http://gbv.eblib.com/patron/FullRecord.aspx?p=2059287

World Health Organization. (2018). *More active people for a healthier world: Global action plan on physical activity 2018-2030. Let`s be active*. World Health Organization.

World Health Organization. (19. März 2020). *eHealth at WHO*. Abruf unter https://www.who.int/ehealth/en/

Wunsch, K., Eckert, T., Fiedler, J., Cleven, L., Niermann, C., Reiterer, H., Renner, B., & Woll, A. (2020). Effects of a Collective Family-Based Mobile Health Intervention Called "SMARTFAMILY" on Promoting Physical Activity and Healthy Eating: Protocol for a Randomized Controlled Trial. *JMIR Research Protocols*, *9*(11), e20534. https://doi.org/10.2196/20534

Wunsch, K., Eckert, T., Fiedler, J., & Woll, A. (2022). Just-in-time adaptive interventions in mobile physical activity interventions - A synthesis of frameworks and future directions. *The European Health Psychologist*, *22*(4), 834–842.

Wunsch, K., Kienberger, K., & Niessner, C. (2022). Changes in Physical Activity Patterns Due to the Covid-19 Pandemic: A Systematic Review and Meta-Analysis. *International journal of environmental research and public health*, *19*(4). https://doi.org/10.3390/ijerph19042250

Yang, C.-H., Maher, J. P., & Conroy, D. E. (2015). Implementation of behavior change techniques in mobile applications for physical activity. *American journal of preventive medicine*, *48*(4), 452–455. https://doi.org/10.1016/j.amepre.2014.10.010

Chapter 2 Key facets of digital health interventions

Paper 1: Key facets to build up eHealth and mHealth interventions to enhance physical activity, sedentary behavior and nutrition in healthy subjects.

Slightly modified version of the published manuscript

Fiedler, J., Eckert, T., Wunsch, K., & Woll, A. (2020). Key facets to build up eHealth and mHealth interventions to enhance physical activity, sedentary behavior and nutrition in healthy subjects - an umbrella review. *BMC public health*, *20*(1), 1605.
https://doi.org/10.1186/s12889-020-09700-7

Abstract

Electronic (eHealth) and mobile (mHealth) health interventions can provide a large coverage, and are promising tools to change health behavior (i.e. physical activity, sedentary behavior, and healthy eating). However, the determinants of intervention effectiveness in primary prevention have not been explored yet. Therefore, the objectives of this umbrella review were to evaluate intervention effectiveness, to explore the impact of pre-defined determinants of effectiveness (i.e. theoretical foundations, behavior change techniques, social contexts, or just-in-time adaptive interventions), and to provide recommendations for future research and practice in the field of primary prevention delivered via e/mHealth technology.

PubMed, Scopus, Web of Science, and the Cochrane Library were searched for systematic reviews and meta-analyses (reviews) published between January 1990 and May 2020. Reviews reporting on e/mHealth behavior change interventions in physical activity, sedentary behavior, and/or healthy eating for healthy subjects (i.e. subjects without physical or physiological morbidities which would influence the realization of behaviors targeted by the respective interventions) were included if they also investigated respective theoretical foundations, behavior change techniques, social contexts or just-in-time adaptive interventions. Included studies were ranked concerning their methodological quality and qualitatively synthesized.

The systematic search revealed eleven systematic reviews and meta-analyses of moderate quality. The majority of original research studies within the reviews found e/mHealth interventions to be effective, but the results showed a high heterogeneity concerning assessment methods

and outcomes, making them difficult to compare. Whereas theoretical foundation and behavior change techniques were suggested to be potential positive determinants of effective interventions, the impact of social context remains unclear. None of the reviews included just-in-time adaptive interventions.

Findings of this umbrella review support the use of e/mHealth to enhance physical activity and healthy eating and reduce sedentary behavior. The general lack of precise reporting and comparison of confounding variables in reviews and original research studies as well as the limited number of reviews for each health behavior constrains the generalization and interpretation of results. Further research is needed on study-level to investigate effects of versatile determinants of e/mHealth efficiency, using a theoretical foundation and additionally explore the impact of social contexts and more sophisticated approaches like just-in-time adaptive interventions.

Trial registration

The protocol for this umbrella review was a priori registered with PROSPERO: CRD42020147902.

Keywords

telemedicine; health behavior; primary prevention; exercise; sedentary behavior; food and nutrition; umbrella review; psychology social; just-in-time adaptive intervention; psychological theory

Background

Physical activity (PA), a reduction of sedentary behavior (SB) and healthy eating (i.e. enhanced fruit and vegetable intake (FVI), reduced sugar and saturated fat intake among others) (HE) are key strategies in the primary prevention of noncommunicable diseases like cardiovascular diseases, diabetes, cancer and obesity, which were responsible for 41 million deaths worldwide in 2016 (World Health Organization, 2018). Despite this knowledge, the levels of PA and HE are often insufficient in our modern society throughout all age groups (Aune et al., 2017; Blair, 2009; Moore & Thompson, 2015; Nielsen et al., 2014; Woll et al., 2011), while SB, such as excessive sitting during worktime (e.g. deskwork) and during leisure time (e.g. watching television), increased over the past years (Owen et al., 2010; Tremblay et al., 2011). As a result, guidelines concerning PA, SB and HE are put into place, but the sole presence of these recommendations is not sufficient to change health behavior and to reduce the financial and health burden worldwide (World Health Organization, 2014). Working towards achieving these guidelines is important throughout all stages of life and can be seen as a long-term investment which seems to be easier to achieve for healthy people since obesity or other morbidities add further barriers which restrict engagement in healthy behaviors (Baird et al., 2017). Focusing on primary prevention in healthy participants can therefore be a sustainable way to reduce the prevalence of noncommunicable diseases. One promising strategy for primary care prevention might be the usage of electronic (eHealth) and mobile (mHealth) health interventions. eHealth interventions comprise “the use of information and communication technologies for health” (World Health Organization, 2020), while mHealth interventions refer to “medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices” (World Health Organization, 2011). With 4.5 billion active internet users in 2020 worldwide (*Global Digital Population 2020 | Statista*, 2020, April), the potential coverage of e/mHealth tools coupled with intuitive and autonomous control of the device by the end user hold great promise. This is especially true for the younger and digital native generations who are known to interact frequently with e/mHealth (Naszay et al., 2018). For the establishment of e/mHealth in primary prevention, several methodological issues such as the need for accurate and validated measuring tools for a better comparison of different e/mHealth approaches and dose/response relationship for interventions require further investigation (McClung et al., 2018).

Theoretical foundation of interventions, as depicted by behavior change theories (e.g. self-determination theory (Ryan & Deci, Edward, L., 2000), theory of planned behavior (Ajzen, 1991), transtheoretical model (Prochaska, James O., and Wayne, F. Velicer, 1997) or social cognitive theory (Bandura, 1989)), and by behavior change techniques (BCTs) (Michie et al., 2011; Michie et al., 2013) were shown to be important facets for intervention effectiveness (Prestwich et al., 2014; Webb et al., 2010). Additionally, health behaviors are usually linked to social contexts and affected by social relations (Umberson et al., 2010). Thus, facets like information about and interacting with other users or peers (Morrison et al., 2012) might also have an important impact on intervention effectiveness and might help to sustain successful behavior change (Glanz et al., 2008). This has been especially true for adolescents as their sufficient level of PA, SB and HE strongly depends on their families, schools, and peers (Viner et al., 2012). Therefore, the integration and documentation of social contexts are important to assess the influence on and enhance the effectiveness of sustainable health behavior change. Furthermore, individual tailoring based on theoretical constructs was shown to be positively associated with effective interventions (Morrison et al., 2012). Delivering these interventions during the most promising time for the desired behavior (e.g. PA and HE) or during the most vulnerable time for unhealthy behavior (e.g. SB), implementation of the so called just-in-time (adaptive) intervention (JITAI) (Hardeman et al., 2019; Schembre et al., 2018) and ecological momentary intervention (EMI) (Heron & Smyth, 2010) are promising new approaches for effective e/mHealth interventions. With the development of new generations of a variety of sensors (Schembre et al., 2018) and the integration of machine learning approaches (Gonul et al., 2019a), the advances in individual tailoring are rapidly evolving and appear to be auspicious facets to implement in behavior change interventions.

Existing umbrella reviews concerning mHealth in general revealed only limited evidence to be effective to change a variety of behaviors (Marcolino et al., 2018), while the use of text messages has shown effectiveness for several health outcomes (Hall et al., 2015). There is an abundance of mHealth interventions for diabetes which led to clinically relevant improvements (Hood et al., 2016; Kitsiou et al., 2017). Existing umbrella reviews in the area of digital behavior change interventions expressed the need to examine the key contents of effective interventions in different settings (e.g. home, work, or school based interventions) (Bertoncello et al., 2018), and to consider various facets for an effective implementation (Ross et al., 2016). An overview of efficient intervention components has only been composed for non-e/mHealth interventions promoting PA, SB, and HE (Biddle et al., 2014; Brand et al., 2014; dos Santos et al., 2019; Greaves et al., 2011). Key determinants of effectiveness in these overviews were the use of theoretical foundations (Biddle et al., 2014; Brand et al., 2014; Greaves et al., 2011), BCTs (Greaves et al., 2011), social contexts (Biddle et al., 2014; Brand et al., 2014; dos Santos et al., 2019; Greaves et al., 2011) and using prompts and feedback (Biddle et al., 2014; Brand et al., 2014; Greaves et al., 2011). Taken together, there is a research gap for e/mHealth interventions concerning facets of effectiveness with a focus on health behavior change in primary prevention.

In order to determine if these facets (i.e. theoretical foundations, BCTs, social contexts, JITAIs) were incorporated in recent e/mHealth interventions of primary prevention and with which magnitude they contributed to intervention effectiveness (in addition to methodological facets), a systematic summary of research by conducting an umbrella review (Fusar-Poli & Radua, 2018) is needed.

Methods

This umbrella review was registered a priori with PROSPERO (International prospective register of systematic reviews, registration number CRD42020147902). It was conducted based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher et al., 2009).

Study aim

The present umbrella review aimed to systematically summarize the results from systematic reviews and meta-analyses concerning the effectiveness of e/mHealth interventions to promote PA, reduce SB and promote HE as a primary care strategy in healthy participants. Further, the umbrella review aims to identify the impact of theoretical foundations, BCTs, social contexts, and JITAIs on the effectiveness of e/mHealth interventions. Moreover, the recommendations for future research provided by the included reviews were analyzed and expanded to provide an overview of needs to be addressed in future developments of e/mHealth interventions.

Data Sources and Search Strategy

A systematic search for reviews published in English between 01.01.1990 and 16.08.2019 was conducted using the four databases PubMed, Scopus, Web of Science, and the Cochrane Library for systematic reviews and meta-analyses. The search was conducted by one author and repeated prior to submission on 20.05.2020 (JF). The search terms were reviewed by two authors (JF, KW) and included the following key constructs as well as numerous synonyms thereof: (eHealth OR mHealth) AND (PA OR SB OR HE) AND (theoretical foundation/BCT OR social OR JITAI/EMI). Additionally, a forward- / backward-search was conducted on the reference lists of included reviews. Please see additional file 1 for detailed search strategy of all databases.

Review selection

Following the systematic search, literature was imported to the reference management software *CITAVI* 6. After duplicates were removed, two reviewers (JF, KW) independently examined titles and abstracts. Full texts of relevant review articles were obtained and assessed based on the inclusion and exclusion criteria described below (JF, TE). Reasons for exclusion at this stage were recorded and are displayed in the PRISMA-flow chart (Figure 1). Any disagreements between authors were resolved by consensus and/or discussion with a third author (KW or TE).

Inclusion and exclusion criteria

Inclusion and exclusion criteria were selected based on PICOS (1. Population, 2. Intervention, 3. Comparison, 4. Outcome, and 5. Study type) (Moher et al., 2009).

1. Population inclusion: Healthy participants of all ages with no physical or physiological morbidities including obesity (BMI > 30kg/m²) which would influence the realization of behaviors (i.e. PA, SB, and HE) targeted by the respective interventions. If a review included patient groups or participants with any physical or psychological morbidities and provided a subgroup analysis or reported the results for the healthy population separately, the review was also included. Exclusion: Participants with any physical or psychological morbidities including obesity (BMI > 30kg/m²), clinical settings, and studies focusing on populations, whose PA, SB, or HE was influenced by disease specific recommendations or health status.

2. Intervention inclusion: e/mHealth interventions where the primary outcome measure was PA (e.g. steps, moderate, vigorous, or moderate to vigorous (MVPA)) and/or SB (e.g. sitting time, screen time) and/or HE (e.g. FVI, fat consumption) were selected. Exclusion: Studies without an intervention, with no e/mHealth interventions, with mixed interventions if e/mHealth were not analyzed separately.

3. Comparison: Included reviews were not limited to comparator studies.

4. Outcome inclusion: Effectiveness for PA and/or SB and/or HE as the main outcome. Effectiveness had to be displayed or discussed with regard to at least one of the following aspects: a) theoretical foundation or BCTs, b) JITAI/EMI, and c) social context (e.g. social network, family/peer group/school setting). Exclusion: Studies that focused on other health outcomes like weight loss, quality of life, or had multiple additional health behaviors not related to PA, SB, or HE (e.g. smoking, drinking) as main outcomes. Studies without discussion/results for any of the following: theoretical foundation or BCTs, JITAI/EMI, or social context.

5. Study type inclusion: Systematic reviews and meta-analyses based on the PRISMA statement. Exclusion: Non-systematic reviews (e.g. narrative reviews, qualitative reviews, scoping reviews).

Study quality assessment

Review quality was rated independently by one author (JF) and a research assistant using the assessment of multiple systematic reviews (AMSTAR) tool (Shea et al., 2007). Any disagreements were resolved by discussion until a common consent was found.

Data extraction

Data was extracted from the included studies by JF using a predefined Excel sheet. Data extraction and coding were checked by a research assistant. Any disagreements were resolved by discussion until a common consent was found. The following data were extracted: author and year, type of review, aim, mHealth/eHealth tools, country (where the included studies were conducted), main outcomes (constructs and parameters), time period searched and time period of included studies, included study designs, number of studies, number, and age of participants, intervention duration, quality of included studies indicated by the reviews, included theory/BCT, included social/JITAI/EMI, reported effectiveness, recommendations for future research as stated by the authors.

Analysis

We used the term review to describe systematic reviews and meta-analysis together, and distinguished between the terms study and publication, since the reviews included multiple publications about one study and thus relate to the same sample. Due to the heterogeneity of methods and reported values, a quantitative pooling of data was not feasible. Descriptive data were extracted and displayed as rounded percentage for a better comparison (e.g. 12/20 (55%) studies were effective). This led to rounding errors in some cases, thus the sum of percentages did not always add up to 100%. A further facet that needs consideration is that some reviews included multiple health outcomes at a time, hence they were mentioned repeatedly in the detailed results for PA, SB, and HE. Additionally, the total number of included publications in the reviews has been used for the results which led to some studies being included two or three times. Between-group effects were indicated as temporary if significant differences between the groups were only present at one and not at all timepoints following the intervention. Effect measures from included meta-analyses were reported in greater detail than systematic reviews due to the additional information provided by the quantitative report and subgroup analyses. Standardized mean differences (SMD), also known as Cohen's d, were classified with 0.2 as small, 0.5 as moderate, and 0.8 as large (Cohen, 1988). Hedge's g was also interpreted by the same rule of thumb. Heterogeneity was reported using the I^2 value, where values of 0% to 40% may indicate no important, 30% to 60% indicate moderate, 50% to 90% substantial, and 75% to 100% considerable heterogeneity (Deeks, Jonathan J, Higgins Julian PT, Altman Douglas G, and Cochrane Statistical Methods Group, 2019). Due to inconsistent reporting, additional values for significance of heterogeneity like Q and X^2 were not reported.

Results

Out of the 3895 reviews initially located and downloaded, 587 doublets were removed. During title and abstract screening, additional 3233 studies were excluded, with 75 studies remaining for full text screening. Sixty-four of these articles were excluded due to above mentioned exclusion criteria. This resulted in a total of 11 systematic reviews and meta-analyses which were included in this umbrella review (Böhm et al., 2019; Buckingham et al., 2019; Direito et al., 2017; Ferrer & Ellis, 2017; Hamel et al., 2011; McIntosh et al., 2017; Muellmann et al., 2018; Nour et al., 2016; Rocha & Kim, 2019; Schoeppe et al., 2016; Stephenson et al., 2017) (for more details see Flow-Chart in Figure 1). The updated search located 472 additional articles which were all excluded after title and abstract screening.

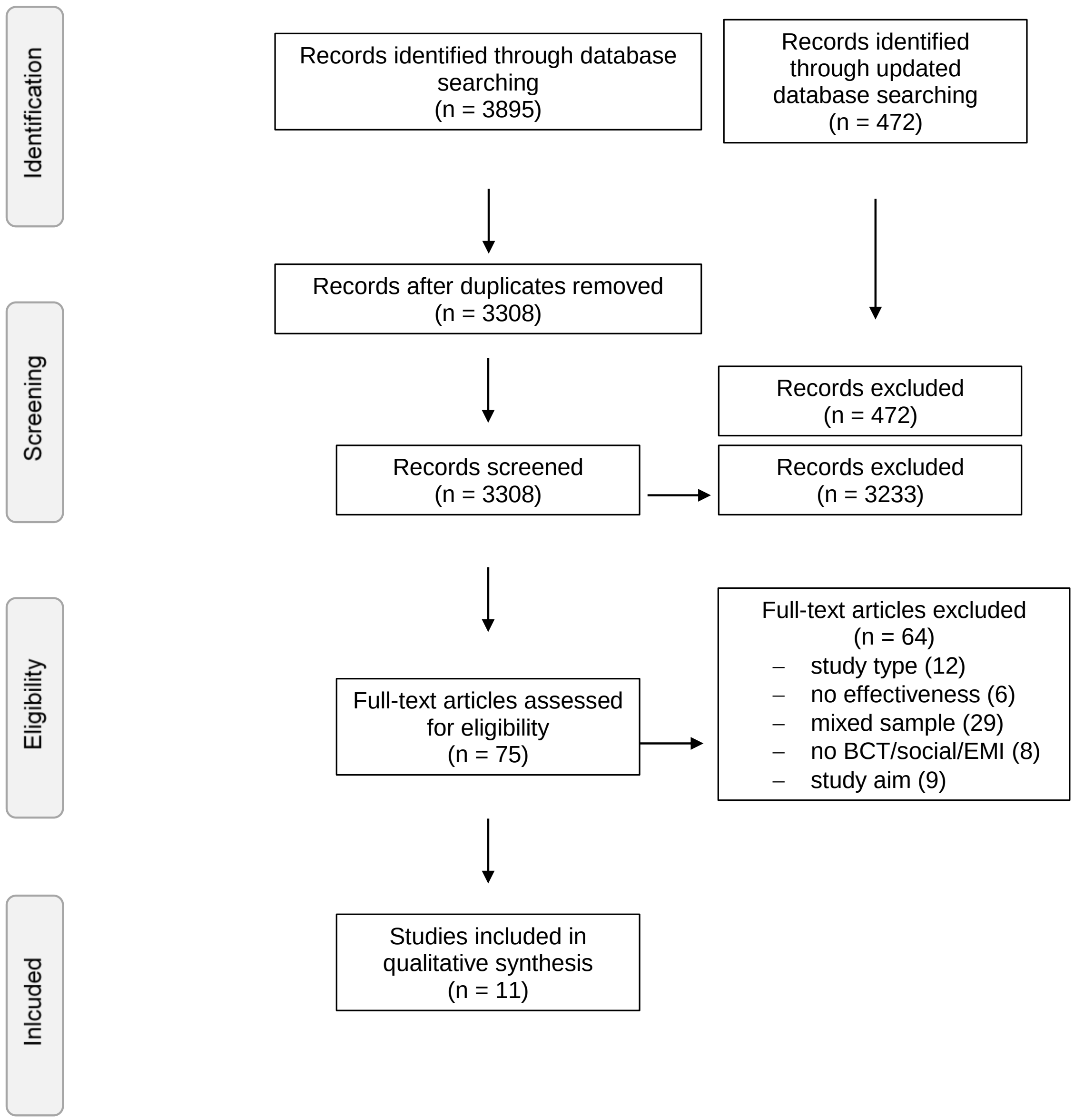

Figure 1. PRISMA Flow Chart of the study selection process.

Description of the included studies

The 11 reviews included a total of 195 publications (182 studies) published between 1998 and 2018, with 167 of these publications being included once throughout the reviews and 13 publications being included in two or three reviews, accounting for 28 publications.

The included original research studies were mainly conducted in USA and Canada and Europe, and the most common study designs were randomized control trials (RCTs). The duration of interventions ranged from one session to 24 months, with the majority (92%) of interventions lasting at least four weeks. Sample sizes ranged from 458 (Ferrer & Ellis, 2017) to 73,417 participants (Buckingham et al., 2019) for the reviews and added up to 114,430 participants throughout all studies. The full details of the study characteristics of articles included in the umbrella review are displayed in Table 1.

Table 1 Study characteristics of articles included in the umbrella review

Author type of review	Aim	Target population Setting	Country (number of studies)	N studies (N participants, age [years])	**PA/SB/HE** Main outcomes inclusion criteria (device-measured/ self-reported/ both)	Intervention effectiveness for **PA/SB/HE**	AMSTAR Review Quality
Böhm et al. 2019 (Böhm et al., 2019) Systematic review	To examine mHealth effectiveness for PA	children and/or adolescents NA	Australia (2), New Zealand (1), Canada (1), Israel (1), Poland (1), USA (1)	7 (1164, 8 - 18)	**PA** Measured at least 1 PA-related variable as the outcome, (5/ 0/ 2)	**PA**: Ø	6/11 medium
Buckingham et al. 2019 (Buckingham et al., 2019) Systematic review	To examine mHealth effectiveness, feasibility and acceptability for PA and SB	NA workplace	USA (11), Australia (5), Canada (2), Netherlands (2), Belgium (1), Singapore (1), Finland (1), Norway (1), in multiple countries (1)	30 publications, 25 (73417, 18+)	**PA/SB** Any quantitative measure as primary outcome (21/ 4/ 0)	**PA**: 14/25 (56%) ↑ over time or vs control 6/25 (24%) Ø 3/25 (12%) ↓ over time or vs control 2/25 (8%) N/A Between-group difference for ↑ of around 847 (95% CI 68 - 1625) to 2183 (95% CI 992 - 3344) steps/day **SB**: 4/10 (40%) ↓over time or vs control 3/10 (30%) Ø 3/10 (30%) ↑over time or vs control	7/11 medium

Direito et al. 2017 (Direito et al., 2017) Systematic review and meta-analysis	To examine mHealth effectiveness for PA and SB To investigate relationship between the effect size and the nature of PA/SB outcomes and to code the BCTs	NA NA	USA (11), Australia (3), United Kingdom (3), Austria (1), Portugal (1), Ireland (1), Canada (1)	21 (1701, 8.4 - 71.7)	**PA/SB** Duration or an estimate of energy expenditure (9/ 12/ 4) 1 not validated	**PA:** *Total PA* (7 studies): Ø SMD=0.14, 95 % CI −0.12 - 0.41, I^2=60 % *MVPA* (9 studies): Ø SMD=0.37, 95 % CI −0.03 - 0.77, I^2=78% *Walking* (8 studies): Ø SMD=0.14, 95%CI −0.01 - 0.29, I^2=0 % **SB** (5 studies): ↓ SMD=−0.26, 95 % CI −0.53 - −0.00, I^2=0%	8/11 medium
Ferrer et al. 2017 (Ferrer & Ellis, 2017) Systematic review	To examine eHealth effectiveness for PA	NA Facebook	Not reported	8 (458, all ages (*M* =24.3, *SD* = 7,5) 93,4% female	**PA** Either as a primary or a secondary outcome measure (2/ 4/ 2)	**PA:** 2/5 (40%) RCTs ↑ group by time interaction for steps per week and light PA participation 0/5 (0%) RCTs reported significant main effects for group, 4/5 (80%) RCTs reported significant main effects for time 1/3 (33%) non-RCTs studies ↑ group by time interactions for self-reported total PA 1/3 (33%) ↑ total steps during the social condition 1/3 (33%) ↑ self-reported mean minutes per week for all categories of PA	4/11 medium

Hamel et al. 2011 (Hamel et al., 2011) Systematic review	To examine eHealth effectiveness for PA and HE	preadolescents and adolescents, 8 – 18 years NA	USA (11), Belgium (3)	14 (6123 reported, 9 - 18)	**PA** PA or a PA-related health change as an outcome variable (2/ 10/ 2)	**PA:** 6/9 (67%) school-based interventions either ↑ PA in the intervention conditions and/or ↓ weight or BMI 2/5 (40%) home-based interventions either ↑ PA or ↓ BMI	4/11 medium
McIntosh et al. 2017 (McIntosh et al., 2017) Systematic review	To examine eHealth effectiveness for PA	young people attending school, college or university NA	USA (4), Netherlands (2), Thailand (1), Japan (1), Canada (1), Europe (1)	10 (5352, young people)	**PA** Primary or secondary outcome (8/ 1/ 1)	**PA:** 8/10 (80%) ↑ over time or vs control	5/11 medium
Muellemann et al. 2018 (Muellmann et al., 2018) Systematic review	To examine eHealth effectiveness for PA To compare effectiveness with either no intervention or a non-eHealth intervention.	older adults, 55 years or above NA	USA (11), Netherlands (3), Belgium (1), Spain (1), Australia (1), New Zealand (1), Malaysia (1)	25 publications, 20 (6671, 56 - 79.8 years)	**PA** Intervention effectiveness for any measure of PA (5/ 13/ 2)	**PA:** 9/9 (100%) web-based interventions ↑ over time or vs control 4/7 (57%) telephone-based interventions ↑ over time or vs control 3/4 (75%) text messaging-based interventions ↑ group over time or vs control	7/11 medium
Nour et al. 2016 (Nour et al., 2016) Systematic review and meta-analysis	To examine e/mHealth efficacy for HE	young adults, 18 to 35 years NA	USA (8), Australia (4), New Zealand (1), Malaysia (1)	14 (7984, *M* = 20.8)	**HE** Primary or secondary aim of increasing FVI (0/ 14/ 0), 5 not validated	**HE:** *FVI* (8 studies): ↑ SMD=0.22 95% CI 0.11 - 0.33, I^2=68.5% *Vegetable intake* (5 studies): ↑ SMD=0.15 95% CI 0.04 - 0.28, I^2=31.4%	9/11 high

Rocha et al. 2019 (Rocha & Kim, 2019) Meta-analysis	To examine eHealth effectiveness for HE To investigate the relationship of effectiveness and intervention characteristics (eHealth tool, tailoring, BCTs, and age group)	NA NA	USA (10), Netherlands (3), Scotland (1), Belgium (1), Portugal (1), Italy (1), Sweden (1), New Zealand (1)	19 (6894, *M* = 4,5 to 57,75)	**HE** Reporting FVI results quantitatively (0/ 19/ 0), 5 not validated	**HE:** *FVI* (19 studies): ↑ g=0.26 95% CI 0.17, 0.35, I^2=62.77, p<.001	6/11 medium
Schoeppe et al. 2016 (Schoeppe et al., 2016) Systematic review	To examine mHealth effectiveness for PA, SB and HE	children and/or adults NA	USA (9), Australia (6), Canada (3), Switzerland (2), Netherlands (2), Ireland (2), Italy (1), Israel (1), New Zealand (1)	30 publications, 27 (2699, 8 - 71)	**PA/SB/HE** Efficacy for behavior change. All types and units of measurements (8/ 13/ 6)	**PA SB and/or HE**: 19/27 (70%) ↑ in behavioral and related health outcomes either over time or vs control 5/10 (50%) ↑ single health behavior interventions vs control 7/17 (41%) ↑ multiple health behavior interventions vs control 8/13 (62%) ↑app in conjunction with other intervention strategies vs control 5/14 (36%) ↑ stand-alone app interventions vs control	4/11 medium
Stephenson et al. 2017 (Stephenson et al., 2017) Systematic review and meta-analysis	To examine e/mHealth for SB To identify the BCTs used within interventions	adults, 18 years or above NA	Not reported	17 (1967, *M* = 20,4 - 64,1)	**SB** Device-measured or self-reported or proxy measure of SB (8/ 6/ 3)	**SB** (15 studies): ↓ −41.28 min/day 95% CI -60.99 - −21.58, I^2=77% at end point follow-up	5/11 medium

Abbreviations: *HE* healthy eating, *M* mean, *NA* not available, *PA* physical activity, *RCT* randomized control trial, *SB* sedentary behavior, *SD* standard deviation, *USA* United States of America

Two reviews focused on children and adolescents (Böhm et al., 2019; Hamel et al., 2011), four focused adults (Buckingham et al., 2019; Muellmann et al., 2018; Nour et al., 2016; Stephenson et al., 2017) and five included participants of all ages (Direito et al., 2017; Ferrer & Ellis, 2017; McIntosh et al., 2017; Rocha & Kim, 2019; Schoeppe et al., 2016). Five systematic reviews focused on PA outcomes (Böhm et al., 2019; Ferrer & Ellis, 2017; Hamel et al., 2011; McIntosh et al., 2017; Muellmann et al., 2018), one meta-analysis focused on SB outcomes (Stephenson et al., 2017), and one meta-analysis and one systematic review included both PA and SB outcomes (Buckingham et al., 2019; Direito et al., 2017). HE was the main outcome in two meta-analyses (Nour et al., 2016; Rocha & Kim, 2019), and one systematic review included PA, SB, and HE as main outcomes (Schoeppe et al., 2016).

Eight reviews reported the use of theoretical frameworks (Böhm et al., 2019; Buckingham et al., 2019; Ferrer & Ellis, 2017; Hamel et al., 2011; McIntosh et al., 2017; Muellmann et al., 2018; Nour et al., 2016; Schoeppe et al., 2016), and 78/125 (62%) publications in these reviews reported the use of a theoretical foundation. The most common reported theories were social cognitive theory (n = 29), transtheoretical model (n = 16), theory of planned behavior (n = 10), self-determination theory (n = 10) and I-change model (n = 7). Four reviews (Buckingham et al., 2019; Direito et al., 2017; Rocha & Kim, 2019; Stephenson et al., 2017) coded the use of BCTs using a taxonomy of behavior change (Michie et al., 2011; Michie et al., 2013) and two reviews (Böhm et al., 2019; Ferrer & Ellis, 2017) reported BCTs without coding them. The BCTs, which were most frequently reported by the reviews, were goal setting (n = 5), self-monitoring (n = 4), social support (n = 4), prompts/cues (n = 4), feedback on the behavior (n = 3) and instruction on how to perform the behavior (n = 2). Since the BCTs were neither coded nor reported in a comparable way by the reviews, a more detailed summary was not feasible.

The majority of intervention studies were socially embedded (111/182, 62%). School, university, or college settings were mentioned in 45 studies, workplace in 37 studies, home and/or community-based study populations were reported in 17 studies, while two studies reported a combination of workplace and home setting. A social media setting was mentioned in eight studies, and supermarket and online setting in one study each. Two reviews (Hamel et al., 2011; Schoeppe et al., 2016) examined whether the interventions involved social support from the setting or solely took place in this context. Social support through peers and/or friendly challenges was described in six studies (Schoeppe et al., 2016) and parental involvement in three studies (Hamel et al., 2011). None of the reviews reported the use of JITAI or EMI.

Overall effectiveness

The heterogeneity of the included studies concerning study type, outcome parameter, and assessment method was high. Thus, the overall effectiveness reported in the reviews is displayed in the following paragraph for any significant differences, which were found for the e/mHealth interventions over time or vs. a control group. Of all included studies, 10/182 did not report intervention effectiveness. The remaining 172 studies found a significant benefit for the intervention group over time and/or vs. a control group in 101/172 (59%) cases. No significant differences were found in 68/172 (40%) studies, and 3/172 (2%) resulted in a significant deterioration of the parameter over time and/or vs. control (see Table 1).

Effectiveness vs. Control

The between group differences for the included systematic reviews are displayed in the following chapters and the results of the included meta-analyses are reported in further detail.

PA

PA (i.e. time spent in different PA intensities, step count, PA frequency, PA goal achievement, school related PA, and leisure time PA) was assessed by seven systematic reviews (PA outcome in 106 studies) (Böhm et al., 2019; Buckingham et al., 2019; Ferrer & Ellis, 2017; Hamel et al., 2011; McIntosh et al., 2017; Muellmann et al., 2018; Schoeppe et al., 2016) and one meta-analysis (PA outcome in 20 studies) (Direito et al., 2017). Of the 126 studies included in these reviews, 58 studies used device-measured outcomes, 52 used self-report (1 not validated), and 16 used a combination of both measures.

Systematic reviews concerning PA did not report group differences or did not use a control group in 14/106 studies. The remaining 92 studies found significant group differences in favor of the intervention group in 19/92 (21%) studies, temporary significant group differences in favor of the intervention group in 25/92 (27%) studies and 49/92 (53%) showed no significant differences between the groups. One meta-analysis (Direito et al., 2017) included participants aged from 8.4 to 71.7 years and found no significant pooled effects using a random effect model between the eHealth and a usual/minimal care group for total PA (seven studies, SMD = 0.14, 95 % CI [−0.12, 0.41]; I^2 = 60 %), MVPA (nine studies, SMD = 0.37, 95 % CI [−0.03, 0.77]; I^2 = 78%) and measures of walking (eight studies reporting steps/day and walking duration/day, SMD = 0.14, 95%CI [−0.01, 0.29]; I^2 =0 %). Subgroup analysis between device-measured and self-reported results showed no significant differences in the eHealth group for total PA, MVPA and walking.

SB

SB (i.e. sitting time (overall and occupational), sedentary time (overall and occupational), screen time, and computer activity) was assessed by two systematic reviews (SB outcome in 13 studies) (Schoeppe et al., 2016; Stephenson et al., 2017) and two meta-analyses (SB outcome in 20 studies) (Buckingham et al., 2019; Direito et al., 2017). Of the 33 studies included in these reviews, 15 studies used device-measured outcomes, 16 used self-report (one not validated), and two used a combination of both measures.

The systematic reviews concerning SB included 4/13 studies which did not report group differences or did not involve a control group. The remaining nine studies showed a significant group difference in favor of the intervention group in 2/9 (22%) studies, 6/9 (67%) studies with no significant differences between the groups, and 1/9 (11%) reported a significant group difference in favor of the control group. The first meta-analysis (five studies) (Direito et al., 2017) which included participants aged from 8.4 to 71.7 years found a significant reduction of SB in favor of the intervention group using a random effect model. This pooled effect was negative and small (SMD = −0.26, 95 % CI [−0.53, −0.00]; I^2 =0 %) with no evidence of heterogeneity. Subgroup analysis between device-measured and self-reported results showed no significant differences for the intervention group in SB. The second meta-analysis on SB (15 studies) (Stephenson et al., 2017) included only adults (20.4 to 64.1 years) and showed a significant pooled reduction of SB with a substantial heterogeneity (-41.28 min/day, 95% CI [-0.99, −21.58], I^2 = 77%; n = 1402) in favor of the intervention group at the end point follow-up measurement using a random effect model. Analysis for device-measured (eight studies) results showed a significant pooled reduction of −35.07 min/day with a low heterogeneity (95% CI [-46.57, −23.57], I^2 =21%; n = 595), while self-reported measures (seven studies) led to a significant reduction of −52.66 min/day with a considerable heterogeneity (95% CI, [−93.63, −11.69], I^2 =88%; n = 807) at end point. The comparison between device-measured and self-reported results has not been conducted by this meta-analysis. The additional analysis of short-term measures for overall SB (less than 3 months, 10 studies) showed a significant mean reduction of −42.42 min/day with a substantial heterogeneity (95% CI [-63.21, −21.63], I^2 =61%; n = 760), the medium-term measures (three to six months, five studies) showed a significant mean reduction of −37.23 min/day with a considerable heterogeneity (95% CI [-73.70, −0.75], I^2 =85%; n =691) and the long-term measures (over six months, three studies) showed no significant mean reduction with a low heterogeneity (−1.65 min/day, 95% CI [-14.77, 11.47], I^2 =23%; n =670).

HE

HE (i.e. FVI, vegetable intake, and healthy dietary choices) was assessed by one systematic review (HE outcome in 13 studies) (Schoeppe et al., 2016) and two meta-analyses (HE outcome in 33 studies, focus on FVI) (Nour et al., 2016; Rocha & Kim, 2019). All of the 46 studies included in these reviews used self-reported results, 10 of which were not validated.

The systematic review concerning HE did not report group differences or did not involve a control group for 1/13 studies. The remaining 12 studies found a significant group difference in favor of the intervention group in 2/12 (17%) studies, a temporary significant group difference in favor of the intervention group in 3/12 (25%) studies, and 7/12 (58%) showed no significant differences between the groups. One meta-analysis (Nour et al., 2016) included young adults (M =20.8 years) and showed a significant increase in FVI (eight studies) calculated by a random effect model with a small pooled Cohen's *d* of 0.22 (95% CI [0.11, 0.33]) and a substantial heterogeneity (I^2 = 68.5%). Effects for vegetable intake alone were also assessed (five studies) and the pooled effect showed a negligible effect with low heterogeneity (Cohen's d = 0.15, 95% CI [0.04, 0.28], I^2 = 31.4%). The second meta-analysis (Rocha & Kim, 2019) included participants of all ages (4.5 to 57.75 years) and found a significant increase of FVI in favor of the intervention group using a random effect model with a small Hedge's *g* and substantial between study heterogeneity (g = 0.26, SE = 0.05, 95% CI [0.17, 0.35], I^2 = 62.77). Subgroup analyses revealed that computer-based (i.e. non-Internet based) eHealth interventions (three studies) showed the largest effect (g = 0.44), followed by SMS interventions (three studies) with a Hedge's *g* of 0.41, while internet-based interventions (nine studies) showed a Hedge's *g* of 0.19 and CD-ROM, mobile apps and video game interventions (four studies) showed no significant improvements. The subgroup analysis relating to age groups yielded no significant differences between adults (11 studies), adolescents (four studies), and children (four studies). Interventions including adults and adolescents showed significant improvements in favor of the intervention group with Hedge's g of 0.26 and 0.35 respectively, while interventions conducted with children showed no significant effects.

Determinants of effective Interventions

The extraction of effect sizes regarding the influence of theoretical foundation/BCTs, social influences, and EMI/JITAs on the efficiency of e/mHealth interventions was not feasible so that only descriptive results were reported in this umbrella review (see Table 2).

Table 2 Intervention effectiveness and the reported use of theoretical foundation, BCT, social context and EMI/JITAI in the included reviews

Author Type of review	Intervention duration Included study designs	Theoretical foundation	BCT	Social context	EMI/JITAI
Böhm et al. 2019 (Böhm et al., 2019) Systematic review	1 - 12 months 4 RCTs, 1 RC cross over, 2 before-and-after trials	Total number of theory-based studies NA 4/7 (57%) social cognitive theory	3/7 (43%) additional BCTs	7/7 (100%) recruited in schools with interventions in and outside the school setting Combining school-based interventions with family or community involvement for social support is potentially effective	NA
Buckingham et al. 2019 (Buckingham et al., 2019) Systematic review	1 - 12 months 10 RCTs, 10 prospective cohort studies, 1 combination of methods mentioned above, 3 cluster-RCT, 1 parallel group uncontrolled randomized trial, 1 prospective cluster trial with an asynchronous control group	15/25 (60%) based on a named behavior change theory and/or principles of behavioral economics 2/25 (8%) studies alluded to BCTs or theory in their discussion 8/25 (32%) neither theory nor BCT	Most frequent BCTs: Self-monitoring of the behavior or outcome (n=22, 88% of studies) provision of feedback on the behavior or outcome (n=21, 84%) goal setting for the behavior or outcome (n=17, 68%) social comparison (n=14, 56%) social support (n=12, 48%)	25/25 (100%) public and private sector workplace setting No associations were found between type of workplace and impact on PA	NA

Direito et al. 2017 (Direito et al., 2017) Systematic review and meta-analysis	Median = 9 weeks 21 RTCs	NA	Total BCTs: $M = 5.4/93$ BCTs ($SD = 2.6$, range 0 to 12) More BCTs were employed with intervention groups ($M = 6.9$, $SD = 2.6$, range 2 to 12) than with comparator groups ($M = 3.1$, $SD = 2.2$, range 0 to 10) Most frequent BCTs: 81% goal setting (behavior 74% self-monitoring of behavior 65% social support (unspecified) 55% feedback on behavior 55% instruction on how to perform the behavior 48% adding objects to the environment 45% information about health consequences 45% prompts/cues	NA	NA Prompts and cues common BCT
Ferrer et al. 2017 (Ferrer & Ellis, 2017) Systematic review	3 - 12 weeks 5 RCTs, 2 single group designs, 1 within-subject crossover	4/8 (50%) theory-based interventions 1/8 (13%) theory of planned behavior 1/8 (13%) social cognitive theory 2/8 (25%) not specified	8/8 (100%) interventions used behavior modification strategies including goal setting, self-monitoring, prompts, and social support	8/8 (100%) social-media interventions (facebook) 7/8 (88%) of the Facebook interventions ↑ some type of PA behavior change (i.e., interactions, main effects for time, differences between conditions) 2/8 (25%) ↑ for the treatment group compared to the control group.	NA

Hamel et al. 2011 (Hamel et al., 2011) Systematic review	2 weeks - 2 years 7 RCTs, 5 quasi-experimental, 1 repeated measures, 1 pretest-posttest control group design	9/14 (64%) theory-based interventions Social cognitive theory most frequent Other theories: Theory of reasoned action transtheoretical model health belief model theory of planned behavior attitude, social influence, and self-efficacy model 4/14 (29%) having more than one theory	NA	9/14 (64%) school-based 3/14 (21%) home-based 1/14 (7%) camp- and home-based 1/14 (7%) scout troop and home-based 3/14 (21%) of these included parental involvement School-based interventions were more effective than e.g. home-based interventions and parental support might be important	NA
McIntosh et al. 2017 (McIntosh et al., 2017) Systematic review	6 weeks - 4 months 6 RCTs, 3 before and after quasi-experimental designs 1 cluster-RCT	9/10 (90%) theory-based interventions 5/10 (50%) social cognitive theory 2/10 (20%) theory of planned behavior 1/10 (10%) transtheoretical model 1/10 (10%) SMART goals	NA	10/10 (100%) students attending school, college or university Effect of social context not analyzed	NA
Muellemann et al. 2018 (Muellmann et al., 2018) Systematic review	1 - 24 months 18 RCTs, 2 quasi-experimental design	16/20 (80%) theory-based interventions most common: 9/20 (45%) social cognitive theory 8/20 (40%) transtheoretical model 7/20 (35%) self-determination theory 7/20 (35%) i-change model	NA	NA	NA
Nour et al. 2016 (Nour et al., 2016) Systematic review and meta-analysis	one-off contact - 6 months 14 RCTs	6/14 (43%) theory- or education-based interventions 5/14 (36%) transtheoretical model 2/14 (14%) social cognitive theory 2/14 (14%) theory of planned behavior and the theory of habit formation 7/14 (50%) applied self-efficacy in their intervention	NA	no group intervention 11/14 (79%) university setting 2/14 (14%) N.A. 1/14 (7%) home based Effect of social context not analyzed	NA

Rocha et al. 2019 (Rocha & Kim, 2019) Meta-analysis	one-time session - 24 weeks 14 RCTs, 3 cluster-RCTs, 2 non-randomized studies	NA	1 to 7 BCTs used 20/40 BCT categories identified Most common: 68% provide instruction on how to perform the behavior 47% Provide feedback on performance 26% goal setting on behavior	8/19 (42%) school setting 6/19 (32%) community-based 2/19 (11%) workplace-based 1/19 (5%) clinic-based (prevention) 1/19 (5%) online-based 1/19 (5%) supermarket-based Effect of social context not analyzed	NA 15/19 (79%) tailored interventions 4/19 (21%) were nontailored interventions
Schoeppe et al. 2016 (Schoeppe et al., 2016) Systematic review	1 - 24 weeks 19 RCTs, 4 pre-post studies, 3 controlled trials, 1 randomized trial	15/27 (56%) theory-based interventions 4/27 (15%) transtheoretical model 4/27 (15%) social cognitive theory 3/27 (11%) self-determination theory 2/27 (7%) theory of planned behavior 1/27 (4%) control systems 1/27 (4%) theory of self-regulation 1/27 (4%) behavior change wheel	NA	6/27 (22%) social support 3/6 interaction with peers 4/6 friendly team challenges Effect of social context not analyzed	NA

Stephenson et al. 2017 (Stephenson et al., 2017) Systematic review and meta-analysis	5 days - 24 months 17 RCTs	NA	1 to 15 BCTs used 20/93 BCTs categories identified most common: 88% instruction on how to perform a behavior 71% social support (unspecified) 65% prompts and cues 65% adding objects to the environment	10/17 (59%) workplace setting 5/17 (29%) Community/home setting 2/17 (12%) workplace and community/home setting Effect of social context not analyzed	NA Prompts and cues common BCT

Abbreviations: *BCT* behavior change technique, *CI* confidence interval, *HE* healthy eating, *EMI* ecological momentary intervention, *FVI* fruit and vegetable intake, *g* Hedges' g, I^2 percentage of variation across studies that is due to heterogeneity rather than chance, *JITAI* just-in-time adaptive intervention, *M* mean, *NA* not available *PA* physical activity, *RCT* randomized control trial, *SB* sedentary behavior, *SMD* standardized mean difference, ↑ significant higher value ($p<0.05$), Ø no significant difference, ↓ significant lower value

Theoretical foundation and BCTs were mentioned in all the included reviews (Böhm et al., 2019; Buckingham et al., 2019; Direito et al., 2017; Ferrer & Ellis, 2017; Hamel et al., 2011; McIntosh et al., 2017; Muellmann et al., 2018; Nour et al., 2016; Rocha & Kim, 2019; Schoeppe et al., 2016; Stephenson et al., 2017). One review (McIntosh et al., 2017) related to PA noted that 5/5 (100%) studies based on social cognitive theory led to significant differences over time or vs. control compared to 1/2 (50%) for theory of planned behavior and 1/1 (100%) showing a temporary significant difference directly after the intervention for transtheoretical model. Another review concerning PA (Muellmann et al., 2018) also found theory-based interventions more effective than those without a theoretical foundation. A third review concerning PA (Hamel et al., 2011), which found that 6/9 (67%) theory-based interventions showed significant improvements of the intervention group over time or vs. control, while only 2/5 (40%) without a theoretical foundation led to such improvements, is in line with these findings. The inclusion of BCTs was associated with higher effectiveness of PA, SB, and HE interventions in one review (Schoeppe et al., 2016). However, the question which BCTs are linked to effectiveness has not been answered by this review. Two meta-analyses (Direito et al., 2017; Stephenson et al., 2017) reported the usage of BCTs for PA and SB interventions, but did not link the use of BCTs to effectiveness due to the small number of studies included. For healthy eating behavior, the use of BCTs was one key component of successful interventions, while the impact of using multiple BCTs remained unclear (Nour et al., 2016). Further, a more recent meta-analysis (Rocha & Kim, 2019) revealed that the inclusion of seven to eight BCTs (four studies) resulted in a statistically significant larger effect size (SMD = 0.42, SE = 0.10, 95% CI [0.21, 0.62], $p < .001$) than those involving four to six BCTs (seven studies) and one to three BCTs (seven studies). In a next step, the meta-analysis found no statistically evidence for specific BCTs yielding larger effect sizes.

The influence of a social settings concerning effectiveness has not been reported in detail by the included reviews and two reviews (Hamel et al., 2011; Schoeppe et al., 2016) reported on the matter at all. The integration of eHealth interventions in school settings was reported to lead more often (6/9, 67%) to positive effects on PA or weight reduction in comparison to home-based interventions (2/5, 40%) (Hamel et al., 2011). Another possible influence on effectiveness mentioned in this review was parental influence (Hamel et al., 2011). The second review about mHealth interventions points out that efficient interventions often include social support related to peers and friendly team challenges among many other facets (Schoeppe et al., 2016). However, since both reviews did not report effect sizes, and there were a variety of other possible facets contributing to effectiveness, the magnitude of the potential influence for social settings remains unclear.

Since none of the reviews reported the use of EMI/JITAIs, the question concerning their effectiveness has to be left unanswered by this umbrella review.

Study Quality

Mean study quality of the included reviews as assessed by the AMSTAR tool (Shea et al., 2007) (maximum score 11, score ratings: low = 0-3, medium = 4-7, and high = 8-11 (Sharif et al., 2013)) was medium (M = 5,9/11) while one review scored high (9/11) (Nour et al., 2016). None of the included reviews reported the conflict of interest of the included studies and only one review provided a list of all included and excluded studies (Nour et al., 2016). For the score of every criterion see additional file 2. Risk of bias ratings conducted by the authors of the included reviews was mainly medium to high with some studies of low risk (see Table 3).

Table 3 Time period, Intervention tools, quality of included studies in the reviews, and recommendations for future research

Author type of review	Time Period Searched (included studies)	mHealth/eHealth tools	Quality of included studies	Recommendations for future research
Böhm et al. 2019 (Böhm et al., 2019) Systematic review	January 2012 to June 2018 (2014 - 2016)	Mobile phones, smartphones, tablets, or wearables	Tool: Cochrane Handbook for Systematic Reviews of Interventions Risk of bias: 2/5 (40%) medium 3/5 (60%) high	1) PA intervention programs for children/adolescents with a greater BMI z-score 2) intervention programs with a longer period of time (≥6 months) 3) sufficiently large number of participants (≥250) 4) bypass self-reported measurements 5) implement theoretical frameworks and BCTs 6) follow-up beyond postintervention 7) age- and sex-specific interventions 8) engagement of children and adolescents with wearable activity trackers 9) impact of social support (school/family) 10) multicomponent interventions 11) cost-effectiveness analyses
Buckingham et al. 2019 (Buckingham et al., 2019) Systematic review	January 2007 to February 2018 (2009 - 2018)	mHealth interventions: mobile phone, smartphone apps, personal digital assistants, tablets, wearable activity monitors/ trackers	Tool: Effective Public Health Practice Project Quality rating: 1/25 (4%) strong, 9/25 (36%) moderate, 15/25 (60%) weak	1) larger samples and more diverse workspace settings 2) report intervention components and outcomes in greater detail 3) SB in addition to PA, and bypass self-report 4) no-intervention control or a reliable baseline measurement 5) wider impact on health and wellbeing 6) mixed and qualitative methods 7) adverse events associated with mHealth use 8) mHealth vs multi-component interventions 9) subgroup differences
Direito et al. 2017 (Direito et al., 2017) Systematic review and Meta-Analysis of RCTs	From earliest available to January 2015 (2007 - 2014)	mHealth interventions: mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants	Tool: Cochrane Collaboration's tool No total rating: High Risk of Bias for blinding, unclear allocation, other biases were low for most studies	1) long-term effectiveness and cost-effectiveness of mHealth interventions 2) dose-response relationship between intervention exposure and outcomes 3) report intervention components and outcomes in greater detail 4) efficacy of more advanced technology than SMS

Ferrer et al. 2017 (Ferrer & Ellis, 2017) Systematic review	not specified (2010 - 2014)	Facebook based interventions	Not assessed	1) no-intervention control 2) target a broader diversity of participants 3) attrition rates for varying durations of interventions 4) theory-based content and measure the effects of those mediators 5) effectivity of social support 6) validate self-report measures against device-measured outcomes of PA 7) match the PA assessment method to the stated goals and outcomes of the intervention 8) long term follow-up
Hamel et al. 2011 (Hamel et al., 2011) Systematic review	1998 to 2010 (1999 - 2009)	Computer- and web-based interventions	Tool: Critical Appraisal Skills Programme of the Public Health Resource Unit Quality rating: No summary presented	1) bypass self-report 2) sex specific interventions 3) involve support persons (e.g. parents or peers) and analyze effectivity 4) integrate into existing school curriculum 5) include a theoretical framework 6) individual tailoring
McIntosh et al. 2017 (McIntosh et al., 2017) Systematic review	2010 to July 2016 (2010 - 2014)	Web-based or eHealth interventions	Tool: based on the critical appraisal for public health checklist Quality rating: 3/10 (30%) high 7/10 (70%) moderate	1) longer follow-up 2) address bias incorporated with self-reporting methods 3) utilize theoretical foundation for eHealth interventions 4) relationship of confounding facets to effectiveness 5) conduct power analysis of studies 6) scale up interventions
Muellemann et al. 2018 (Muellmann et al., 2018) Systematic review	from earliest available to April 2017 (1997 - 2017)	eHealth interventions: computer, telephone, smartphone, or tablet	Tool: Cochrane Collaboration's tool for assessing risk of bias Risk of bias: 1/20 (95%), low 19/20 (95%) moderate to high	1) eHealth interventions vs non-eHealth interventions promoting PA in older adults

Nour et al. 2016 (Nour et al., 2016) Systematic review and Meta-Analysis	1990 to August 2015 (2007 - 2014)	eHealth- and mHealth-based interventions: texting, email, mobile phone apps, phone calls, or websites	Tool: Cochrane Collaboration's tool for assessing risk of bias Risk of bias rating: majority of the studies unclear to high risk (attrition bias) 2/14 (14%) studies additionally high detection bias	1) longer follow-up in intervention 2) secondary outcomes (e.g.) weight and indicators of cardiovascular health) 3) focus primarily on vegetables 4) combine efficacious strategies and repeat exposure at a later date 5) develop validated tools for measuring vegetable intake in young adults 6) quantify a serving of vegetables 7) implement Biomarkers (e.g. vitamin C and beta-carotene) 8) more diverse samples 9) cost effectiveness for upscaling interventions 10) conduct process evaluations
Rocha et al. 2019 (Rocha & Kim, 2019) Meta-Analysis	1999 to July 2018 (1999 - 2017)	eHealth interventions: mobile devices (apps, text messages via cellphone), web or internet-based programs, computer-based programs (non-Internet based), and video games.	Tool: guided by the Cochrane's Risk of Bias Tool for RCTs Quality rating: 5/19 (26%) good 12/19 (63%) fair 2/19 (11%) poor	1) tailor based on distal correlates and proximal determinants of dietary habits 2) link the types of BCTs implemented in the eHealth interventions to effectiveness 3) develop validated tools for measuring FVI 4) report intervention components and outcomes in greater detail 5) use of the CALO-RE taxonomy for uniformity in the reporting of BCTs

Schoeppe et al. 2016 (Schoeppe et al., 2016) Systematic review	January 2006 to November 2016 (2010 - 2016)	mHealth (App interventions): stand-alone intervention using apps only, or a multi-component intervention including apps	Tool: 25-point criteria adapted from the CONSORT checklists Quality rating: 11/27 (40%) high 8/27 (30%) fair 8/27 (30%) low	1) test the efficacy of specific app features and BCTs 2) efficacy of stand-alone app intervention vs multi-component app interventions 3) efficacy of app vs website, print-based and face-to-face interventions 4) utilize larger sample sizes 5) tailor app interventions to specific population groups with high app usage (e.g., women, young people) 6) report app usage statistics using device and self-report measures 7) optimal duration and intensity of app interventions 8) user engagement and retention in app interventions 9) relationship between user engagement and intervention efficacy (considering socio-demographic and psychosocial facets)
Stephenson et al. 2017 (Stephenson et al., 2017) Systematic Review and Meta-analysis	from earliest available to June 2016 (2012 - 2016)	Computer, mobile or wearable technology	Tool: Cochrane Collaboration's risk of bias tool Risk of bias: 1/17 (6%) low 3/17 (18%) unclear 13/17 (76%) high	1) focus on attrition rates 2) improve reporting of BCTs 3) improve detection bias by using objective measurement tools of SB 4) conduct extended follow-up 5) include outcome measures that will be of interest to workplaces and policy makers 6) use adaptive interventions

Abbreviations: *AMSTAR* assessment of multiple systematic reviews, *App* smartphone application, *BCT* behavior change technique, *CONSORT* consolidated standards of reporting trials, *eHealth* electronic health, *FYI* fruit and vegetable intake, *mHealth* mobile health, *PA* physical activity, *SB* sedentary behavior

Discussion

This umbrella review provided an overview of e/mHealth interventions concerning PA, SB, and HE for primary prevention with a special focus on potentially important facets and their contribution to intervention effectiveness. To avoid an overwhelming heterogeneity in the included reviews, these facets have been pre-defined based on the current literature and previous umbrella reviews as theoretical foundation, BCTs, social context, and the use of JITAI/EMI. To the best of our knowledge, this umbrella review is the first to systematically analyze the potential impact for those predefined facets.

Effectiveness of e/mHealth interventions

Overall, findings of this umbrella review suggested that a majority (59%) of e/mHealth interventions were effective (including interventions eliciting short-term effects and interventions without control-group comparison). Since multiple studies reported a high heterogeneity (with low to high quality ratings), this result has to be interpreted with caution. Results of the systematic reviews including a control-group indicated that PA interventions were more often effective (48%) than interventions concerning HE (42%) and SB (22%). However, more than 50% of these effects for PA and HE interventions were only temporary and one SB study outcome (11%) was even in favor of the control group. In contrast to systematic reviews, quantitative findings of the included meta-analyses did not indicate any significant benefit for PA while SB and HE interventions showed significant small effects. A reason for the lack of effectiveness in the only meta-analysis concerning PA (Direito et al., 2017) may be that solely one original research study included a true control group and e/mHealth to usual/minimal care. Furthermore, the post hoc exploratory sensitivity analysis displayed two of the included studies as being the main reason for heterogeneity in this meta-analysis. After removal of these studies, results indicated a small but significant effect for MVPA (but none for total PA) and thereby partially support the findings of the systematic reviews that e/mHealth interventions can be effective tools to change all three health behaviors.

One facet which could have influenced the results of this umbrella review is the use of different assessment methods in the studies, as self-report measures are commonly reported to overestimate PA compared to sensor-based PA (Dyrstad et al., 2014). Considering the fact that some studies even used non-validated self-report tools in PA, SB, and HE interventions, these facets could have highly influenced findings. In the present umbrella review, the comparison of self-reported and device-measured outcomes showed no significant differences for PA and SB in one meta-analysis (Direito et al., 2017), and another meta-analysis reported lower heterogeneity and a descriptive difference of SB reduction for device-measured results (Stephenson et al., 2017). While all other reviews reported on the use of self-reported and device-measured results, the examination of influence on effectiveness has not been conducted and thus no assumptions about a potential impact of the measuring method could be made. However, examining the impact of the measurement method could be important, since self-reported and device-measured results often differ concerning the construct (e.g. measuring habitual PA or sport related PA) and the time epoch (e.g. regular/last week/month PA recall via questionnaire or measured PA during a defined time via accelerometry) (Dyrstad et al., 2014). Furthermore, the earliest study included in the reviews was published in 1997 and the complexity and capacity of sensors evolved rapidly since that time (Burchartz et al., 2020), allowing for more precise measurements and the combination of PA data with physiological parameters like heartrate or blood sugar (Reichert et al., 2020). This potential influence of different sensors on intervention effectiveness however, has not been considered in the reviews. Future reviews should specifically compare results derived by self-reports to device-measured outcomes and assess the impact of the complexity of sensors in order to further investigate the true impact of assessment methods and ease the interpretation of results.

The sustainability of intervention effects was reported to be low in the reviews for PA and HE, and quantification of one meta-analysis (Stephenson et al., 2017) showed that the effects of SB interventions diminish after 6 months, which is in accordance with other research (Prochaska, James O., and Wayne, F. Velicer, 1997). Intervention duration and engagement are also important facets influencing intervention effectiveness (Vandelanotte et al., 2007), but the influence remains unclear due to a lack of reporting by the included reviews. Future reviews should consider this link, especially if they are comparing sustainability of intervention effects over time.

The use of eHealth compared to mHealth might also influence the effectiveness. However, results are inconclusive since most reviews did only assess the intervention type but not compare the impact on effectiveness. One meta-analysis (Direito et al., 2017) which quantified the results found computer-based interventions to cause superior effectiveness compared to web- and app-based interventions. However, since mHealth is a more recent development and the amount of evidence is limited, this trend might be modified with more sophisticated approaches and more study results in the future (Vandelanotte et al., 2016). There is a clear need to include the comparison of effectiveness across devices in PA and SB interventions along with the influence of the age of participants in order to enhance and specify future interventions.

Influence of theoretical foundation, BCT, social aspects, and JITAIs on effectiveness

The diversity of results supported the importance to consider the underlying mechanisms for effective e/mHealth interventions in order to further develop the field of digital behavior change in general and in the area of primary prevention in particular.

The use of BCTs as a sub-section of theoretical foundation provided the most distinct picture and was highly associated with effective interventions for PA, SB, and HE interventions in one systematic review (Schoeppe et al., 2016). This finding was further supported by the two meta-analyses concerning HE (Nour et al., 2016; Rocha & Kim, 2019). The use of more BCTs enhanced intervention effectiveness for HE, whereas the impact of specific BCTs or combinations of BCTs remains unknown (Rocha & Kim, 2019), which has been a common finding in reviews on HE interventions (Villinger et al., 2019). The meta-analyses concerning PA and SB did not report the impact of BCTs on effectiveness which should be addressed by future research. Support for the use of a theoretical foundation for effective e/mHealth interventions concerning PA was found in three reviews (Hamel et al., 2011; McIntosh et al., 2017; Muellmann et al., 2018), and there were indications that social cognitive theory might be especially effective (McIntosh et al., 2017). The overall higher effectiveness for theoretical founded interventions supported the findings of a previous review about internet interventions (not focused on primary prevention) (Webb et al., 2010) but in contrast to our results, the theory of planned behavior was found to be more effective than social cognitive theory. Since direct comparisons of theory vs. no theory in the included reviews were scarce and only descriptive, there is a need for further investigation and better documentation of theoretical backgrounds in intervention studies in order to draw a clear conclusion. The lack of reporting regarding the impact of theoretical foundation on effectiveness for SB and HE should additionally be addressed by further research. A further aspect to consider in future studies is the compatibility of static behavior change theories to the technological advances which has not been addressed by the included reviews. While dynamically changing theories like the adapted versions of the Theory of Planned Behavior (Navarro-Barrientos et al., 2011) or the Social Cognitive

Theory (Martin et al., 2014) has been promoted in the development of JITAIs (Gonul et al., 2019b) the impact on intervention effectiveness should be assessed in future interventions.

In contrast to the potential impact of theoretical foundation and BCTs, most reviews did neither report nor analyze the association of embedding interventions into social contexts (e.g. involving family, peers, or co-workers in the intervention) and intervention effectiveness. Only three reviews (Ferrer & Ellis, 2017; Hamel et al., 2011; Schoeppe et al., 2016) reported on that matter but were unable to specify the influence due to a small sample and/or multiple other parameters linked to effectiveness. The importance of getting a better impression of social influences should however not be underestimated in order to conduct effective interventions in the future (Bandura, 1998). Including social facets can have an essential influence on intervention effectiveness and should be considered in future research (Wunsch et al., under review). Furthermore, intervention designs comparing e/mHealth interventions with clearly defined and controlled social contexts (e.g. social comparison, cooperative approaches) might help to gain evidence on the impact of social context.

No mention at all was found for the use of EMI/JITAI in this umbrella review. With the possibility to tailor and to continually adapt interventions to each person’s needs, as well as to deliver support at the most promising moment, there is a clear need for examination of this important field in the future (Hardeman et al., 2019; Heron & Smyth, 2010; Schembre et al., 2018).

Strengths and Limitations

The main strengths of this umbrella review consisted in summarizing the knowledge about the impact of multiple facets of effective behavior change interventions, derived from current literature, on effectiveness. Following a pre-registered protocol and systematically summarizing the evidence on the effectiveness of e/mHealth interventions in primary prevention ensured a replicable approach. Using a systematic search with pre-defined terms, following the PRISMA guidelines for reporting, and using AMSTAR for quality assessment thereby enhanced the transparency of the results. The inclusion of systematic reviews and meta-analyses following PRISMA guidelines ensured a solid foundation of higher quality reviews and a systematic reporting of the original research results.

Nonetheless, there are several limitations concerning the current umbrella review that need to be considered. First, the results of this umbrella review highly depended on the detailed reporting of the desired parameters in the reviews. Even if the original research studies reported on the issue but the reviews did not, the result has not been considered for this umbrella review. Even though the included reviews had to follow the PRISMA statement themselves, the quality of reviews was medium with a high discrepancy of included original research studies ranging from low to high scores and including several non-RCT studies. This might have impacted the conclusions of this umbrella review as well. The fact that 13 publications were included twice or more might also bias the evidence since those studies get a higher impact on the overall results. Finally, important studies might not be included in any review article yet since the conduction and publication of reviews produces a certain time lag compared to the present evidence.

Implications for practice and research

Results of this umbrella review can serve as a theoretical basis to conduct both, original research and review articles in the field of primary prevention using e/mHealth. Researchers should address the main research gaps, namely the impact of different theoretical foundations for interventions in different contexts, the adequate amount and types of BCTs, the impact of social context, and enhancing interventions with JITAIs, by conducting original research studies or especially focused reviews to close research gaps. For practitioners, we recommend to implement theoretical foundation and BCTs to their e/mHealth interventions in order to enhance intervention effectiveness. Furthermore, e/mHealth interventions should be adapted once further evidence emerges in order to maximize the usefulness of this fast-changing field of behavior change.

Future Directions

Even though the included reviews were conducted over the course of nearly a decade and thus represent different stages of e/mHealth tools, recommendations for future research given by the authors of the included reviews have a lot in common (for more details see Table 3). Based on these recommendations, a clear need for PA and SB studies is stated to bypass self-report and use validated and comparable device-measured outcomes instead (Böhm et al., 2019; Buckingham et al., 2019; Ferrer & Ellis, 2017; Hamel et al., 2011; McIntosh et al., 2017; Nour et al., 2016; Rocha & Kim, 2019; Stephenson et al., 2017). Mainly including device-measured outcomes will lead to a more comprehensive picture of intervention effectiveness even though other challenges arise from that approach (e.g. comparison of different epoch lengths (Fabre et al., 2020)). The most promising aspect of device-measured outcomes and accelerometry in particular is the assessment of valid PA and SB data in real-time, resulting in a variety of outcome parameters which have the potential to be easily compared throughout different studies (Burchartz et al., 2020). While device-measured assessments for HE are rarely used (but becoming more and more available (Bandodkar & Wang, 2014)), HE interventions should only include validated tools and be aware of the advantages of each assessment to ensure the quality of results (Rollo et al., 2016).

In order to analyze the influences of different intervention aspects on effectiveness, a uniform and full reporting of the intervention components (theoretical foundation, BCTs, social aspects, etc.), methods and outcomes is needed (Buckingham et al., 2019; Direito et al., 2017; Ferrer & Ellis, 2017; Rocha & Kim, 2019; Schoeppe et al., 2016; Stephenson et al., 2017). Additionally, an exploration of the adequate dose and length of interventions (Böhm et al., 2019; Direito et al., 2017; Ferrer & Ellis, 2017; Nour et al., 2016; Schoeppe et al., 2016; Stephenson et al., 2017), the influence of social support (Böhm et al., 2019; Ferrer & Ellis, 2017; Hamel et al., 2011; McIntosh et al., 2017; Schoeppe et al., 2016) as well as individual tailoring (e.g. using JITAIs to deliver sex-, age- or BMI-specific interventions adapting to personal preferences) (Böhm et al., 2019; Hamel et al., 2011; Rocha & Kim, 2019; Schoeppe et al., 2016) is needed for a better understanding of the determinants of effectiveness. Here, machine learning principles can enhance intervention effectiveness by allowing a highly personalized adaptation to the users' needs and environmental requirements (Gonul et al., 2019a). Future e/mHealth studies for behavior change should also conduct a priori power analyzes to include appropriate sample sizes in order to enhance the value of the results (Böhm et al., 2019; Buckingham et al., 2019; McIntosh et al., 2017; Schoeppe et al., 2016) and assess cost-effectiveness (Böhm et al., 2019; Direito et al., 2017; Nour et al., 2016).

Conclusions

In summary, e/mHealth interventions can be effective tools for primary prevention in behavior change of PA, SB, and HE, but the evidence for effectiveness is still limited. Theoretical foundation and the use of BCTs are promising determinants of effectiveness. However, there is still a research gap which theory and which BCTs are the most promising for primary prevention and for the inclusion of social contexts, JITAIs, and other facets like the optimal dose and length of interventions. Therefore, future studies should limit methodological issues (e.g. non-validated tools) and use appropriate assessments (depending on the outcome variable of choice), and a more comprehensive and standardized way of reporting. In doing so, the benefit of the main advantages of e/mHealth, namely the large coverage, potential cost effectiveness, and high adaptability to individual preferences and environmental facets, can be utilized to enhance behavior change in primary prevention.

References

Ajzen, E. (1991). The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211.

Aune, D., Giovannucci, E., Boffetta, P., Fadnes, L. T., Keum, N., Norat, T., Greenwood, D. C., Riboli, E., Vatten, L. J., & Tonstad, S. (2017). Fruit and vegetable intake and the risk of cardiovascular disease, total cancer and all-cause mortality-a systematic review and dose-response meta-analysis of prospective studies. *International Journal of Epidemiology*, *46*(3), 1029–1056. https://doi.org/10.1093/ije/dyw319

Baird, J., Jacob, C., Barker, M., Fall, C. H. D., Hanson, M., Harvey, N. C., Inskip, H. M., Kumaran, K., & Cooper, C. (2017). Developmental Origins of Health and Disease: A Lifecourse Approach to the Prevention of Non-Communicable Diseases. *Healthcare (Basel*, *Switzerland)*, *5*(1). https://doi.org/10.3390/healthcare5010014

Bandodkar, A. J., & Wang, J. (2014). Non-invasive wearable electrochemical sensors: a review. *Trends in Biotechnology*, *32*(7), 363–371. https://doi.org/10.1016/j.tibtech.2014.04.005

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist*, *44*(9), 1175.

Bandura, A. (1998). Health promotion from the perspective of social cognitive theory. *Psychology & Health*, *13*(4), 623–649. https://doi.org/10.1080/08870449808407422

Bertoncello, C., Colucci, M., Baldovin, T., Buja, A., & Baldo, V. (2018). How does it work? Factors involved in telemedicine home-interventions effectiveness: A review of reviews. *PloS One*, *13*(11), e0207332. https://doi.org/10.1371/journal.pone.0207332

Biddle, S. J. H., Petrolini, I., & Pearson, N. (2014). Interventions designed to reduce sedentary behaviours in young people: A review of reviews. *British Journal of Sports Medicine*, *48*(3), 182–186. https://doi.org/10.1136/bjsports-2013-093078

Blair, S. N. (2009). Physical inactivity: the biggest public health problem of the 21st century. *British Journal of Sports Medicine*, *73*(1), 1–2.

Böhm, B., Karwiese, S. D., Böhm, H., & Oberhoffer, R. (2019). Effects of mobile health including wearable activity trackers to increase physical activity outcomes among healthy children and adolescents: Systematic review. *JMIR MHealth and UHealth*, *7*(4), e8298. https://doi.org/10.2196/mhealth.8298

Brand, T., Pischke, C. R., Steenbock, B., Schoenbach, J., Poettgen, S., Samkange-Zeeb, F., & Zeeb, H. (2014). What works in community-based interventions promoting physical activity and healthy eating?

A review of reviews. *International Journal of Environmental Research and Public Health*, *11*(6), 5866–5888. https://doi.org/10.3390/ijerph110605866

Buckingham, S. A., Williams, A. J., Morrissey, K., Price, L., & Harrison, J. (2019). Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: A systematic review. *Digital Health*, *5*, 1-50. https://doi.org/10.1177/2055207619839883

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

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences: Second Edition*. Erlbaum.

Deeks, Jonathan J, Higgins Julian PT, Altman Douglas G and Cochrane Statistical Methods Group. (2019). *Chapter 10: Analysing data and undertaking meta-analyses: In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019).* https://training.cochrane.org/handbook/current/chapter-10

Direito, A., Carraca, E., Rawstorn, J., Whittaker, R., & Maddison, R. (2017). mHealth Technologies to Influence Physical Activity and Sedentary Behaviors: Behavior Change Techniques, Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Annals of Behavioral Medicine*, *51*(2), 226–239. https://doi.org/10.1007/s12160-016-9846-0

dos Santos, P. C., Barbosa Filho, V. C., da Silva, J. A., Bandeira, A. d. S., Minatto, G., & da Silva, K. S. (2019). What Works in Sedentary Behavior Interventions for Youth: A Review of Reviews. *Adolescent Research Review*, *4*(3), 267–292. https://doi.org/10.1007/s40894-018-0105-4

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

Fabre, N., Lhuisset, L., Bernal, C., & Bois, J. (2020). Effect of epoch length on intensity classification and on accuracy of measurement under controlled conditions on treadmill: Towards a better understanding of accelerometer measurement. *PloS One*, *15*(1), e0227740. https://doi.org/10.1371/journal.pone.0227740

Ferrer, D. A., & Ellis, R. (2017). A Review of Physical Activity Interventions Delivered via Facebook. *Journal of Physical Activity & Health*, *14*(10), 823–833. https://doi.org/10.1123/jpah.2016-0534

Fusar-Poli, P., & Radua, J. (2018). Ten simple rules for conducting umbrella reviews. *Evidence-Based Mental Health*, *21*(3), 95–100. https://doi.org/10.1136/ebmental-2018-300014

Glanz, K., Rimer, B. K., & Viswanath, K. (Eds.). (2008). *Health behavior and health education: theory, research, and practice*. John Wiley & Sons.

Global digital population 2020 | Statista. (2020, April). https://www.statista.com/statistics/617136/digital-population-worldwide/

Gonul, S., Namli, T., Huisman, S., Laleci Erturkmen, G. B., Toroslu, I. H., & Cosar, A. (2019b). An expandable approach for design and personalization of digital, just-in-time adaptive interventions. *Journal of the American Medical Informatics Association: JAMIA*, *26*(3), 198–210. https://doi.org/10.1093/jamia/ocy160

Greaves, C. J., Sheppard, K. E., Abraham, C., Hardeman, W., Roden, M., Evans, P. H., & Schwarz, P. (2011). Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. *BMC Public Health*, *11*(1), 1–12. https://doi.org/10.1186/1471-2458-11-119

Hall, A. K., Cole-Lewis, H., & Bernhardt, J. M. (2015). Mobile text messaging for health: A systematic review of reviews. *Annual Review of Public Health*, *36*, 393–415. https://doi.org/10.1146/annurev-publhealth-031914-122855

Hamel, L. M., Robbins, L. B., & Wilbur, J. (2011). Computer- and web-based interventions to increase preadolescent and adolescent physical activity: a systematic review. *Journal of Advanced Nursing*, *67*(2), 251–268. https://doi.org/10.1111/j.1365-2648.2010.05493.x

Hardeman, W., Houghton, J., Lane, K., Jones, A., & Naughton, F. (2019). A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. *The International Journal of Behavioral Nutrition and Physical Activity*, *16*(1), 31. https://doi.org/10.1186/s12966-019-0792-7

Heron, K. E., & Smyth, J. M. (2010). Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. *British Journal of Health Psychology*, *15*(Pt 1), 1–39. https://doi.org/10.1348/135910709X466063

Hood, M., Wilson, R., Corsica, J., Bradley, L., Chirinos, D., & Vivo, A. (2016). What do we know about mobile applications for diabetes self-management? A review of reviews. *Journal of Behavioral Medicine*, *39*(6), 981–994. https://doi.org/10.1007/s10865-016-9765-3

Kitsiou, S., Paré, G., Jaana, M., & Gerber, B. (2017). Effectiveness of mHealth interventions for patients with diabetes: An overview of systematic reviews. *PloS One*, *12*(3), e0173160. https://doi.org/10.1371/journal.pone.0173160

Marcolino, M. S., Oliveira, J. A. Q., D'Agostino, M., Ribeiro, A. L., Alkmim, M. B. M., & Novillo-Ortiz, D. (2018). The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. *JMIR MHealth and UHealth*, *6*(1), e23. https://doi.org/10.2196/mhealth.8873

Martin, C. A., Rivera, D. E., Riley, W. T., Hekler, E. B., Buman, M. P., Adams, M. A., & King, A. C. (2014). A dynamical systems model of Social Cognitive Theory. In *American Control Conference (ACC), 2014: 4-6 June 2014, Portland, Oregon, USA.* IEEE. https://doi.org/10.1109/acc.2014.6859463

McClung, H. L., Ptomey, L. T., Shook, R. P., Aggarwal, A., Gorczyca, A. M., Sazonov, E. S., Becofsky, K., Weiss, R., & Das, S. K. (2018). Dietary Intake and Physical Activity Assessment: Current Tools, Techniques, and Technologies for Use in Adult Populations. *American Journal of Preventive Medicine*, *55*(4), e93-e104. https://doi.org/10.1016/j.amepre.2018.06.011

McIntosh, J. R. D., Jay, S., Hadden, N., & Whittaker, P. J. (2017). Do E-health interventions improve physical activity in young people: a systematic review. *Public Health*, *148*, 140–148. https://doi.org/10.1016/j.puhe.2017.04.001

Michie, S., Ashford, S., Sniehotta, F. F., Dombrowski, S. U., Bishop, A., & French, D. P. (2011). A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: The CALO-RE taxonomy. *Psychology & Health*, *26*(11), 1479–1498. https://doi.org/10.1080/08870446.2010.540664

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M. P., Cane, J., & Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*, *46*(1), 81–95. https://doi.org/10.1007/s12160-013-9486-6

Moher, D., Liberati, A., Tatzlaff, J., Altman, D. G., & Prisma Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264-269.

Moore, L. V., & Thompson, F. E. (2015). Adults meeting fruit and vegetable intake recommendations—United States, 2013. *MMWR. Morbidity and Mortality Weekly Report*, *64*(26), 706.

Morrison, L. G., Yardley, L., Powell, J., & Michie, S. (2012). What design features are used in effective e-health interventions? A review using techniques from Critical Interpretive Synthesis. *Telemedicine Journal and E-Health: The Official Journal of the American Telemedicine Association*, *18*(2), 137–144. https://doi.org/10.1089/tmj.2011.0062

Muellmann, S., Forberger, S., Moellers, T., Broering, E., Zeeb, H., & Pischke, C. R. (2018). Effectiveness of eHealth interventions for the promotion of physical activity in older adults: A systematic review. *Preventive Medicine*, *108*, 93–110. https://doi.org/10.1016/j.ypmed.2017.12.026

Naszay, M., Stockinger, A., Jungwirth, D., & Haluza, D. (2018). Digital age and the Public eHealth perspective: Prevailing health app use among Austrian Internet users. *Informatics for Health & Social Care*, *43*(4), 390–400. https://doi.org/10.1080/17538157.2017.1399131

Navarro-Barrientos, J.-E., Rivera, D. E., & Collins, L. M. (2011). A dynamical model for describing behavioural interventions for weight loss and body composition change. *Mathematical and Computer Modelling of Dynamical Systems*, *17*(2), 183–203. https://doi.org/10.1080/13873954.2010.520409

Nielsen, S. J., Rossen, L. M., Harris, D. M., & Ogden Cynthia L. (2014). *Fruit and vegetable consumption of U.S. youth, 2009–2010: NCHS data brief, no 156*. Hyattsville. National Center for Health Statistics.

Nour, M., Chen, J., & Allman-Farinelli, M. (2016). Efficacy and External Validity of Electronic and Mobile Phone-Based Interventions Promoting Vegetable Intake in Young Adults: Systematic Review and Meta-Analysis. *Journal of Medical Internet Research*, *18*(4). https://doi.org/10.2196/jmir.5082

Owen, N., Healy, G. N., Matthews, C. E., & Dunstan, D. W. (2010). Too much sitting: The population health science of sedentary behavior. *Exercise and Sport Sciences Reviews*, *38*(3), 105–113. https://doi.org/10.1097/JES.0b013e3181e373a2

Prestwich, A., Sniehotta, F. F., Whittington, C., Dombrowski, S. U., Rogers, L., & Michie, S. (2014). Does theory influence the effectiveness of health behavior interventions? Meta-analysis. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, *33*(5), 465–474. https://doi.org/10.1037/a0032853

Prochaska, James O., and Wayne, F. Velicer (1997). The transtheoretical model of health behavior change. *Amercian Journal of Health Promotion*, *12*(1), 38–48.

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., Brüβler, 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

Rocha, N. P. R., & Kim, H. (2019). eHealth Interventions for Fruit and Vegetable Intake: A Meta-Analysis of Effectiveness. *Health Education & Behavior: The Official Publication of the Society for Public Health Education.* Advance online publication. https://doi.org/10.1177/1090198119859396

Rollo, M. E., Williams, R. L., Burrows, T., Kirkpatrick, S. I., Bucher, T., & Collins, C. E. (2016). What Are They Really Eating? A Review on New Approaches to Dietary Intake Assessment and Validation. *Current Nutrition Reports*, *5*(4), 307–314. https://doi.org/10.1007/s13668-016-0182-6

Ross, J., Stevenson, F., Lau, R., & Murray, E. (2016). Factors that influence the implementation of e-health: A systematic review of systematic reviews (an update). *Implementation Science: IS*, *11*(1), 146. https://doi.org/10.1186/s13012-016-0510-7

Ryan, R. M., & Deci, Edward, L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68.

Schembre, S. M., Liao, Y., Robertson, M. C., Dunton, G. F, Kerr, J., Haffey, M. E., Burnett, T., Basen-Engquist, K., & Hicklen, R. S. (2018). Just-in-Time Feedback in Diet and Physical Activity Interventions: Systematic Review and Practical Design Framework. *Journal of Medical Internet Research*, *20*(3), e106. https://doi.org/10.2196/jmir.8701

Schoeppe, S., Alley, S., van Lippevelde, W., Bray, N. A., Williams, S. L., Duncan, M. J., & Vandelanotte, C. (2016). Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. *The International Journal of Behavioral Nutrition and Physical Activity*, *13*. https://doi.org/10.1186/s12966-016-0454-y

Sharif, M. O., Janjua-Sharif, F. N., Hesham, A., & Farooq, A. (2013). Systematic Reviews Explained: AMSTAR—How to Tell the Good From the Bad and the Ugly. *Oral Health Dent Manag.*, *12*(1), 9–16.

Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., Porter, A. C., Tugwell, P., Moher, D., & Bouter, L. M. (2007). Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology*, *7*(1), 10. https://doi.org/10.1186/1471-2288-7-10

Stephenson, A., McDonough, S. M., Murphy, M. H., Nugent, C. D., & Mair, J. L. (2017). Using computer, mobile and wearable technology enhanced interventions to reduce sedentary behaviour: a systematic review and meta-analysis. *The International Journal of Behavioral Nutrition and Physical Activity*, *14*(105). https://doi.org/10.1186/s12966-017-0561-4

Tremblay, Mark S., LeBlanc, A. G., Kho, Michelle E., Saunders, Travis J., Larouche, R., Colley, Rachel, C., Goldfield, G., & Gorber, S. C. (2011). Systematic review of sedentary behaviour and health indicators in school-aged children and youth. *The International Journal of Behavioral Nutrition and Physical Activity*, *8*(98). https://doi.org/10.1186/1479-5868-8-98

Umberson, D., Crosnoe, R., & Reczek, C. (2010). Social Relationships and Health Behavior Across Life Course. *Annual Review of Sociology*, *36*, 139–157. https://doi.org/10.1146/annurev-soc-070308-120011

Vandelanotte, C., Müller, A. M., Short, C. E., Hingle, M., Nathan, N., Williams, S. L., Lopez, M. L., Parekh, S., & Maher, C. A. (2016). Past, Present, and Future of eHealth and mHealth Research to Improve Physical Activity and Dietary Behaviors. *Journal of Nutrition Education and Behavior*, *48*(3), 219-228. https://doi.org/10.1016/j.jneb.2015.12.006

Vandelanotte, C., Spathonis, K. M., Eakin, E. G., & Owen, N. (2007). Website-delivered physical activity interventions a review of the literature. *American Journal of Preventive Medicine*, *33*(1), 54–64. https://doi.org/10.1016/j.amepre.2007.02.041

Villinger, K., Wahl, D. R., Boeing, H., Schupp, H. T., & Renner, B (2019). The effectiveness of app-based mobile interventions on nutrition behaviours and nutrition-related health outcomes: A systematic review and meta-analysis. *Obesity Reviews: An Official Journal of the International Association for the Study of Obesity*, *20*(10), 1465–1484. https://doi.org/10.1111/obr.12903

Viner, R. M., Ozer, E. M., Denny, S., Marmot, M., Resnick, M., Fatusi, A., & Currie, C. (2012). Adolescence and the social determinants of health. *The Lancet*, *379*(9826), 1641–1652. https://doi.org/10.1016/S0140-6736(12)60149-4

Webb, T. L., Joseph, J., Yardley, L., & Michie, S. (2010). Using the internet to promote health behavior change: A systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *Journal of Medical Internet Research*, *12*(1), e4. https://doi.org/10.2196/jmir.1376

Woll, A, Kurth, B.-M., Opper, E., Worth, A., & Bös, K. (2011). The 'Motorik-Modul' (MoMo): Physical fitness and physical activity in German children and adolescents. *European Journal of Pediatrics*, *170*(9), 1129–1142. https://doi.org/10.1007/s00431-010-1391-4

World Health Organization. (2011). *Mhealth: Second Global Survey on eHealth*. World Health Organization. http://gbv.eblib.com/patron/FullRecord.aspx?p=851142

World Health Organization. (2014). *Global Status Report on Noncommunicable Diseases 2014*. World Health Organization. http://gbv.eblib.com/patron/FullRecord.aspx?p=2059287

World Health Organization. (2018). *Global Health Estimates 2016: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2016*. https://www.who.int/healthinfo/global_burden_disease/estimates/en/

World Health Organization. (2020, March 19). *eHealth at WHO*. https://www.who.int/ehealth/en/

Wunsch, K, Eckert, T, Fiedler, J., Cleven, L, Niermann, C, Reiterer, H, Renner, B, & Woll, A (2020). Effects of a Collective Family-Based Mobile Health Intervention Called "SMARTFAMILY" on Promoting Physical Activity and Healthy Eating: Protocol for a Randomized Controlled Trial. In: *JMIR research protocols* 9 (11), e20534. DOI: 10.2196/20534.

Chapter 3 Choice of physical activity measurement

Paper 2: Comparison of Self-Reported and Device-Based Measured Physical Activity Using Measures of Stability, Reliability, and Validity in Adults and Children.

Slightly modified version of the published manuscript

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). https://doi.org/10.3390/s21082672

Abstract

Quantification of physical activity (PA) depends on the type of measurement and analysis method making it difficult to compare adherence to PA guidelines. Therefore, test-retest reliability, validity, and stability for self-reported (i.e. questionnaire and diary) and device-based measured (i.e. accelerometry with 10/60 second epochs) PA was compared in 32 adults and 32 children from the SMART*FAMILY* study to examine if differences in these measurement tools are systematic. PA was collected during two separate measurement weeks and the relationship for each quality criteria was analyzed using Spearman correlation. Results showed the highest PA values for questionnaires followed by 10-second and 60-second epochs measured by accelerometers. Levels of PA were lowest when measured by diary. Only accelerometry demonstrated reliable, valid, and stable results for the two measurement weeks, the questionnaire yielded mixed results and the diary showed only few significant correlations. Overall, higher correlations for the quality criteria were found for moderate than for vigorous PA and the results differed between children and adults. Since the differences were not found to be systematic, the choice of measurement tools should be carefully considered by anyone working with PA outcomes, especially if vigorous PA is the parameter of interest.

Keywords: self-report, device-based measured, physical activity, reliability, validity, stability

Introduction

Insufficient physical activity (PA) is a high-risk factor for non-communicable diseases in modern society (Lee et al., 2012) and is linked to an overall increased mortality rate (Kohl et al., 2012). This in turn leads to a high economic burden worldwide (Ding et al., 2016) and calls for a systematic approach to increase PA. To counteract the insufficient PA levels, the world health organization (WHO) has continuously put PA guidelines into place (Bull et al., 2020). One of the main challenges with these guidelines is to classify someone as sufficiently active since PA can be measured in various ways yielding different outcomes (Dyrstad et al., 2014; Hagstromer et al., 2010; Hukkanen et al., 2018; Skender et al., 2016). Unfortunately, there is no basic solution as van Hees described in a recent blog post (van Hees, 2021). According to him, PA is defined by the measurement method used and guidelines represent the average results of a variety of different methods which is not feasible to apply to the conception of intervention studies (van Hees, 2021). This is especially important if the study aims to use a personalized approach like just-in-time adaptive interventions (Hardeman et al., 2019) or aims to compare the result to common PA guidelines which both strongly depend on comparable data. The previous World Health Organization (WHO) guidelines (World Health Organization, 2010) were mainly based on data revealed by studies using subjective assessment methods, e.g. questionnaire data from the Global Physical Activity Questionnaire (Armstrong & Bull, 2006) and the frequently used International Physical Activity Questionnaire (IPAQ) (Craig et al., 2003) with both showing low to strong correlations (lower correlations for moderate PA (MPA) than vigorous PA (VPA)) in previous studies (Fiona C. Bull et al., 2009; Herrmann et al., 2013). Using self-report assessments is convenient in large samples but the results are inconsistent due to either over- or underreporting of PA (Skender et al., 2016). To counteract recall-bias, other self-report, as well as methods like PA diaries and, with new technological advances, ecological momentary assessments are used due to their timeliness and a smaller Black-box due to multiple measurements, which in turn increases the burden for participants due to more frequent reports (Dunton, 2017; Hukkanen et al., 2018; Sattler et al., 2021). The most recent WHO guidelines adapted the recommendations according to findings from studies using accelerometry, pedometer, and other device-based measurements which found that PA bouts of less than 10 minutes (not monitored by most PA questionnaires) also qualify to boost health benefits (Bull et al., 2020; Troiano et al., 2020) while they remove the recall bias (Anastasopoulou et al., 2014).

Furthermore, the rising interest in short-time intermitted V PA for health benefits with a reduced time requirement is especially difficult to monitor using self-reports (Stamatakis et al., 2020). However, to gain reliable measurements of PA, the sensors should be worn long enough to accurately represent the measurement duration of interest (e.g. eight hours a day for at least four days can accurately represent one measurement week) (Burchartz et al., 2020; Stamatakis et al., 2020). Even though a wear time of 24 hours per day is described to be the most accurate assessment method for overall behavior throughout the day (i.e. sleep, sedentary behavior, and PA) (Burchartz et al., 2020), accelerometers are rarely attached to the body for this duration and therefore measured PA can be impacted by wear-time bias (a wear-time of 8 to 10 hours is commonly assumed to be sufficient (Burchartz et al., 2020) but PA can occur during the non-wear-time and therefore PA is likely to be underestimated as compared to real PA during 24 hours). While wearables like Fitbit can easily be attached to the body for 24 hours, they are mainly designed for commercial purposes, show limited validity and reliability, and can only provide accurate step counts in adults under certain conditions (no mobility limitation and worn at the torso) but not for energy estimation (Feehan et al., 2018). In accelerometers, the use of different sensors, algorithms, cut points (point which determines PA intensity) and epoch lengths (e.g. raw data, 1 second, 10 seconds, 60 seconds) used in measurement and analysis of PA has a high impact on PA estimations, depends on the age group (e.g. recommendation for the use of shorter epochs in children (Edwardson & Gorely, 2010)) and complicates comparison between different studies (Edwardson & Gorely, 2010; Migueles et al., 2019; Orme et al., 2014). Here, the choice of epoch length is especially important in detecting VPA and inactivity in children (Edwardson & Gorely, 2010) due to their highly intermittent PA patterns (Livingstone et al., 2003). Even though PA patterns in adults are often linear (i.e. less short duration and high-intensity PA) and PA, therefore is believed to be not as susceptible as children's PA, the use of different epoch lengths alone can also change moderate to vigorous PA estimations in adults due to the smoothing of PA intensity with the use of longer epochs (i.e. 10-second vs 60-second epochs) (Edwardson & Gorely, 2010; Orme et al., 2014). Thus, each approach has its own challenges and there is currently no gold standard to measure PA if using accelerometry, questionnaires, or diaries (Slootmaker et al., 2009), even though best practices to handle these issues are currently discussed (Burchartz et al., 2020). Comparing these measurement methods (i.e. accelerometry, diary, and questionnaire) is therefore challenging and requires further data on their relationship between each measurement method and it is important to evaluate if the relationship is consistent over time (i.e. over two measurement periods) under consideration of multiple aspects.

To gain further insights into the issue at hand, statistical quality criteria of the different methods have to be considered (Patterson, 2000). Thereby, test-retest reliability (in the following referred to as reliability) shows good to excellent intraclass correlation coefficients for accelerometry (for everyday activity) (Sirad, John, R. et al., 2011). Furthermore, the evaluation of the reliability of self-report questionnaires (i.e. GPAQ (Herrmann et al., 2013) and IPAQ (Craig et al., 2003)) and a PA diary (Williams et al., 1989) also indicate good to excellent reliability, while others found poor reliability in some metrics of the GPAQ (Rivière et al., 2018). Measures of agreement (in the following referred to as validity) expressed by the correlation between self-reported and device-based measured PA, show an overall low agreement, are influenced by age and gender and self-reported results are often overreported when compared to device-based measurements especially for VPA (Fiona C. Bull et al., 2009; Herrmann et al., 2013; Skender et al., 2016). One study which compared PA data measured by accelerometry, diary, questionnaire, and interview in adults (*N*=1916) found that the comparison between the device-based measured and self-reported meeting of PA recommendation at one measurement period yielded only 12% agreement based on pairwise comparisons (Hukkanen et al., 2018). Other studies that analyzed test-retest reliability at several timepoints and included diaries (Pols, Margreet, A., Peeters, Peter, H., M., Ocké, Marga, C., Bueno-de-Mesquita, H., Bas, et al., 1997; Pols, Margreet, A., Peeters, Peter, H., M., Ocké, Marga, C., Slimani, et al., 1997) or accelerometer (Cust, Anne, E. et al., 2008; Lubans, David, R. et al., 2008; Trinh, Oanh, T., H. et al., 2009) to analyze the validity at one or all measurement periods also reported good test-retest reliability but only acceptable or comparable validity showing that comparing PA results of different measurement methods should be done with caution. These discrepancies indicate the difficulty of the interpretation for sufficient PA using different methods.

However, even though differences between measurement tools are frequently reported throughout several studies (Dyrstad et al., 2014; Hagstromer et al., 2010; Hukkanen et al., 2018; Slootmaker et al., 2009), there are, to the best of our knowledge, currently, no studies analyzing whether these differences are systematic (i.e. high correlation for the paired differences between the measurement methods between two separate measurement weeks) for PA levels measured via accelerometry, diary, and questionnaire for adults, children, and adolescents (in the following referred to as children). If these differences could be shown to be consistent over time, this would strengthen the interpretation and comparison, and use of different PA data in intervention studies, in between different studies, and regarding PA guidelines. Here, longitudinal data may represent a more consistent picture of PA, allowing to detect time-stable differences regarding the amount of PA between the different methods.

Therefore, the current study aimed to examine the stability of the pairwise differences between three PA measurement methods (i.e. accelerometry, diary, and questionnaire) and the influence of different evaluation techniques (i.e. epoch lengths of 10 seconds and 60 seconds for accelerometer data for MPA and VPA in adults and children between two independent measurement weeks in an explorative manner. A secondary aim was to analyze the reliability of the above-mentioned measurement methods and to assess the validity of those methods.

Materials and Methods

Participants and procedure

Participants were eligible for this study if they represented a family with at least one child and one adult who were living in a common household. In total, 74 adults and 74 children participated in the SMART*FAMILY* (SF) trial which consists of a theory- and evidence-based mHealth intervention and targets health behavior change in families (further information are described in the study protocol (Wunsch et al., 2020)) and all participants of the control group (32 adults age 37 – 55 years and 32 children age 5 – 19 years) were eligible for the present study. Full ethical approval was obtained for SF. All participants, children, and legal guardians provided written informed consent prior to commencing the study by signing the informed consent form (The International Registered Report Identifier (IRRID) for the SF study is RR1-10.2196/20534.). The trial was conducted in accordance with the Declaration of Helsinki.

Participants were recruited in schools, school holiday programs, music schools, sports clubs, via personal communication and via newspapers and email-distribution lists. Participants were cluster-randomized to an intervention group and a control group. Whereas the intervention group received a three-week mobile health intervention between the two measurements, families of the control group had a three-week waiting period without any intervention. Baseline (T_0) and postintervention (T_1) data of the control group were used for this study because the intervention might have influenced PA sampling at T_1. Data collection at T_0 and T_1 consisted in the measurement of PA by accelerometer, diary, and questionnaire over one week which was identical for T_0 and T_1 (which were at least three weeks apart). For children, the inclusion of questionnaire data was not feasible for this study due to the use of a questionnaire without the indication of minutes per week for PA (Sixty-Minute Screening Measure (Prochaska, Judith, J. et al., 2001)) which is also not comparable to the new PA guidelines which recommend an average of 60 minutes PA per day for children (Bull et al., 2020).

Measurements

Accelerometer

Hip-worn (right side) 3-axial accelerometers (Move 3 / Move 4, Movisens GmbH, Karlsruhe, Germany) were used to continuously record PA (see supplement Figure 1). These accelerometers are scientific research instruments with a measurement range of ±16g, an output rate of 64 Hz, physical dimensions of 62,3mm x 38,6mm x 11,5mm, weight of 25g, and custom epoch lengths (i.e. 10 sec and 60 sec). Data is recorded in a rare format (64 Hz) and afterward summarized in the epoch lengths of choice. Epoch lengths were chosen to represent the most common used epoch length (60 sec) which was mainly used due to limited storage in the past, and a shorter epoch length (10 sec) as shorter epoch lengths are believed to be more appropriate to estimate VPA and to assess PA in children due to intermittent movement behavior (Burchartz et al., 2020). Validity has been evaluated for a previous version of the accelerometer (Move 2) which uses comparable digital signal processing as the move 3/4 (Jörg Ottenbacher, personal communication, March 16, 2021) and has been considered accurate for assessing steps (Anastasopoulou et al., 2013) and energy estimation (Anastasopoulou et al., 2014; Härtel et al., 2011) in adults. Handling of the accelerometer was explained and demonstrated by a study instructor and participants were instructed to wear the accelerometer during wake-time and to remove it only for taking a shower, swimming or during certain sports involving bodily contact to minimize the probability of injuries. Outcomes for the accelerometer which were used for this study were MPA (3.0-5.9 MET) and VPA (≥ 6 MET) (light PA was not considered because the questionnaire has no comparable measure) for all participants. MET values were calculated based on activity class (based on acceleration and barometric signals) which determines the estimation model. Afterwards, movement acceleration, altitude change, and demographics were combined in the model for the MET estimation (Härtel et al., 2011) (see supplement Figure 2).

Accelerometer data were included if a minimum wear time of at least 8 hours a day for at least 4 of the 7 days during the measured week was obtained. Non-wear time was calculated on the accelerometer in 30-second intervals. The non-wear time detection was based on an algorithm that used accelerometry and temperature signal over a 10-minute window to distinguish between wear time, non-wear time, and sleep as described elsewhere (Barouni et al., 2020). For valid measurements, the average of MPA and VPA per valid day was multiplied by 7 to represent the total minutes per week.

Diary

A daily PA diary was filled in by all participants complementary to wearing accelerometers indicating date, time and type of activity, duration, and perceived intensity on every single day within the two measurement weeks. Each activity was recorded separately and participants were instructed to rate the respective PA intensity as either light (no perspiration or shortness of breath), moderate (some perspiration and shortness of breath) or vigorous (profound perspiration and shortness of breath). Participants were asked to report all PA with a duration of more than 10 minutes. Analogous to accelerometry outcomes, MPA and VPA were summarized as total minutes per week.

Questionnaire

At the end of each measurement week, adults were asked to fill in the German short version of the IPAQ (Mäder et al., 2006) which is available at the IPAQ website (*Downloadable Questionnaires - International Physical Activity Questionnaire*, 2021), asking retrospectively for activities during the previous week. The results of the question relating to minutes spent in MPA (comprising of moderate activity and walking (IPAQ Research Committee)) and VPA were calculated for this study by multiplying the reported amount of days with the reported duration of the indicated activity per day. Therefore, the outcomes MPA and VPA were also recorded as total minutes per week. Children completed the Sixty-Minute Screening Measure (Prochaska, Judith, J. et al., 2001) for moderate to vigorous physical activity which yields binary results (sufficiently active vs insufficiently active according to the previous WHO guidelines (World Health Organization, 2010)) and was not included in this study to maintain total minutes per week as a unit. Therefore, all results referring to the questionnaire are limited to adults. Additionally, questions about age and anthropometry were included in the questionnaire among others (see the study protocol for detailed information (Wunsch et al., 2020)).

Statistical analysis

To compare the mean differences for the four PA measures (i.e. accelerometry with 10 sec and 60-sec epoch lengths, diary, and questionnaire) between T_0 and T_1, the differences between both measurement time points were calculated in total minutes per week for MPA and VPA for all combinations (i.e. six combinations for adults and three combinations for children) and defined as new parameters (ranging from -607.17 to 398.29 min/week) at each measurement week. If one of the original parameters included missing data, the parameter expressing the difference was also considered as missing data for the participant. Additionally, test-retest reliability for each parameter between T_0 and T_1 and a validity measure by pairwise comparison of all parameters at both T_0 and T_1 were calculated (see Figure 1).

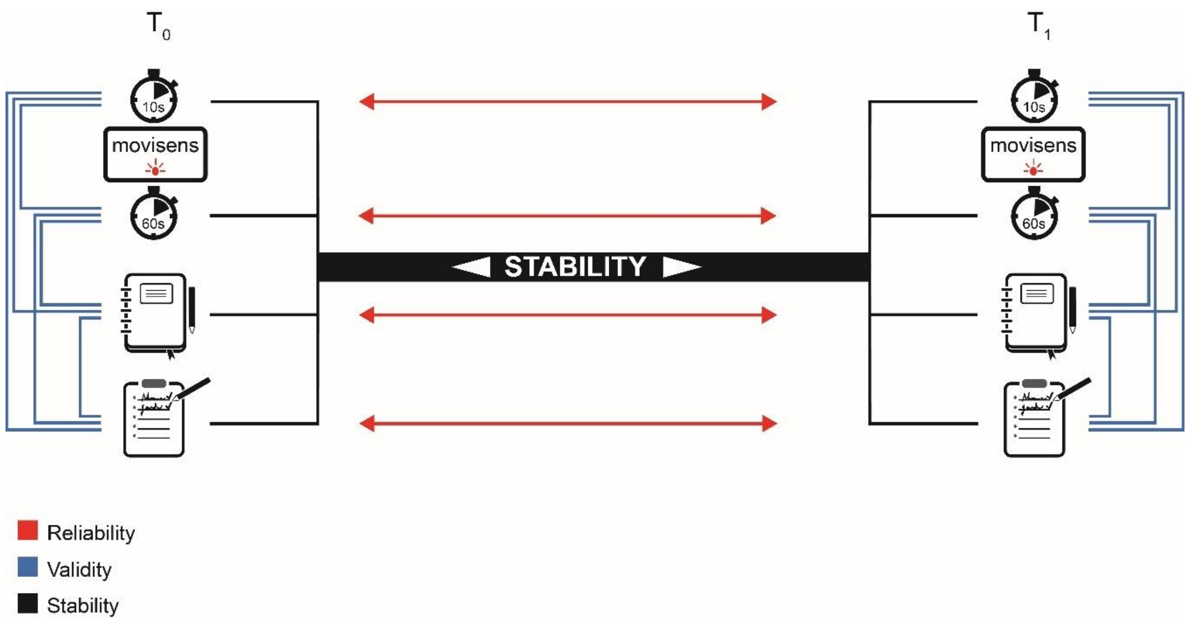

Figure 1. Study design. Displayed are the calculated combinations for validity (blue brackets) and reliability measures (red arrows) for the secondary aims concerning the parameters (from top to bottom) accelerometry using 10-second epochs and 60-second epochs, a physical activity diary, and the International Physical Activity Questionnaire. The main aim consisted in comparing the difference in total minutes for each bracket from T_0 to T_1 (black).

Figure 2 has been created using RStudio (R Core Team, 2020) and the ggplot2 package (Wickam, 2016), following the instructions of Allan and colleagues (Allan et al., 2019). Statistical analyses were performed in RStudio using the RVAideMemoire package (Maxime Hervé, 2020). Descriptive characteristics of all participants are displayed as means with standard deviation (SD). The degree of agreement for all calculations was assessed using the Spearman correlation coefficient (r_s) by the cor.test() function in RStudio since the data differed significantly from a normal distribution in the Kolmogorov Smirnov test, which was confirmed via visual inspection of the distribution in histograms. First, r_s values between T_0 and T_1 were calculated between the pairwise differences of all parameters to indicate the stability of these differences (main aim). Then, r_s values between T_0 and T_1 were computed for each parameter separately to indicate test-retest reliability (secondary aim). Afterwards, r_s values for the pairwise comparison between all combinations of parameters at both T_0 and T_1 were computed for a measure of validity (secondary aim). Afterward, Confidence intervals were added by using bootstrapping (n=1000). All calculations were performed for children and adults separately and pairwise deletion was used for each calculation.

r_s were interpreted under consideration of the 95% confidence intervals as recommended by Schober, Boer, and Schwarte (Schober et al., 2018). The level for significance was set a priori to .05 and was based on the correlation and not on the confidence intervals.

Results

Participant characteristics

The data of 32 adults and 32 children was used in this study. Characteristics of the participants are presented in Table 1.

Table 1. Characteristics of the participants. Displayed are the number of participants (N), means, and standard deviations (SD) for the parameters gender (male/female), age in years (y), height in centimeter (cm), and weight in kilogram (kg).

Parameter	**Adults**		**Children**	
	N	**Mean (SD)**	**N**	**Mean (SD)**
Gender (m/f)	11/21	-	15/17	-
Age (y)	31	47.90 (4.44)	32	13.22 (2.94)
Height (cm)	31	170.42 (8,52)	31	162.68 (17.61)
Weight (kg)	31	72.74 (13.27)	29	51.10 (14.10)
BMI (kg/m²)	31	24.97 (3.62)	29	18.96 (2.94)

Physical activity outcomes

The descriptive results of PA measurements at T_0 and T_1 and corresponding reliability and validity measures (r_s) are presented in supplement 1 (supplement Tables 1 to 4). Figure 2 A/B visualize the descriptive PA level measured by each measurement tool for adults and Figures 2 B/D for children. Overall, the descriptive values show the highest PA values for the IPAQ, followed by accelerometry with 10-second epochs and 60-second epochs, and the lowest PA values are reported for the PA diary. These results are consistent for MPA and VPA in both adults and children except for VPA in children where the PA diary shows the highest PA values. MPA in T_0 is higher in all measures compared to T_1 whereas VPA values are only consistently lower in children at T_1.

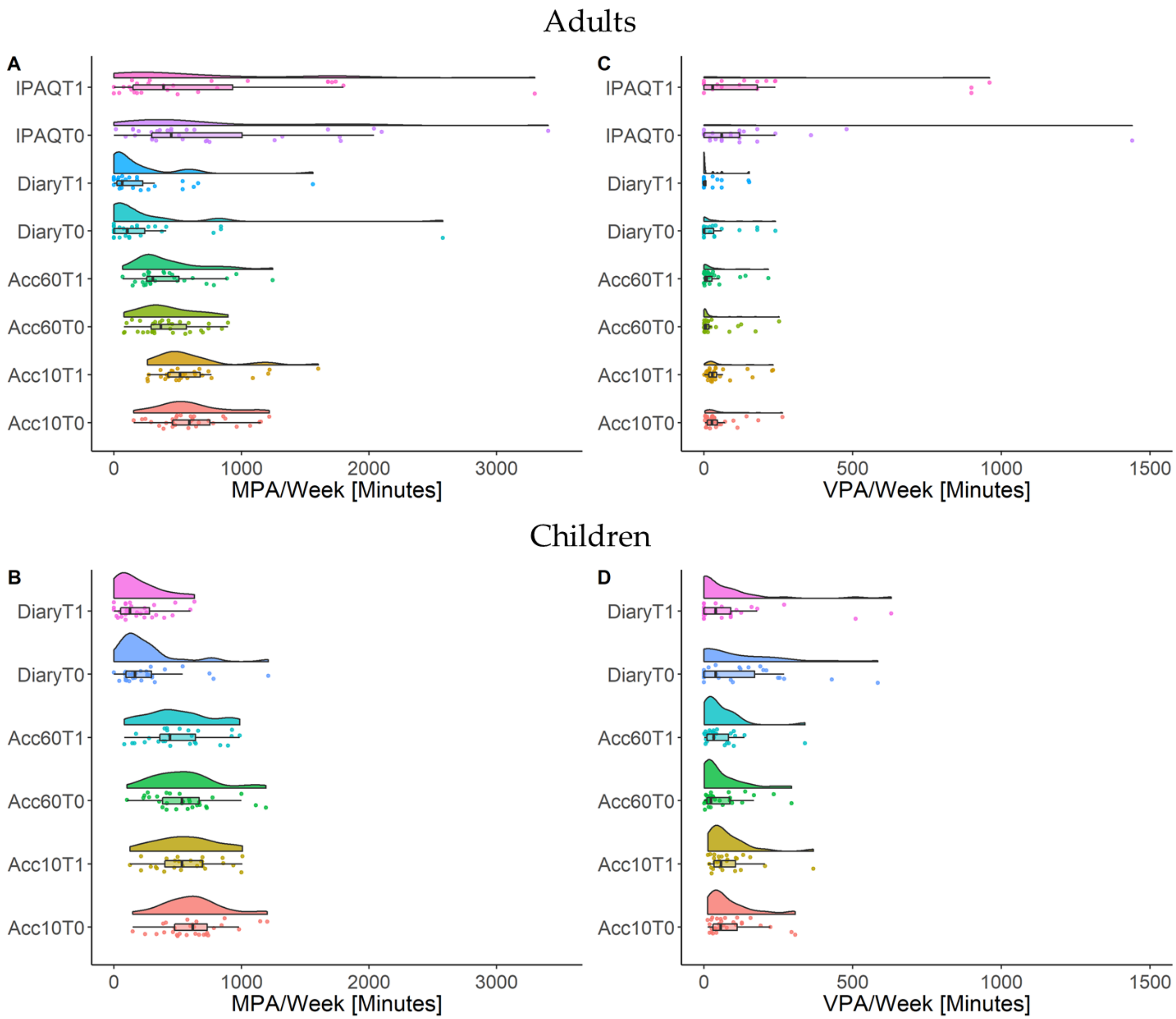

Figure 2. Descriptive means of moderate physical activity (MPA) in adults (A) and children (B) as well as vigorous physical activity (VPA) in adults (C) and children (D). Displayed are the results (independent measurements, distribution, and box-plots) of the physical activity diary (Diary), accelerometry with 60-second epochs (Acc60), accelerometry with 10-second epochs (Acc10), and the International Physical Activity Questionnaire (IPAQ) for two independent measurement weeks (T0 and T1) in minutes per week.

Stability between the differences of the parameters at the two measurement weeks

Table 4 presents the r_s for the differences in minutes per week of all parameters compared from T_0 to T_1 for adults while table 5 shows the results for children.

The differences in the amount of PA gathered by accelerometers using 10 second and 60-second epoch lengths showed a significant relationship for both adults and children in MPA and VPA between T_0 to T_1.

Significant associations of the differences between accelerometry and diary were found for MPA, but not for VPA, measured by 10-second epochs, and PA diary for adults. For children, there was a significant relationship between the differences of accelerometry using 10-second epochs and the PA diary for VPA, but not for MPA, between T_0 and T_1 with a lower confidence limit below zero.

The differences between accelerometry and the IPAQ were significantly related for both 60 second and 10-second epochs concerning MPA but not for VPA.

No significant association at all was found for the differences of the diary and IPAQ between T_0 and T_1.

Table 2. Results of stability of differences calculated by Spearman´s rho (r_s) for adults. Displayed are the number of participants (n), mean differences between the measurement tools (Mean), their standard deviations (SD), and their minimum (Min) and maximum (Max) values at two measurements three weeks apart (T_0 and T_1) as well as the differences in the error between T_1 and T_0 as mean, SD and percentage (%). Additionally, r_s with corresponding *p*-value (* for $p < .05$) and 95% confidence interval via bootstrapping (CI) for differences between the measurement tools: accelerometry with 60 seconds epoch length (Acc 60) and 10 seconds epoch length (Acc 10), physical activity diary (Diary), the International Physical Activity Questionnaire (IPAQ) for moderate (MPA) and vigorous (VPA) physical activity.

		Adults			
	n	**T_0 Mean (SD) [Min-Max]**	**T_1 Mean (SD) [Min-Max]**	**T_1 - T_0 Mean (SD) difference [%]**	**r_s (*p*-value) [CI]**
MPA (min/wk)					
Acc 10 - Acc 60	28	206.47 (74.33) [122.50-456.40]	181.40 (102.78) [-207.67-364.00]	-25.06 (107.39) [-12.93%]	.604 (.001*) [.242-.842]
Acc 10 - Diary	25	394.80 (533.57) [-1513.00-1152.20]	398.29 (514.08) [-1196.60-1566.83]	3.49 (70.37) [1.41%]	.579 (.002*) [.111-.879]
Acc 10- IPAQ	25	-195.42 (764.58) [-2262.00-873.00]	-125.04 (730.91) [-2744.00-631.50]	70.37 (854.72) [43.92%]	.613 (.001*) [.268-.841]
Acc 60 - Diary	25	185.97 (528.88) [-1839.00-755.00]	216.77 (488.43) [-1324.80-1208.67]	30.80 (595.35) [15.30%]	.699 (<.001*) [.314-.890]
Acc 60 - IPAQ	25	-393.48 (779.17) [-2558.00-654.00]	-303.05 (769.65) [-2978.00-622.83]	-107.63 (852.18) [25.97%]	.598 (.001*) [.266-.828]
Diary - IPAQ	24	-607.17 (858.53) [-3224.00-480.00]	-466.08 (743.78) [-2640.00-750.00]	-141.08 (1061.75) [26.29%]	.072 (.737) [-.431-.603]
VPA (min/wk)					
Acc 10 - Acc 60	28	15.39 (11.79) [-16.00-43.00]	23.37 (31.34) [1.00-176.17]	7.98 (32.51) [41.18%]	.569 (.002*) [.162-.858]
Acc 10 - Diary	25	22.74 (87.04) [-222.00-204.00]	35.50 (67.60) [-119.67-228.67]	12.75 (99.56) [43.82%]	.197 (.344) [-.249-.558]
Acc 10 - IPAQ	26	-79.57 (292.69) [-1417.00-129.00]	-117.10 (281.61) [-942.50-228.67]	-37.53 (266.22) [38.17%]	.279 (.167) [-.113-.625]
Acc 60 - Diary	25	6.93 (86.72) [-239.00-193.00]	10.86 (52.11) [-130.17-157.00]	3.93 (89.12) [44.18%]	.155 (.459) [-.266-.558]

Acc 60 - IPAQ	26	-94.96 (292.49) [-1432.00-118.00]	-140.26 (275.82) [-957.67-52.50]	-60.67 (262.90) [38.52%]	.142 (.490) [-.233-.509]
Diary - IPAQ	25	-99.44 (289.09) [-1440.00-120.00]	-152.92 (292.10) [-960.00-17.00]	-53.48 (261.13) [42.38%]	.219 (.293) [-.221-.625]

Table 3. Stability of differences as calculated by Spearman´s rho (r_s) for children. Displayed are the number of participants (n), mean differences between the measurement (Mean), their standard deviations (SD), and their minimum (Min) and maximum (Max) values at two measurements three weeks apart (T_0 and T_1) as well as the differences in the error between T_1 and T_0 as mean, SD and percentage (%). Additionally, r_s with corresponding *p*-value (* for $p < .05$) and 95% confidence interval via bootstrapping (CI) for differences between the measurement tools: accelerometry with 60 seconds epoch length (Acc 60) and 10 seconds epoch length (Acc 10), physical activity diary (Diary) for moderate (MPA) and vigorous (VPA) physical activity.

			Children		
	n	**T_0 Mean (SD) [Min – Max]**	**T_1 Mean (SD) [Min – Max]**	**T_1 - T_0 Mean (SD) difference [%]**	**r_s (*p*-value) [CI]**
MPA (min/wk)					
Acc 10 - Acc 60	24	78.77 (48.73) [-16.33-171.50]	61.10 (54.36) [-69.00-177.00]	-17.67 (37.36) [25.27%]	.785 (<.001*) [.548-.920]
Acc 10 - Diary	23	376.80 (348.08) [-470.00-1037.83]	342.78 (326.38) [-272-979.00]	-34.02 (402.33) [9.46%]	.363 (.090) [-.152-.806]
Acc 60 - Diary	23	299.96 (364.70) [-538.00-1027.33]	283.99 (331.10) [-300.75-967.00]	-15.97 (380.57) [5.47%]	.383 (.072) [-.083-.754]
VPA (min/wk)					
Acc 10 - Acc 60	24	29.43 (19.36) [9.33-102.67]	25.57 (21.25) [5.60-105.00]	-3.86 (12.81) [14.04%]	.448 (.028*) [.010-.762]
Acc 10 - Diary	23	-37.24 (174.17) [-541.25-295.17]	-16.02 (157.83) [-550.67-131.83]	21.23 (207.55) [79.68%]	.417 (.049*) [-.053-.786]
Acc 60 - Diary	23	-67.17 (167.10) [-564.00-234.50]	-40.83 (152.54) [-560.00-105.00]	26.35 (203.06) [48.78%]	.275 (.205) [-.227-.659]

Test-retest reliability

Adults' MPA measured by accelerometry at T_0 shows a significant relationship with MPA measured by accelerometry at T_1 with the lower confidence interval limit of $r_s > .5$ at both epoch lengths. For children only VPA at 10-second epochs showed a similar effect. The other accelerometry measures also show significant correlations but lower confidence limits.

For the diary-based PA, adults PA has no significant relationship between T_0 and T_1 while childrens' PA shows a significant relationship with the lower confidence limit of around 0.

PA measured by the IPAQ at T_0 has a significant relationship with PA measured by the IPAQ at T_1 for MPA with a lower confidence limit of around .1.

Validity

Additional analysis of pairwise r_s between all methods at each T_0 and T_1 showed a significant relation between 10 and 60-second epochs at T_0 and T_1 for both children and adults with the lower confidence limit above .7 for both MPA and VPA (see Supplement Tables 3 and 4). The IPAQ showed a significant relationship to accelerometry for VPA (compared to 10-sec epochs and PA diary) with lower confidence limits of around 0 only at T_0. No further significant relations were found between the parameters at neither measurement week.

Discussion

This study aimed to examine the reliability, validity, and stability of a PA questionnaire, a PA diary, and accelerometry using 10 and 60-second epochs for MPA and VPA in adults and children over two measurement weeks. The main result evoked the stability of differences to be an interesting additional measure for the comparison of different measurement methods not necessarily being in concordance with reliability and measures of validity. Overall, descriptive results consistently showed that self-reports via questionnaire revealed by far the highest PA amounts, followed by accelerometry with 10 and 60-second intervals. The lowest amounts were detected for PA measured via diary for both MPA and VPA in adults and children with a large variance in the results of each measurement tool. Only device-based measured PA showed reliable, valid, and stable results for the two measurement weeks for both epoch lengths. The IPAQ yielded mixed results and the PA diary showed few significant relations for stability in adults and mixed results in children.

Quality criteria

Stability

The comparison of the pairwise differences between T_0 and T_1 showed stable results for almost all comparisons in adults' MPA. For children, stable MPA differences were found for 10 and 60-second epoch lengths and the diary (which did not reach statistical significance but indicates r_s of 0.4). VPA mainly showed a significant correlation between the two epoch lengths in adults and children, while 10-second epochs were only associated with the diary in children. Taken together, the stability results showed significant results for all parameters also demonstrating high reliability and validity, which was to be expected. No stability was found for parameters with low reliability and significant validity (i.e. adults VPA 10 second epochs to IPAQ and diary to IPAQ). However, some measures also showed significant stability where no validity, and in one case where neither reliability nor validity, was found to be significant. This indicates the importance of the relationship over time because these results would have been missed without a stability measure.

These results show that the relation between the measures including self-report differed between the measurement weeks and are therefore not stable over time which gives reason for concern in the comparability of these measures (as indicated by validity as well). However, the comparison of device-measured PA and the diary in children indicates some stability which might show that the diary is more feasible for children than for adults and strengthens the point that children's structured PA might be easier to determine using self-report. Additionally, the descriptive values showed that only the comparison of 10 to 60-second intervals in children yielded minimum and maximum values without a change of signs indicating that these differences were exposed to intraindividual variations for most of the device based measured results and are not 100% consistent even though they are highly related (i.e. comparison of 10 to 60-second intervals). These findings have to be treated with caution and need to be reevaluated due to the limitations listed below.

Reliability

Test-retest reliability of both epoch lengths indicated that the present data of two measurement weeks represented comparable weeks of everyday life concerning PA. This was partially confirmed by the IPAQ (only for MPA) and the PA diary (only for children). This finding differed from other studies which found the IPAQ and PA diaries to also yield reliable results (Craig et al., 2003; Williams et al., 1989). The true amount of PA remains unknown, as is the case with all estimates, but the reliability of accelerometry can be seen as a benchmark indicating that both weeks are comparable. This, however, thrives the question of why the self-reported measures showed limited reliability in our sample. One reason could be that the perception of PA load changed for participants between T_0 and T_1, e.g. because they were bored of the repeated questions or they reflected more about their PA the second time. The reason why children's PA diaries show some reliability might be that they showed higher VPA in all comparable measures. This might be indicative for the circumstance that the children in our sample engaged to a high amount in structurally organized VPA (e.g. training in a sports club, school sport lessons) which is easier to document in a diary than short and intermittent bouts of occasional VPA during everyday situations (e.g. playground), which was previously reported to represent the nature of VPA in children (Edwardson & Gorely, 2010). Finally, the actual PA might also have changed, even though the device-measured results were reliable, which can, however, not be evaluated in the present study because no gold standard of PA measurements has been assessed.

Validity

Unsurprisingly, PA evaluated by 10 and 60-second epoch showed a high and consistent validity among each other for both MPA and VPA at T_0 and T_1 for adults and children despite their indicated total amount of PA differed descriptively. Total values were consistently higher for 10-second intervals than for 60-second intervals, depended on the population and PA intensity (differences: adult MPA: 42-47%, VPA: 46-82%; children: MPA: 14-18%, VPA: 50-56%), even though adults are thought to have longer intervals of PA which should be stable for different epochs (Orme et al., 2014). To illustrate this issue with an example: If a person is moving up one level of stairs rather fast in 20 seconds and stops at the top to have a conversation with a colleague, the use of 10-second epoch length would detect 10 to 20 seconds of VPA while the use of 60-second epochs would calculate the mean over this longer time period and end up with light or MPA (the total metabolic equivalent (MET) would not differ between the epochs, but classification would). Therefore, the high changes in MPA for adults, in this case, may have arisen from a switch of MPA to light PA in the longer epochs because of multiple occasions where MPA lasted less than one minute (e.g. walking short distances in the office).

This supports the importance to consider the impact of epoch lengths on PA outcomes as mentioned in previous studies (Edwardson & Gorely, 2010; Orme et al., 2014). Here, in line with the credo of "every move counts" (Bull et al., 2020) it is recommended to choose shorter epoch lengths as these might capture short bouts of MPA and VPA more sensitively than longer epoch lengths. Concerning the validity of the self-reported measures, the IPAQ was associated with 10-second epochs, and the PA diary at T_0 for adults´ VPA and no further comparison of measurement tools showed a significant result, not supporting weak correlations found in previous studies (Craig et al., 2003; Hagstromer et al., 2010). Descriptive results from the PA diary indicate the lowest reported MPA and VPA (except for children VPA), while the IPAQ showed the highest PA results, which was also found by Hukkanen and colleagues in adults (Hukkanen et al., 2018). One reason for the difference between the self-report measures in the current study could be that participants were instructed to classify their PA in the diary as light, moderate or vigorous while only MPA and VPA were included in this study. Since the IPAQ, has no measure for light PA, participants might have classified their light PA in the diary as MPA in the IPAQ. This might be responsible for the high MPA values reported by the IPAQ but this does not explain the high discrepancy in results for VPA in adults where the PA diary also showed the lowest values. This implies the importance that all measurement methods in a study include the same outcome variables, especially if the measurement methods are compared to each other. A further complication with the PA diary was that there has been no indication included if the diaries have been filled out daily or at the end of the week and that there was no distinction between missing values and no PA. Furthermore, indicating if the accelerometer was worn during the PA which was documented in the diary would have allowed a more detailed impression of discrepancies and the true value of PA during the week. This will be accounted for in the SF2.0 study (Wunsch et al., 2020).

General discussion

The results of the current study are mainly in accordance with the current literature indicating higher reliability than validity for the three measurement tools (Dyrstad et al., 2014; Fiona C. Bull et al., 2009; Hagstromer et al., 2010; Herrmann et al., 2013; Hukkanen et al., 2018; Skender et al., 2016; Slootmaker et al., 2009; Williams et al., 1989) and revealing that the epoch length influences PA estimations (Edwardson & Gorely, 2010; Orme et al., 2014). In contrast to earlier results, we found limited reliability for the PA diary and the questionnaire in adults' VPA. The inclusion of stability shows more stable results in children than in adults, especially for the diary, and adults' VPA is only stable if the two epoch lengths are compared. This has important implications for the use of these measurement tools. Based on current results, future research should further explore the stability between different measurement tools over time to gain further knowledge about the relationship, trying to find a solution to compare single measurement methods to the mixed-method approach in the WHO guidelines. Moreover, different assessment methods should be used which can complement each other like ecological momentary assessment and accelerometry. Researchers should be aware of the limitations of and within each measurement tool and ensure that it is the best fit for the purpose in question. Hence, differences between adults and children in PA research should be considered to deepen the understanding of these differences. Future studies should also aim to create comparable data sets with clear and thorough reporting of outcomes to enable the merging of data in order to be able to compare more subgroups and different settings. Here, it might be helpful to provide a relative amount of PA compared to the wear time or 24-hour measurements in order to compare results between participants in greater detail in future studies. In order to confirm and refine our findings, a replication study with data from the SF2.0 trial with a feasible questionnaire for children and the GPAQ for adults will be conducted in the future (Wunsch et al., 2020).

Strengths and Limitations

The main strength of the current study is the concomitant use of all measurement methods (i.e. accelerometry, questionnaire, and diary) within the same time frame and that they were repeated in the same manner without any intervention in between. Furthermore, both measurement weeks represented an everyday week (i.e. no measurements during holidays), which enhanced comparability. Including reliability and validity as secondary aims in this study helped to interpret and understand the stability results more accurately. This is especially important as results showed reliability and validity to differ from stability results in some cases. Furthermore, the inclusion of data from both adults and children allowed us to analyze differences between these populations. Here, further distinctions between children and adolescents will be interesting to examine in a larger sample. Finally, the use and detailed reporting of multiple measurement tools strengthened the explanatory power of results and allowed for comparison with existing research.

However, there are some limitations to mention. First of all, the true value of PA is unknown and each measure is just an estimate of PA as, for example, no 24-hour measurements of energy expenditure via indirect calorimetry or throughout observation of activity patterns has been recorded (Burchartz et al., 2020). Evaluating the relationship between the different parameters is even more complicated as most questionnaires and PA diaries only ask to report PA with a duration of at least 10 minutes (or even asking for PA over or under 60 minutes (Prochaska, Judith, J. et al., 2001)). With the new WHO guidelines for PA (Bull et al., 2020), self-reports have to be adapted to indicate guideline adherence in larger samples and to be comparable to accelerometry data (Troiano et al., 2020). This might be achieved by removing the wording of reporting only 10 minutes of PA which, however, would increase participant burden and limit adherence due to more detailed reporting requirements and the possible benefit will have to be evaluated in future trials (Troiano et al., 2020).

Another limitation is the rather small sample size which is further divided into adults and children. This limits the generalizability of results and the exploration of subgroups e.g. divided by gender or evaluating results for children or groups of different PA levels separately. Comparisons were also limited as there was no feasible questionnaire included for children in this study. Furthermore, the four-day measurement criteria including eight hours a day might have impacted the measured PA values even though it is assumed to be a sufficient measurement duration (Jacobi et al., 2009), increased the convenience for participants, and allows for reduced loss of data while maintaining reliable data (Colley et al., 2010; Toftager et al., 2013). Furthermore, sedentary behavior was not included in this study, even though the updated WHO recommendations include these important measures (Bull et al., 2020). However, hip-worn accelerometers only capture inactive behavior, but not sedentary patterns (e.g. sitting, lying (Giurgiu et al., 2020)) as has been discussed elsewhere (Kuster et al., 2020). To gain a fair impression of these parameters and to cover all 24 hours of the day, future studies should include the comparison of the outcomes for sedentary behavior and light PA under consideration of non-wear time within a 24-hour measurement approach to evaluate shifts between physical activity levels (e.g. if a higher amount of VPA occurs due to less non-wear time or less SB) (Rowlands et al., 2019).

Finally, because data differed significantly from a normal distribution and especially VPA was skewed due to many low values, no intraclass correlation coefficients could have been calculated, which would have been more accurate as they comprise the total mean value of the measure in the equation (Koo & Li, 2016; Liu et al., 2016). Due to the large number of comparisons, the use of Bland Altman plots as an alternative method for such comparisons (Giavarina, 2015) would be fairly interpretable and was therefore not feasible in our study which used an explorative approach. Future studies should consider a more specific approach with fewer comparisons by formulating clear hypotheses for the present results (e.g. stability for MPA in adults) and use Bland Altman plots to analyze the data in greater detail.

Conclusions

Based on the results of the current study, a comparison between PA estimations (especially for VPA) measured by different tools should be carried out with caution and only if all measurement methods include the same outcome parameters over the same period of time. Here, it needs to be stressed that everyone working with PA values (e.g. scientists planning and conducting PA studies, practitioners giving detailed health-related PA advice, and consumers trying to estimate if they are sufficiently active compared to the guidelines) should carefully consider the measurement tool to be suitable for the purpose in question because considerable discrepancies in results can be detected. Furthermore, it is crucial to use standardized reporting to enhance the comparability of the data (e.g. for future meta-analyses) (Fiedler et al., 2020). Finally, self-reported measures can offer additional contextual information of PA in a timely manner by using e.g. ecological momentary assessments (Reichert et al., 2020; Sattler et al., 2021) to further refine our understanding of PA and may lay the foundation for personalized intervention approaches like just-in-time adaptive interventions (Hardeman et al., 2019) in the future.

References

Allan, M., Poggiali, D., Whitaker, K., & Marshall, Tom, Rhys, Kievit, Rogier, A. (2019). Raincloud plots: a multi-platform tool for robust datavisualization: [version 1; peer review: 2 approved]. *Wellcome Open Research*, *4*(63). https://doi.org/10.12688/wellcomeopenres.15191.1

Anastasopoulou, P., Härtel, S., & Hey, S. (2013). A Comparison of Two Commercial Activity Monitors for Measuring Step Counts During Different Everyday Life Walking Activities. *International Journal of Sports Science and Engineering*, *7*(1), 31–35.

Anastasopoulou, P., Tubic, M., Schmidt, S., Neumann, R., Woll, A., & Härtel, S. (2014). Validation and comparison of two methods to assess human energy expenditure during free-living activities. *PloS One*, *9*(2), e90606. https://doi.org/10.1371/journal.pone.0090606

Armstrong, T., & Bull, F. C. (2006). Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ). *Journal of Public Health*, *14*(2), 66–70. https://doi.org/10.1007/s10389-006-0024-x

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

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C., Chaput, J.-P., Chastin, S., Chou, R., Dempsey, P. C., DiPietro, L., Ekelund, U., Firth, J., Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., . . . Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British Journal of Sports Medicine*, *54*(24), 1451–1462. https://doi.org/10.1136/bjsports-2020-102955

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

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

Craig, C. L, Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J. F., & Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine and Science in Sports and Exercise*, *35*(8), 1381–1395. https://doi.org/10.1249/01.MSS.0000078924.61453.FB

Cust, Anne, E., Smith, Ben, J., Chau, J. Y., van der Ploeg, H. P., Friedenreich, C. M., Armstrong, Bruce, K., & Bauman, A. E. (2008). Validity and repeatability of the EPIC physical activity questionnaire: a validation study using accelerometers as an objective measure. *International Journal of Behavioral Nutrition and Physical Activity*, *5*(1), 33. https://doi.org/10.1186/1479-5868-5-33

Ding, D., Lawson, K. D., Kolbe-Alexander, T. L., Finkelstein, E. A., Katzmarzyk, P. T., van Mechelen, W., & Pratt, M. (2016). The economic burden of physical inactivity: a global analysis of major non-communicable diseases. *The Lancet*, *388*(10051), 1311–1324. https://doi.org/10.1016/S0140-6736(16)30383-X

Downloadable questionnaires - International Physical Activity Questionnaire. (2021, March 4). https://sites.google.com/site/theipaq/questionnaire_links

Dunton, G. F (2017). Ecological Momentary Assessment in Physical Activity Research. *Exercise and Sport Sciences Reviews*, *45*(1), 48–54. https://doi.org/10.1249/JES.0000000000000092

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

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

Feehan, L. M., Geldman, J., Sayre, E. C., Park, C., Ezzat, A. M., Yoo, J. Y., Hamilton, C. B., & Li, L. C. (2018). Accuracy of Fitbit Devices: Systematic Review and Narrative Syntheses of Quantitative Data. *JMIR MHealth and UHealth*, *6*(8), e10527. https://doi.org/10.2196/10527

Fiedler, J., Eckert, T., Wunsch, K., & Woll, A. (2020). Key facets to build up eHealth and mHealth interventions to enhance physical activity, sedentary behavior and nutrition in healthy subjects - an umbrella review. *BMC Public Health*, *20*(1), 1605. https://doi.org/10.1186/s12889-020-09700-7

Fiona C. Bull, Tahlia S. Maslin, & and Timothy Armstrong (2009). Global Physical Activity Questionnaire (GPAQ): Nine Country Reliability and Validity Study. *Journal of Physical Activity and Health*, *6*, 790–804.

Giavarina, D. (2015). Understanding Bland Altman analysis. *Biochemia Medica*, *25*(2), 141–151. https://doi.org/10.11613/BM.2015.015

Giurgiu, M., Bussmann, J. B., Hill, H., Anedda, B., Kronenwett, M., Koch, E. D., Ebner-Priemer, U. W., & Reichert, M. (2020). Validating Accelerometers for the Assessment of Body Position and Sedentary

Behavior. *Journal for the Measurement of Physical Behaviour*, *3*(3), 253–263. https://doi.org/10.1123/jmpb.2019-0068

Hagstromer, M., Ainsworth, B. E., Oja, P., & Sjostrom, M. (2010). Comparison of a Subjective and an Objective Measure of Physical Activity in a Population Sample. *Journal of Physical Activity and Health*, *7*(4), 541–550. https://doi.org/10.1123/jpah.7.4.541

Hardeman, W., Houghton, J., Lane, K., Jones, A., & Naughton, F. (2019). A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. *The International Journal of Behavioral Nutrition and Physical Activity*, *16*(1), 31. https://doi.org/10.1186/s12966-019-0792-7

Härtel, S., Gnam, J.-P., Löffler, S., & Bös, K. (2011). Estimation of energy expenditure using accelerometers and activity-based energy models—validation of a new device. *European Review of Aging and Physical Activity*, *8*(2), 109–114. https://doi.org/10.1007/s11556-010-0074-5

Herrmann, S. D., Heumann, K. J., Der Ananian, C. A., & Ainsworth, B. E. (2013). Validity and Reliability of the Global Physical Activity Questionnaire (GPAQ). *Measurement in Physical Education and Exercise Science*, *17*(3), 221–235. https://doi.org/10.1080/1091367X.2013.805139

Hukkanen, H., Husu, P., Sievänen, H., Tokola, K., Vähä-Ypyä, H., Valkeinen, H., Mäki-Opas, T., Suni, J. H., & Vasankari, T. (2018). Aerobic physical activity assessed with accelerometer, diary, questionnaire, and interview in a Finnish population sample. *Scandinavian Journal of Medicine & Science in Sports*, *28*(10), 2196–2206. https://doi.org/10.1111/sms.13244

IPAQ Research Committee. *Guidelines for data processing and analysis of the International Physical Activity Questionnaire (IPAQ)-short and long forms.* https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnx0aGVpcGFxfGd4OjE0NDgxMDk3NDU1YWRlZTM

Jacobi, D., Charles, M.-A., Tafflet, M., Lommez, A., Borys, J.-M., & Oppert, J.-M. (2009). Relationships of self-reported physical activity domains with accelerometry recordings in French adults. *European Journal of Epidemiology*, *24*(4), 171–179. https://doi.org/10.1007/s10654-009-9329-8

Kohl, H. W., Craig, C. L, Lambert, E. V., Inoue, S., Alkandari, J. R., Leetongin, G., & Kahlmeier, S. (2012). The pandemic of physical inactivity: global action for public health. *The Lancet*, *380*(9838), 294–305. https://doi.org/10.1016/S0140-6736(12)60898-8

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kuster, R. P., Grooten, W. J. A., Blom, V., Baumgartner, D., Hagströmer, M., & Ekblom, Ö. (2020). Is Sitting Always Inactive and Standing Always Active? A Simultaneous Free-Living activPal and ActiGraph Analysis. *International Journal of Environmental Research and Public Health*, *17*(23). https://doi.org/10.3390/ijerph17238864

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

Liu, J., Tang, W., Chen, G., Lu, Y., Feng, C., & Tu, X. M. (2016). Correlation and agreement: Overview and clarification of competing concepts and measures. *Shanghai Archives of Psychiatry*, *28*(2), 115–120. https://doi.org/10.11919/j.issn.1002-0829.216045

Livingstone, M. B. E., Robson, P. J., Wallace, J. M. W., & McKinley, M. C. (2003). How active are we? Levels of routine physical activity in children and adults. *The Proceedings of the Nutrition Society*, *62*(3), 681–701. https://doi.org/10.1079/PNS2003291

Lubans, David, R., Sylva, K., & Osborn, Z. (2008). Convergent Validity and Test–Retest Reliability of the Oxford Physical Activity Questionnaire for Secondary School Students. *Behaviour Change*, *25*(1), 23–34.

Mäder, U., Martin, B. W., Schutz, Y., & Marti, B. (2006). Validity of four short physical activity questionnaires in middle-aged persons. *Medicine and Science in Sports and Exercise*, *38*(7), 1255–1266. https://doi.org/10.1249/01.mss.0000227310.18902.28

Maxime Hervé. (2020). *RVAideMemoire: Testing and Plotting Procedures for Biostatistics*. https://CRAN.R-project.org/package=RVAideMemoire

Migueles, J. H., Cadenas-Sanchez, C., Tudor-Locke, C., 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

Orme, M., Wijndaele, K., Sharp, Steven, J., Westgate, K., Ekelund, U., & Brage, S. (2014). Combined influence of epoch length, cut-point and bout duration on accelerometry-derived physical activity. *International Journal of Behavioral Nutrition and Physical Activity*, *11*(1), 34. https://doi.org/10.1186/1479-5868-11-34

Patterson, P. (2000). Reliability, Validity, and Methodological Response to the Assessment of Physical Activity via Self-Report. *Research Quarterly for Exercise and Sport*, *71 Suppl 2*, 15–20. https://doi.org/10.1080/02701367.2000.11082781

Pols, Margreet, A., Peeters, Peter, H., M., Ocké, Marga, C., Bueno-de-Mesquita, H., Bas, Slimani, N., Kemper, Han, C., G., & Collette, Hubertine, J., A. (1997). Relative Validity and Repeatability of a New Questionnaire on Physical Activity. *Preventive Medicine*, *26*(1), 37–43.

Pols, Margreet, A., Peeters, Peter, H., M., Ocké, Marga, C., Slimani, N., Bueno-de-Mesquita, H., Bas, & Collette, Hubertine, J., A. (1997). Estimation of Reproducibility and Relative Validity of the Questions Included in the EPIC Physical Activity Questionnaire. *International Journal of Epidemiology*, *26*(1), 181-189.

Prochaska, Judith, 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), 554–559. https://doi.org/10.1001/archpedi.155.5.554

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

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., Brüβler, 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

Rivière, F., Widad, F. Z., Speyer, E., Erpelding, M.-L., Escalon, H., & Vuillemin, A. (2018). Reliability and validity of the French version of the global physical activity questionnaire. *Journal of Sport and Health Science*, *7*(3), 339–345. https://doi.org/10.1016/j.jshs.2016.08.004

Rowlands, A. V., Dawkins, N. P., Maylor, B., Edwardson, C. L., Fairclough, S. J., Davies, M. J., Harrington, D. M., Khunti, K., & Yates, T. (2019). Enhancing the value of accelerometer-assessed physical activity: Meaningful visual comparisons of data-driven translational accelerometer metrics. *Sports Medicine - Open*, *5*(1), 47. https://doi.org/10.1186/s40798-019-0225-9

Sattler, M. C., Ainsworth, B. E., Andersen, L. B., Foster, C., Hagströmer, M., Jaunig, J., Kelly, P., Kohl Iii, H. W., Matthews, C. E., Oja, P., Prince, S. A., & van Poppel, M. N. M. (2021). Physical activity self-reports: Past or future? *British Journal of Sports Medicine.* Advance online publication. https://doi.org/10.1136/bjsports-2020-103595

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia and Analgesia*, *126*(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Sirad, John, R., Forsyth, A., Oakes, Michael, J., & Schmitz, Kathryn, H. (2011). Accelerometer Test-Retest Reliability by Data Processing Algorithms: Results From the Twin Cities Walking Study. *Journal of Physical Activity and Health*, *8*(5), 668–674. https://doi.org/10.1123/jpah.8.5.668

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

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

Stamatakis, E., Huang, B.-H., Maher, C., Thøgersen-Ntoumani, C., Stathi, A., Dempsey, P. C., Johnson, N., Holtermann, A., Chau, J. Y., Sherrington, C., Daley, A. J., Hamer, M., Murphy, M. H., Tudor-Locke, C., & Gibala, M. J. (2020). Untapping the Health Enhancing Potential of Vigorous Intermittent Lifestyle Physical Activity (VILPA): Rationale, Scoping Review, and a 4-Pillar Research Framework. *Sports Medicine (Auckland, N.Z.).* Advance online publication. https://doi.org/10.1007/s40279-020-01368-8

Toftager, M., Kirstensen, 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*(1), 140. https://doi.org/10.1186/1479-5868-10-140

Trinh, Oanh, T., H., Nguyen, Nguyen, Do, van der Ploeg, H. P., Dibley, Michael, J., & Bauman, A. E. (2009). Test-Retest Repeatability and Relative Validity of the Global Physical Activity Questionnaire in a Developing Country Context. *Journal of Physical Activity and Health*, *6*(s1), 46–53. https://doi.org/10.1123/jpah.6.s1.s46

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 van Hees, V. (2021, January 5). How to measure PA guideline adherence? *Accelting*. https://accelting.com/updates/how-to-measure-pa-guideline-adherence/

Wickam, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* [Computer software]. Springer-Verlag. New York. https://ggplot2.tidyverse.org

Williams, E., Klesges, Robert, C., Hanson, Cindy, L., & Eck, Linda, H. (1989). A prospective study of the reliability and convergent validity of three physical activity measures in a field research trial. *Journal of Clinical Epidemiology*, *42*(12), 1161–1170.

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

Wunsch, K., Eckert, T., Fiedler, J., Cleven, L., Niermann, C., Reiterer, H., Renner, B., & Woll, A. (2020). Effects of a Collective Family-Based Mobile Health Intervention Called "SMARTFAMILY" on Promoting Physical Activity and Healthy Eating: Protocol for a Randomized Controlled Trial. *JMIR Research Protocols*, *9*(11), e20534. https://doi.org/10.2196/20534

Chapter 4 Considerations for just-in-time adaptive interventions

Paper 3: Just-in-time adaptive interventions in mobile physical activity interventions.

Slightly modified version of the published manuscript

Wunsch, K.*, Eckert, T.*, Fiedler, J.*, & Woll, A. (2022). Just-in-time adaptive interventions in mobile physical activity interventions - A synthesis of frameworks and future directions. *The European Health Psychologist*, *22*(4), 834–842. https://doi.org/10.5445/IR/1000143174

*Authors contributed equally to this work

Abstract

Mobile health (mHealth) solutions seem to be a promising approach to tackle sedentary lifestyle in modern society. They have the potential to identify situations when people are likely to engage in an unhealthy behaviour or when they face opportunities to perform healthy behaviours. These situations can serve as triggers to manipulate current behaviour, defined as just-in-time adaptive interventions (JITAIs) by using real-time behavioural data. The current position paper aims to provide a "think piece" by synthesizing evidence into a short conceptual overview of JITAI research by creating a framework and discussing future directions of JITAI research with a focus on PA interventions.

In conclusion, JITAIs are a promising feature in mHealth applications, however showing a lack of theoretical underpinning until today. To summarize evidence on JITAI implementation research and to provide some guidance, the following key features were identified: a JITAI should 1) correspond to real-time needs; 2) adapt to input data; 3) be system-triggered; 4) be goal-oriented; and 5) be customized to user preferences. These features aim to provide first insights into how to guide researchers and practitioners when developing and reporting JITAI features implemented in mHealth interventions. Concluding from the existing knowledge, the potential of machine learning and deep learning principles for JITAIs regarding mHealth should be further explored and established.

Introduction

Physical activity (PA) plays an important role in the prevention of noncommunicable diseases like cardiovascular diseases, diabetes and obesity (Penedo & Dahn, 2005). Levels of PA, however, are frequently found to be insufficient in modern society (Blair, 2009; Woll et al., 2011). Here, mobile Health (mHealth) interventions might be a promising approach to change PA behaviour and to reduce sedentary behaviour patterns (SBP) operationalized by minimal PA (i.e. PA of less than 1,5 MET) (Fiedler et al., 2020). Several key aspects have been shown to increase intervention efficacy when included in mHealth app development. One of these key components refers to the provision of behaviour change support in real time that is matched to when users are most capable of or in need of this support (Schembre et al., 2018). Various publications have used different terms to describe interventions that adapt the provision of support to an individual's changing internal and contextual state. Analogous to Hardeman and colleagues (2019) as well as Nahum-Shani and colleagues (2018), the term just-in-time adaptive intervention (JITAI) is used throughout this position paper, referring to the potential to immediately intervene in situations when people are either likely to engage in an unhealthy behaviour or when they face opportunities to perform healthy behaviours and adapt these interventions to tailoring variables (e.g. user preferences or sensor input).

The current position paper aims to a) summarize existing conceptualizations of JITAIs, to b) provide a comprehensive overview of JITAI features and mechanisms and to c) provide future directions concerning the implementation of JITAIs in mHealth research.

Theoretical foundations of JITAIs

In recent years, many widely used theories were adapted to explain within-person behavioural variability in order to support new technology-driven interventions that can adapt over time to a person's real-time behaviour and needs (e.g. the Dynamical System Model of Social Cognitive Theory; Martín et al., 2014). Since feedback as a self-regulating strategy is an important component of successful behaviour change, tailored just-in-time feedback depicts a key facet of JITAIs besides timeliness, goal-orientation, personalization and action-orientation (Schembre et al., 2018). In sum, theories indicate that feedback should be personalized, goal-oriented and that it should be presented when attention could be refocused to enhance the likelihood of goal attainment. Here, N-of-1 methodology can be insightful in order to evaluate individual trajectories and antecedents of behaviour change alongside JITAIs (Kwasnicka & Naughton, 2020; McDonald et al., 2017).

Additionally, studies using ecological momentary assessments (EMAs), which are implemented to assess a desired outcome in a specific situation and the natural setting (Stone & Shiffman, 1994), grew rapidly during the past years (Reichert et al., 2020). The results of these studies can provide the foundation for more sophisticated JITAIs (Dunton, 2017; Spruijt-Metz & Nilsen, 2014) and for the application of advanced methods like machine learning algorithms (Kim et al., 2019; Maher et al., 2021; Rozet et al., 2019). By applying such algorithms, researchers aim to automatically detect meaningful patterns in behavioural data which is not feasible with pre-defined specifications due to the complexity and adaptivity of these patterns (Shalev-Shwartz & Ben-David, 2014).

Integration of JITAIs into mHealth interventions

With the continuously growing field of mHealth and a high variety of different sensors and communication devices, the opportunities for the development and implementation of JITAIs are manifold (Reichert et al., 2020). JITAIs are especially useful for behavioural interventions to enhance PA and reduce SBP since they offer new types of timely and adaptive support in the users' natural environment. Therefore, bias due to retrospective measurement methods can be diminished and data of continuously measurements can be obtained. This is especially important as changing contexts (e.g. environmental factors) are highly associated with intervention effectiveness (Hardeman et al., 2019; C. K. Miller, 2019). Although a recent review points to the potential benefit of JITAIs as a key facet within mHealth intervention development (Fiedler et al., 2020), the current evidence on the effectiveness of JITAIs on PA and SBP is limited (Hardeman et al., 2019; Miller, 2019). Most existing JITAI studies show considerable methodological constraints regarding effectiveness measures, i.e. regarding sample size, study design and reporting of JITAI features. Due to the novelty of this research topic, most studies focus on feasibility rather than on the examination of effectiveness in order to aggregate basic knowledge about JITAIs. As an example for a study investigating effectiveness, the MyBehaviour study is interleaving machine learning mechanisms with multi-modal contextualised JITAI components (Rabbi et al., 2015). Here, automatically adapting PA and dietary behaviour advice was integrated into a smartphone application. In addition, PA energy expenditure was calculated and combined with caloric advice. Moreover, environmental information (location) was included for PA advice (Rabbi et al., 2015). Another example study is the SMART*FAMILY* study which includes a JITAI (e.g. provide prompts) along with several other Behaviour Change Techniques (BCTs, e.g. provide information, goal setting, social support). Here, participants received a behavioural support message (i.e. push notification) if they were not sufficiently active (i.e. 100 steps or 2 minutes above 2 MET) during the past hour in order to reduce SBP and enhance PA (Wunsch et al., 2020). Thoroughly, existing studies point to a high acceptance of JITAIs by participants (Hardeman et al., 2019) and to an improvement of user engagement and adherence (Schembre et al., 2018). This, in turn, led to increased awareness of PA opportunities, increased PA and reduced time spent engaging in SBP (Hardeman et al., 2019) in participants using JITAI interventions as compared to no-JITAI users or no-intervention controls.

Theoretical conceptualization of JITAIs

In this position paper, three recent frameworks of JITAIs are presented and synthesized. Hardeman and colleagues (2019) defined three key features that define JITAIs: 1) the provision of behavioural support that directly corresponds to a need in real-time; 2) the adaptation of content or timing of support according to data collected by the corresponding input system since support was initiated; and 3) the system-triggered support. Nahum-Shani and colleagues (2018) distinguish between proximal outcomes (short term goals which can act as mediators to the distal outcome, e.g. daily step count or daily SBP periods), and distal outcomes (behavioural outcome of choice, e.g. increased PA level or decreased SBP level). These authors defined four key facets of JITAIs: 1) decision points (frequency of opportune moments to change the target behaviour and therefore the time at which an intervention decision is made); 2) intervention options (actions to be performed at a decision point); 3) tailoring variables (as obtained via active or passive assessments of individual information, determining intervention delivery); and 4) decision rules (link between the intervention options and the tailoring variables provide the intervention at each decision point). Based on this conceptual framework, Gonul and colleagues (2019) additionally introduced machine learning strategies to individualize decision rules for intervention implementation (i.e. selecting BCTs) based on goal achievement.

Synthesis of theoretical foundations – A holistic and comprehensive conceptual framework for the implementation of JITAIS

As these above-mentioned conceptualizations (i.e. Gonul et al., 2019; Hardeman et al., 2019; Nahum-Shani et al., 2015) build upon different approaches (content, methodology), these conceptual frameworks are synthesized in the following paragraphs in order to provide a holistic and comprehensive overview of JITAI features and mechanisms.

Based on these frameworks, JITAI features were combined and synthesized, attaining a total of five factors which should be taken into account when constituting JITAIs for mHealth research: JITAIs should 1) correspond to real-time needs; 2) adapt to input data; 3) be system-triggered; 4) be goal-oriented; and 5) be customized to user preferences (see Figure 1). The former three factors are needed in order for an intervention to be defined as a JITAI intervention (Hardeman et al., 2019), whereas number 4) and 5) are additional factors which should be included whenever possible to enhance the likelihood of effectiveness and the quality of future interventions in terms of individual user-tailoring (i.e. personalized prevention / medicine). Subsequently, *Tailoring Variables* (e.g. GPS, sensor input data etc.) and *Decision Points and Rules* were added to the framework.

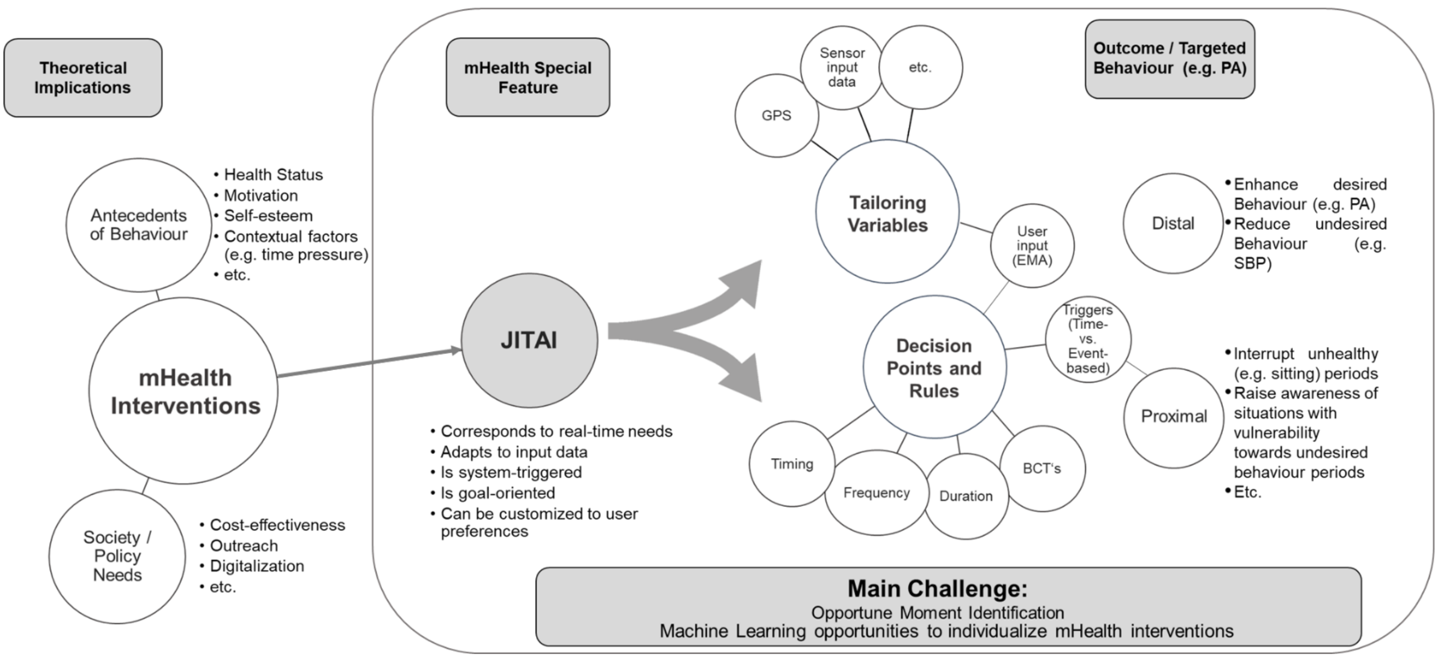

Figure 1. Conceptual framework of JITAIs.

On the left, this figure indicates the *Theoretical Implications* of mHealth for certain *Outcome* variables (on the right). Here, just-in-time adaptive interventions (*JITAIs*) as an *mHealth Special Feature* are described thoroughly concerning their key facets *Tailoring Variables* and *Decision Points and Rules* for *Targeted Behaviour* attainment. *Note.* PA: physical activity; SBP: sedentary behaviour pattern, BCTs: Behaviour Change Techniques, GPS: Global Positioning System; EMA: Ecological Momentary Assessments

Hereafter, italic terms refer to Figure 1. *Theoretical implications* comprising of different *Antecedents of Behaviour* (e.g. mood, sleep, weather, location, opportunity for walking in green areas) and *Society / Policy Needs* determine the content of mHealth interventions. A special feature of such interventions are *JITAI*s, which use different information (i.e. *Tailoring Variables*) to compile a JITAI, e.g. data derived from a sensor, or user input data. Then, *Decision Points* are set in order to determine the points in time when a specific JITAI is triggered. The *Decision Rules* include the designation of principles like *Timing* (e.g. no JITAI at night), *Frequency* (e.g. no JITAI if another JITAI appeared just a couple of minutes ago), *Duration* (e.g. if a JITAI is ignored for a defined amount of time, it won't occur again for a given period of time), and *BCT*-related decision rules (e.g. if the BCT "comparison with others" is completed by the user, a JITAI appears). *User Input* (i.e. no *Trigger* during the next two hours) then lead to the decision if the JITAI is triggered and which *Trigger* will be executed. Beyond these detailed determinations, *Tailoring Variables* and *Decision Points* and *Rules* should finally be defined in order to evaluate whether a *Proximal* goal (e.g. interruption of sitting time) is reached or not and to decide when an additional trigger is necessary and promising in order to reach a more *Distal* goal (i.e. long-term behaviour change).

In the following, an example for a mHealth application using a JITAI for the distal outcome to reduce SBP (which could be based on findings of a recent EMA study (Giurgiu et al., 2020)) by targeting the proximal outcome to interrupt inactive periods will be provided for a more comprehensive understanding of the interconnection of all facets. In a basic version, this JITAI is triggered if a) a connected sensor (e.g. an accelerometer) registers a prolonged period of a SBP (sensor input leading to a Decision Point) and if b) the user is not sleeping (e.g. it is not night-time), didn't receive a JITAI during the past 30 minutes, has not been sufficiently active on that day already (i.e. has already reached his or her step goal), and has no meeting or important appointment based on calendar entries (Decision Rules based on User Input and Tailoring Variables). If all Decision Rules are met at that certain Decision Point, the JITAI trigger will be sent in a moment where the user is likely to engage in an unhealthy behaviour and the intervention is promising for him / her to change this behavior. This basic version could then be adapted according to user preferences and other variables (weather etc.) using machine learning principles.

Taken together, JITAIs aim to positively affect a *Targeted Behaviour*, i.e. PA or SBP based on well-aligned and user-specific adaptability. Setting up *Proximal* targets (i.e. short-term goals which can act as mediators to the *Distal* outcome) can help to achieve a long-term, i.e. *Distal* goal of enhancing PA and / or reducing SBP. Preliminary study results suggest that aiming at short-term goals, receiving feedback, targeting daily life activities as well as the explanation of the reason for reminders and triggers leads to a high acceptance of JITAIs by participants (Hardeman et al., 2019). Hence, implementing these features may improve user engagement and adherence and therefore enhance behaviour change (Schembre et al., 2018). Pilot and feasibility studies also revealed increased awareness of opportunities (e.g. to use active transportation opportunities), a reduction of SBP (e.g. to interrupt screen time periods) and enhanced PA levels, which underlines the potential of JITAIs to change health behaviours (Hardeman et al., 2019).

Opportunities and Challenges of implementing JITAIs in mHealth research

The implementation of JITAIs into mHealth interventions hold promising prospects for health behaviour change. Especially the ongoing development of more advanced and smaller devices to continuously and objectively assess PA and SBP (as well as other health-related variables) and the synthesis of gathered activity-data with additional sensory information (e.g. GPS, ECG, blood-sugar, etc.) further indicate the potential to adapt interventions individually to the user (Reichert et al., 2020).

However, the identification of *Decision Points* and *Rules* (i.e. *Opportune Moment Identification*) for behavioural support depicts the *Main Challenge* of implementing JITAIs (Gonul et al., 2019). Until today, the identification of the optimal number and timing of treatments generated by the JITAI, which are accepted by and effective for users, still remains unknown and most likely depends on the *Proximal* goal and the population of choice. Too frequently sent JITAIs within a specific context, such as the working environment or within school times, may lead to disengagement and/or low adherence and may increase the risk of intervention fatigue. With respect to the implementation of evaluation studies, researchers are advised to use conceptual foundations of JITAI research to determine the critical parameters and choices for participants which are most promising in various settings (e.g. concerning population, duration and aim of the study, and the *Targeted Behaviour*).

Additionally, there is still a need to construct personalized JITAIs comprising the inclusion of behaviour-related (e.g. inactivity) and context-related information (e.g. weather). Here, computational science and machine learning principles offer a new perspective to personalized mHealth interventions (Gonul et al., 2019). Machine learning strategies can include a variety of *Decision Points* into intervention development allowing for context-sensitive and therefore individually tailored and timely flexible support in contrast to fixed algorithms (“if then functions”). Automated system identification modelling can help to identify person-specific *Decision Points and Rules* referring to intrapersonal states and environmental conditions (Conroy et al., 2020). This allows for individually tailored feedback increasing the likelihood of high adherence, user acceptance and higher levels of PA compared to fixed conventional behavioural support. However, a precise forecast of individual behaviour based on system identification modelling requires an extensive data collection prior to intervention onset to gather training data sets derived from different sources and populations. This may impact cost-effectiveness and feasibility of study implementation within a given timeframe for researchers. Some technological aspects also need to be considered when implementing JITAIs into mHealth research, including a short durability of electronic devices due to battery requiring demands (e.g. geolocation features). Furthermore, the necessity of continuous wireless connection between sensors and mHealth devices have to be kept in mind for the development of JITAIs and mHealth interventions in general (Hardeman et al., 2019), as they potentially mitigate user satisfaction and are a source of missing data. Additionally, feasibility studies are warranted in target groups including persons without experience in using digital media, such as older adults. These individuals potentially need additional personal assistance or monitoring to assure safety during PA (K. J. Miller et al., 2014).

Conclusion and Future Directions of JITAI research

The current position paper summarized the knowledge from existing frameworks about JITAIs and synthesized and visualized knowledge into a comprehensive and holistic framework to inform mHealth practitioners about how to implement and report on JITAIs in upcoming mHealth applications. The complexity of designing personalized interventions requires the transdisciplinary collaboration between engineers, computer scientists and behavioural scientists. One of the most important issues is a clear and uniform reporting, which can be informed by the key components of our framework (see Figure 1). Furthermore, reporting should include a clear depiction of the study design (e.g. outcomes, population and duration), methodological approach of the study (e.g. theory used, BCTs and intervention setting) and *Decision Points* and *Rules* (e.g. precise reporting on algorithms or deep learning mechanisms used) in order to compare different studies and to evaluate best-practice approaches for highest effectiveness.

In conclusion, the framework of the current position paper not only provides a basis for the development of JITAIs but also indicates variables which should be reported by JITAI studies.

Future studies should focus on forming consensus on the different parts of the framework to be able to provide a thorough checklist informing researchers and practioners about gold-standards to deploy when initializing JITAI-based mHealth interventions.

References

Blair, S. N. (2009). Physical inactivity: the biggest public health problem of the 21st century. *British Journal of Sports Medicine*, *41*(1), 1–2.

Conroy, D. E., Lagoa, C. M., Hekler, E. B., & Rivera, D. E. (2020). Engineering Person-Specific Behavioral Interventions to Promote Physical Activity. *Exercise and Sport Sciences Reviews*, *48*(4), 170–179. https://doi.org/10.1249/JES.0000000000000232

Dunton, G. F. (2017). Ecological Momentary Assessment in Physical Activity Research. *Exercise and Sport Sciences Reviews*, *45*(1), 48–54. https://doi.org/10.1249/JES.0000000000000092

Fiedler, J., Eckert, T., Wunsch, K., & Woll, A. (2020). Key facets to build up eHealth and mHealth interventions to enhance physical activity, sedentary behavior and nutrition in healthy subjects - an umbrella review. *BMC Public Health*, *20*(1), 1605. https://doi.org/10.1186/s12889-020-09700-7

Giurgiu, M., Niermann, C, Ebner-Priemer, U., & 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

Gonul, S., Namli, T., Huisman, S., Laleci Erturkmen, G. B., Toroslu, I. H., & Cosar, A. (2019). An expandable approach for design and personalization of digital, just-in-time adaptive interventions. *Journal of the American Medical Informatics Association: JAMIA*, *26*(3), 198–210. https://doi.org/10.1093/jamia/ocy160

Hardeman, W., Houghton, J., Lane, K., Jones, A., & Naughton, F. (2019). A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. *The International Journal of Behavioral Nutrition and Physical Activity*, *16*(1), 31. https://doi.org/10.1186/s12966-019-0792-7

Kim, H., Lee, S [SungHee], Lee, S [SangEun], Hong, S., Kang, H., & Kim, N. (2019). Depression Prediction by Using Ecological Momentary Assessment, Actiwatch Data, and Machine Learning: Observational Study on Older Adults Living Alone. *JMIR MHealth and UHealth*, *7*(10), e14149. https://doi.org/10.2196/14149

Kwasnicka, D., & Naughton, F. (2020). N-of-1 methods: A practical guide to exploring trajectories of behaviour change and designing precision behaviour change interventions. *Psychology of Sport and Exercise*, *47*, 101570. https://doi.org/10.1016/j.psychsport.2019.101570

Maher, J. P., Rebar, A. L., & Dunton, G. F. (2021). The influence of context stability on physical activity and sedentary behaviour habit and behaviour: An ecological momentary assessment study. *British Journal of Health Psychology*. Advance online publication. https://doi.org/10.1111/bjhp.12509

Martín, C. A., Rivera, D. E., Riley, W. T., Hekler, E. B., Buman, M. P., Adams, M. A., & King, A. C. (2014). A dynamical systems model of social cognitive theory. *American Control Conference*, 2407–2412.

McDonald, S., Quinn, F., Vieira, R., O'Brien, N., White, M., Johnston, D. W., & Sniehotta, F. F. (2017). The state of the art and future opportunities for using longitudinal n-of-1 methods in health behaviour research: A systematic literature overview. *Health Psychology Review*, *11*(4), 307–323. https://doi.org/10.1080/17437199.2017.1316672

Miller, C. K. (2019). Adaptive Intervention Designs to Promote Behavioral Change in Adults: What Is the Evidence? *Current Diabetes Reports*, *19*(2), 7. https://doi.org/10.1007/s11892-019-1127-4

Miller, K. J., Adair, B. S., Pearce, A. J., Said, C. M., Ozanne, E., & Morris, M. M. (2014). Effectiveness and feasibility of virtual reality and gaming system use at home by older adults for enabling physical activity to improve health-related domains: A systematic review. *Age and Ageing*, *43*(2), 188–195. https://doi.org/10.1093/ageing/aft194

Nahum-Shani, I., Hekler, E. B., & Spruijt-Metz, D. (2015). Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. *Health Psychology*, *34S*, 1209–1219. https://doi.org/10.1037/hea0000306

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Annals of Behavioral Medicine*, *52*(6), 446–462. https://doi.org/10.1007/s12160-016-9830-8

Penedo, F. J., & Dahn, J. R. (2005). Exercise and well-being: A review of mental and physical health benefits associated with physical activity. *Current Opinion in Psychiatry*, *18*(2), 189–193. https://doi.org/10.1097/00001504-200503000-00013

Rabbi, M., Aung, M. H., Zhang, M., & Choudhury, T. (2015). MyBehavior. In K. Mase, M. Langheinrich, D. Gatica-Perez, H. Gellersen, T. Choudhury, & K. Yatani (Eds.), *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15* (pp. 707–718). ACM Press. https://doi.org/10.1145/2750858.2805840

Reichert, M., Giurgiu, M., Koch, E. D., 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*, 101742. https://doi.org/10.1016/j.psychsport.2020.101742

Rozet, A., Kronish, I. M., Schwartz, J. E., & Davidson, K. W. (2019). Using Machine Learning to Derive Just-In-Time and Personalized Predictors of Stress: Observational Study Bridging the Gap Between Nomothetic and Ideographic Approaches. *Journal of Medical Internet Research*, *21*(4), e12910. https://doi.org/10.2196/12910

Schembre, S. M., Liao, Y., Robertson, M. C., Dunton, G. F., Kerr, J., Haffey, M. E., Burnett, T., Basen-Engquist, K., & Hicklen, R. S. (2018). Just-in-Time Feedback in Diet and Physical Activity Interventions: Systematic Review and Practical Design Framework. *Journal of Medical Internet Research*, *20*(3), e106. https://doi.org/10.2196/jmir.8701

Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press.

Spruijt-Metz, D., & Nilsen, W. (2014). Dynamic Models of Behavior for Just-in-Time Adaptive Interventions. *Pervasive Health*, *37*(11), 1443–1451. https://doi.org/10.1038/ijo.2013.120

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine*, *16*(3), 199–202. https://doi.org/10.1093/abm/16.3.199

Woll, A., Kurth, B.-M., Opper, E., Worth, A., & Bös, K. (2011). The 'Motorik-Modul' (MoMo): Physical fitness and physical activity in German children and adolescents. *European Journal of Pediatrics*, *170*(9), 1129–1142. https://doi.org/10.1007/s00431-010-1391-4

Wunsch, K., Eckert, T., Fiedler, J., Cleven, L., Niermann, C, Reiterer, H., Renner, B., & Woll, A. (2020). SMARTFAMILY: A randomized-controlled trial on a collective family-based mobile health intervention to promote physical activity and healthy eating. *JMIR Research Protocols*. http://dx.doi.org/10.2196/20534

Chapter 5 Effectiveness of engaging with a just-in-time adaptive intervention

Paper 4: A just-in-time adaptive intervention to enhance physical activity in the SMART*FAMILY*2.0 trial.

Janis Fiedler*, Caroline Seiferth*, Tobias Eckert, Alexander Woll and Kathrin Wunsch (in print). A just-in-time adaptive intervention to enhance physical activity in the SMART*FAMILY*2.0 trial. *Sport, Exercise, and Performance Psychology*.

*Authors contributed equally to this work

Abstract

Just-in-time adaptive interventions (JITAIs) are a promising technology-based approach for health behavior change. This examination aimed to evaluate whether a JITAI after a period of inactivity can enhance physical activity in the subsequent hour depending on whether the JITAI has been answered ("engaged" condition) compared to when the trigger was not answered ("not engaged" condition). Data of the three-week intervention period of the SMART*FAMILY*2.0 trial was used for analysis. A total of 80 participants (n = 47 adults, 23 female; n = 33 children, 15 female) with 907 JITAI triggers were included in this examination. A JITAI was sent when the participant has been inactive for at least 60 minutes as indicated by accelerometry. Two multilevel models were calculated for metabolic equivalents (MET) and step count with measurements (level 1) nested in participants (level 2) under consideration of several covariates (i.e. weekday/weekend, time of the day, adult/child). Results indicated significantly higher MET (β = 0.08, p = .014) and step (β = 0.08, p = .022) counts in the subsequent hour for the engaged condition compared to the not engaged condition within-persons (level 1). Engagement with the JITAI implemented in the SMART*FAMILY*2.0 trial yielded promising results concerning physical activity enhancement in the subsequent hour. Here, the inclusion of further constraining factors like the availability of the participant or the inclusion of affective and contextual variables into the design of a JITAI might enhance the engagement in future studies.

Keywords: Physical activity, Mobile Health, individual tailoring

Introduction

One important aspect which is linked to metabolic health in children, adolescents, and adults (Healy et al., 2008; Tremblay et al., 2011) is the avoidance of prolonged phases of physical inactivity (e.g. deskwork or watching TV). This is also implemented in the most recent physical activity (PA) guidelines by the world health organization which recommended to reduce sedentary behavior for health benefits (Bull et al., 2020). In the context of health benefits, the reduction of prolonged inactive phases has been shown to be positively associated with physiological health markers like Body Mass Index, waist circumference, and plasma glucose levels in several studies (Carson et al., 2014; Dunstan et al., 2012; Healy et al., 2008). In modern society, however, values of physical inactivity are rising (Bull et al., 2020; Owen et al., 2010) and effective ways to change health behavior throughout the lifespan are needed.

A promising option to deliver cost-effective interventions with a large coverage that aim to break inactive phases are digital interventions (Vandelanotte et al., 2016). Here, mobile health (mHealth) interventions which are described by the World-Health-Organization as “medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices” (World Health Organization, 2011) are especially promising due to increased access to digital devices worldwide (Statista, 2022). One key facet to design effective mHealth interventions in the context of physical inactive phases is the individual tailoring of the interventions to correspond to the participant's behavior (Fiedler et al., 2020; Wunsch et al., 2022). A special case of individual tailoring are just-in-time adaptive interventions (JITAIs) which can be used to interrupt physical inactive phases and enhance PA by providing tailored messages or reminders for healthy behavior in these moments (Hardeman et al., 2019). JITAIs have the potential to automatically intervene when people are most prone to unhealthy behavior or have an opportunity to engage in healthy behavior and adapt these interventions to tailoring variables (e.g. user preferences or sensor input) (for an overview see Nahum-Shani et al., 2018; Wunsch et al., 2022). For a digital intervention to be defined as JITAI the following requirements need to be fulfilled: 1) correspond to real-time needs; 2) adapt to input data; 3) be system-triggered. This can be extended for the enhancement of effectivity by 4) be goal-oriented; and 5) be customized to user preferences (Wunsch et al., 2022). In this regard, decision points (the points in time when a JITAI can be triggered), decision rules (rules which determine if a JITAI is triggered at a decision point), intervention options (possible actions of the JITAI at a decision point), and tailoring variables (sensor- or user-input that is used for adaptation) depict the key features to design JITAIs (Nahum-Shani et al., 2018). Here, the choice of adequate decision points and rules for opportune moments (i.e. times when participants can engage in healthy behavior) present the main challenges in JITAI design. This is especially important, since too many untimely triggers can affect user satisfaction while not enough triggers might not lead to the desired health behavior change or even have adverse effects by promoting engaging with behavior in unintended moments (Gonul et al., 2019). This adaptation of interventions is getting more sophisticated and promising due to technological advances in PA research (e.g. smaller and more powerful accelerometers) (Burchartz et al., 2020) and the amount of individualization (i.e. adapting the intervention to each participants' needs) has been shown to be linked to intervention effectiveness (Baumann et al., 2022).

Previous studies on JITAIs in the context of PA show promising results for a daily accumulated PA (Rabbi et al., 2015) and that sending higher frequented JITAIs per day (after 30/60 minutes of inactivity compared to 120 minutes) was associated with more frequent walking breaks in daily life (Thomas & Bond, 2015). Feasibility studies point to a high user acceptance and preliminary evidence for effectivity of JITAIs while more detailed evaluation are needed (Hardeman et al., 2019). Today, there is a lack of studies evaluating the importance of user engagement with JITAI triggers regarding PA under free-living conditions or during mobile health interventions (Hardeman et al., 2019). Previous studies either did not investigate user engagement, such as timely answering the trigger by clicking on the notification, or focused on e.g. days of smartphone use (Hardeman et al., 2019). Here, evaluating the momentary effect of engagement directly after the trigger occurs is especially relevant as the JITAI is to be triggered in opportune moments for behavior change and a timely response is assumed to be important (Wunsch et al., 2022). The consideration of different aspects of PA (i.e. device-based measured steps as a measure which is used by most people with fitness trackers or smartwatches as a daily goal, and metabolic equivalents (MET) as an indicator for PA intensity) is important, to distinguish between the implications of JITAIs for health behavior change (Silfee et al., 2018).

Therefore, the current examination aimed to evaluate the effectivity of engaging with a JITAI after a prolonged phase of inactivity (>60 minutes) on device-based measured PA (i.e. MET and step counts) in the hour following the trigger during a real-life intervention setting over 21 days in both children and adults.

It was hypothesized that device-based measured step and MET counts in the 60 minutes following the answering of a JITAI trigger ("engaged" condition) were significantly higher within-persons than step and MET counts in the hour where a trigger has been sent but was not answered within 60 minutes ("not engaged" condition).

Additionally, the persistence of the effect for longer time frames (i.e. 90 and 120 minutes) and the between-person effects were exploratively examined. The relevance of the stated reason for the previous period of inactivity has been explored descriptively.

Methods

Transparency and Openness

Data for the current within-person study origins from the SMART*FAMILY*2.0 study. The detailed study protocol of the SMART*FAMILY*2.0 study can be found elsewhere (Wunsch et al., 2020). The International Registered Report Identifier (IRRID) for the SMART*FAMILY* study is RR1-10.2196/20534 and the protocol for the current examination has been pre-registered and uploaded to the open science framework along with the data and analysis code (https://osf.io/u9ca2/). Full ethical approval was obtained from the Karlsruhe Institute of Technology. All participants provided written informed consent. The trial was conducted in accordance with the Declaration of Helsinki.

Participants and Procedure

Within the main study, only families including at least one parent and at least one child who was 10 years of age or older and who were living together in a common household were eligible for the study. Families have been cluster-randomized into an intervention group and a control group. Both groups participated in a baseline measurement of one week, followed by a three-week intervention/waiting period, a one-week post measurement, and a follow-up questionnaire four weeks after the post measurement. During the intervention period, each participant was provided with a smartphone and simultaneously wore a 3-axial accelerometer placed at the hip which corresponded with the smartphone via Bluetooth Low Energy. Participants only had access to the preinstalled SMART*FAMILY*2.0 application (app) on the provided smartphone. The app included several behavior change techniques (e.g. providing information, and goal setting for weekly step and moderate-to-vigorous PA goals) and participants received ecological momentary assessments (i.e. assessing sleep quality and core affect with 4 single item questions) to collect data as part of the study design and an event-based JITAI after a period of physical inactivity longer than 60-minutes. All participants were instructed on app use by a researcher from the SMART*FAMILY*2.0 study, and were provided with a booklet including precise instructions on how to use the app along with troubleshooting. Participants received a 40€ (US $46.8) online shopping voucher and an activity tracker for every child of the family after completing the three assessments of the main study. Participants were not compensated for answering the JITAI and related questions within the app. Power analysis was conducted a priori and resulted in a required total sample size of N = 156 participants to detect a small-to-medium effect for the main trial (Wunsch et al., 2020). Overall, N = 192 participants were included in the SMART*FAMILY*2.0 trial, indicating sufficient power.

For the current examination, only data of the intervention group (N = 98, 52% adults) during the three-week intervention period has been included. Here, the secondary data analysis focuses on the effect of engaging vs not engaging with the JITAI trigger on subsequent (i.e. 60/90/120 minutes following the trigger) PA in a within-person design (see figure 1).

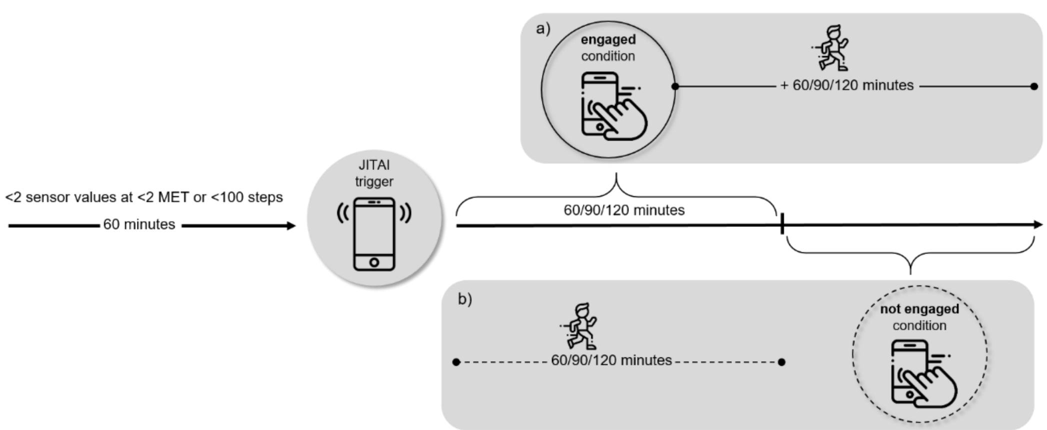

Figure 1. Illustration of 60/90/120-min time windows summarizing physical activity data (step count and METs) when (a) the JITAI trigger was answered within the subsequent 60/90/120 minutes ("engaged" condition) or (b) when the JITAI trigger was not answered within this time window ("not engaged" condition).

Measurements

Questions about age, sex, and anthropometry were included in the questionnaire of the main study at the end of the baseline measurement (Wunsch et al., 2020).

Accelerometry

PA (i.e. step count and MET) was continuously recorded by 3-axial accelerometers (Move 3/Move 4, movisens GmbH, Karlsruhe, Germany). The accelerometers were small-scale (62.3 mm x 38.6 mm x 11.5 mm) and light-weight and were attached by a clip or at a belt to the right hip. Raw data was sampled at an input frequency of 64 Hz and stored on an internal memory card. The accelerometers have been shown to accurately detect step counts (Anastasopoulou et al., 2013) and to validly estimate energy expenditure (Anastasopoulou et al., 2014). Participants were instructed to wear the accelerometer during wake time for the whole intervention period of 21 consecutive days and were told to remove the sensors during showering, swimming, or during contact sports. If participants did not wear the sensor but participated in any exercise, they were instructed to manually record the duration and intensity of the exercise in the SMART*FAMILY*2.0 app (data not included in this examination).

Just-in-time adaptive intervention

We used an event-contingent scheme with JITAI triggers which were sent via the SMART*FAMILY*2.0 app when the participant has been detected to be inactive by the accelerometer for more than 60 minutes (decision rule based on tailoring variables: neither <2 sensor values at >2 MET nor 100 steps registered on the accelerometer; sensor input leading to a decision point). These thresholds were chosen based on previous research pointing out that interrupting inactive phases for at least one minute is associated with health benefits (Carson et al., 2014; Dunstan et al., 2012; Healy et al., 2008) while 60 minutes instead of e.g. 20 minutes were chosen to lower participant burden. The 60-minute time-window following the trigger was chosen to correspond to the minimal frequency of decision points (triggers could only occur every 60 minutes). Triggers regarding inactivity were inhibited for the remaining day if the participant reached at least 60 minutes of moderate-to-vigorous PA on a respective day corresponding to PA guidelines for children (Bull et al., 2020). Furthermore, the trigger only occurred if the participant indicated wakefulness (i.e. has pushed the wake-up button on the app), if there were sufficient (i.e. for at least 50 of 60 minutes) sensor values within the past hour, and there has been no manually adjusted activity for the past 60 minutes within the app (further decision rules). The JITAI was a simple notification stating: ‘You didn't move between 9:00 and 10:14. You should start moving!‘ Along with a single item about the reason for the inactivity with four possible answers (i.e. ‘I engaged in PA but did not wear the sensor‘, ‘I did not have any time‘, ‘I did not feel like doing that‘, ‘I did not feel well‘) and a mood assessment via ecological momentary assessment (not included in this examination). The JITAI notification prevailed until it has been answered and disappeared at midnight if it has not been answered.

Data analysis

Regarding the PA data, raw data has been summarized in 60-second epochs using the software DataAnalyzer, version 1.13.16 (movisens GmbH, Karlsruhe, Germany) and was processed by algorithms into step and MET counts per minute, and non-wear time. MET values were calculated based on activity class (based on acceleration and barometric signals) which determines the estimation model. Then, movement acceleration, altitude change, and demographics were combined in the model for the MET estimation (Härtel et al., 2011). Afterwards, PA and JITAI data have been merged using RStudio (R Core Team, 2021; RStudio Team, 2021). Here, rolling sums for 60, 90, and 120 minutes of the PA data (i.e. summarizing the 60/90/120 values per minute after the trigger) have been calculated. These variables were then matched to the timestamp when the trigger has been sent to the participants (if the JITAI has not been answered within 60/90/120 minutes, this was defined as the "not engaged" condition as it was assumed that participants did not react to the trigger) or to the timestamp when the trigger has been answered by the participant (if the participants engaged with the app by answering clicking on the notification and answering the follow-up questions within 60/90/120 minutes, this was defined as the "engaged" condition). PA data has been considered valid if the sensor has been worn for at least 80% of the respective 60/90/120 minutes. To avoid overlapping periods, the time between the condition ("engaged" or "not engaged") and the condition of the following trigger has been checked and the second trigger has been deleted if the time between the conditions was less than 60/90/120 minutes.

Statistical analysis

Different packages of R (R Core Team, 2021) and RStudio (RStudio Team, 2021) were used for all analyzes. The package ‘ggplot2’ was used for visualizations (Wickham, 2016). Multilevel models were calculated using the package 'lmerTest' (Kuznetsova et al., 2017) with the time of the measurement (level 1) nested in participants (level 2) to identify the within- and between-person effects concerning the research question. The result tables of the regression analyses were generated using the package 'sjPlot' (Lüdecke, 2021). Two final models were calculated for 60, 90, and 120 minutes, one for each PA parameter (sum of step and MET counts per time period) as dependent variables. Intraclass correlation coefficients (ICCs) of the null models indicated that 6% and 8% of variance for the main model (60 minutes) of step and MET counts respectively were due to between-person differences. Therefore, the influence of the hierarchical data structure was confirmed as the majority of the variance was explained by within-person differences. Hence, a multilevel approach was used. In contrast to the preregistration, the inclusion of level 3 (family) was not tested to avoid the overcomplication of the models. Assumptions were checked using the visualization of the 'performance' package (Lüdecke et al., 2021). As visual inspection pointed to no violation of the assumptions, no robust models were calculated. A hierarchical approach was used for the inclusion of the control variables and the model fit was assessed with the Akaike information criterion (AIC).

The dichotomous predictor condition (i.e. "not engaged" = 0, "engaged" = 1) was included at level 1 into the models and centered at the person-mean to estimate within-person effects (Hoffman & Stawski, 2009). Additionally, the control variables weekday or weekend (i.e. wewd, weekday = 0, weekend = 1), and time (i.e. time of the beginning of PA dummy coded as follows: morning (reference) = 00:00:00 to 11:59:59, afternoon = 12:00:00 to 16:59:59, and evening = 17:00:00 till 23:59:59) were included at level 1. The person-mean of the respective condition as well as population (adult = 0, children = 1) were added as a between-person control variables at level 2 into the models. All control variables improved the model fit based on AIC and were therefore included in the final models. In contrast to the preregistration, the reason for the inactivity was not considered as a control variable because, in relation to the trigger, different time-windows were chosen for the "engaged" and "not engaged" condition which allow no direct comparison between conditions (see Figure 1).

Random intercepts were used for all models and the level for significance was set a priori to $\alpha < 0.05$. The equation of the final models was:

Level 1 equation:

$$Y_{ij} = \beta_{0j} + \beta_{1j} * (condition)_{ij} + \beta_{2j} * (wewd)_{ij} + \beta_{3j} * (afternoon)_{ij} + \beta_{4j} * (evening)_{ij} + r_{ij}$$

Level 2 equation:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} * (mean\ condition)_j + \gamma_{02} * (population)_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

Results

Data availability and participant characteristics

Overall, 98 participants were included in the intervention group of the SMART*FAMILY*2.0 study. The average number of days where the app was used for each participant was 17.39 out of 21, equating to 82.85% frequency of daily use (see elsewhere Fiedler et al., 2022). Eighty-four of those participants received and answered at least one JITAI trigger during the 21-day intervention phase (1274 answered JITAI triggers) and were included in this examination. Controlling for sufficient wear time (> 80%) in the 60/90/120 minutes following either the answering of the JITAI ("engaged" condition) or the sending of the JITAI ("not engaged" condition) led to 80/78/77 participants with 907/864/826 observations across both conditions respectively. As the JITAI could be triggered every 60-minutes, data of the 90- and 120-minutes timeframes were controlled for overlapping periods. This led to the final inclusion of 907/810/739 observations (80/80/79% adults) of 80/78/77 (59/60/61% adults) participants respectively. The following sections are referring to the 60-minute timeframe if not specified otherwise. Participant characteristics and PA separated by condition (*n* = 69 "engaged", *n* = 73 "not engaged") are shown in Table 1. Due to the examination design and analysis protocol, individuals can be assigned to both groups. Strictly descriptive, the results showed that an average of 458.19 (*SD* = 813.25) steps were recorded in the hour after the trigger was answered ("engaged") in comparison to 353.55 (*SD* = 562.31) steps when the trigger was not answered ("not engaged").

Table 1 Descriptive data of all participants included in the analyses (N = 80, participants can but do not have to appear in both the "engaged" and the "not engaged" condition data). Displayed are the means(M) and standard deviations (SD) during three weeks for the parameters age, body mass index (BMI) steps, metabolic equivalent (MET) divided by condition ("engaged" or "not engaged"), population (children and adults), and sex (male and female).

condition	"engaged" (397 observations)				"not engaged" (510 observations)			
population	adult		child		adult		child	
sex	female (*n* = 20)	male (*n* = 21)	female (*n* = 17)	male (*n* = 11)	female (*n* = 22)	male (*n* = 23)	female (*n* = 16)	male (*n* = 12)
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Age (years)	44.8 (5.70)	46.5 (4.98)	11.2 (2.81)	13.5 (3.39)	44.4 (5.29)	46.1 (5.02)	11.1 (2.95)	11.3 (4.11)
BMI (kg/m²)	24.0 (4.05)	26.6 (3.09)	17.5 (2.36)	19.3 (2.50)	23.7 (3.81)	26.9 (3.67)	17.2 (2.85)	17.9 (2.65)
steps (counts/hour)	627 (1020)	516 (1180)	391 (433)	952 (1430)	335 (356)	552 (1110)	832 (1400)	480 (714)
MET (counts/hour)	95.5 (34.1)	92.4 (36.9)	86.8 (26.2)	132 (94.4)	82.8 (14.9)	98.6 (55.6)	94.5 (34.9)	98.1 (45.3)

Figure 2 illustrates the average step count separated for both conditions within each person.

Figure 2 Step count in the 60 minutes following the trigger based on the condition (dots for each trigger, X and number for the mean for each condition) for each participant (1-80) for the "not engaged" condition (green) and the "engaged" condition (orange/red) which have been slightly jittered for better visualization. The black lines mark where adult data (n = 47) stops and children data (n = 33) begins.

397 of the 907 observations belong into the "engaged" condition, meaning that the participants received triggers due to inactivity and answered them within 60 minutes. Regarding the reason of inactivity, participants indicated that they did not have the time to be active (289 observations; 75%), did not want to be active (73 observations; 18%), or did not feel good enough to be active (28 observations; 7%). In seven cases, the participants indicated that they had been active but did not wear the sensor. Descriptive results show that the PA tends to be higher after the trigger when participants did not have the time and did not want to be active in the previous minutes (see supplement Figure 1).

Effect of engaging with the just-in-time adaptive intervention on step count

Within-person effects (Level-1)

The results indicate a significantly higher step count in the "engaged" condition compared to the "not engaged" condition within-persons. In detail, if a person was triggered and answered the trigger within 60 minutes ("engaged"), he or she had 113.16 more steps recorded on average ($\beta = 0.08$, $p = .022$) in the hour following the answering compared to the 60 minutes after sending the trigger if the person did not respond within 60 minutes ("not engaged"). No other significant within-person effects were found. The results for 90/120 minutes found a significantly higher step count for "engaged" compared to "not engaged" within-persons. Furthermore, afternoon predicted significantly higher results compared to morning within-persons for 90 and 120 minutes and evening predicted significantly lower results compared to morning for 120 minutes.

Between-person effects (Level-2)

Results showed no significant effect between persons whose data was assigned to the "engaged" condition more often on average compared to persons whose data was assigned to the "not engaged" condition more often. However, significant differences between children and adults ($\beta = 0.08$, $p = .037$) were found which indicate that children had 138.10 more steps recorded in the hour throughout both conditions on average compared to adults. No significant influences of the between-person variables were found for 90 or 120 minutes. Overall, the ICC indicated that 3%/10%/7% of the variance in the models was due to between-person differences and 97%/90%/93% due to within-person variance for 60/90/120 minutes respectively.

Table 2

Multilevel model analysis for the influence of the intervention (i.e. just-in-time adaptive intervention) on step count (steps) in the 60/90/120 minutes following the trigger. Displayed are the within-person results of the person mean centered (mc) condition (control (trigger has not been answered) = 0, JITAI (trigger has been answered) =1), the within-person variable weekend/weekday (wewd, weekday = 0, weekend =1), and time (i.e. dummy coded morning (reference), afternoon, and evening). Additionally, the between-person results of the person mean (pm) condition (control (trigger has not been answered) = 0, JITAI (trigger has been answered) =1), and of population (adult = 0, children = 1) are displayed. All results are displayed using the raw estimates (count per timeframe), the standardized Beta (β), 95% confidence intervals (CI), and standardized (std.) 95% CI. Additionally, the within-person variance (σ^2), the between-person variance ($\tau_{00\ id}$), the intraclass correlation coefficient (ICC), the number of participants (N_{id}), the number of observations, and the marginal and conditional R^2 are displayed.

	steps 60					steps 90					steps 120				
Predictors	*Estimates*	*std. Beta*	*95% CI*	*standardized 95% CI*	*p*	*Estimates*	*std. Beta*	*95% CI*	*standardized 95% CI*	*p*	*Estimates*	*std. Beta*	*95% CI*	*standardized 95% CI*	*p*
(Intercept)	340.56	0.02	197.94 – 483.18	-0.06 – 0.10	<.001	494.43	0.03	246.20 – 742.66	-0.08 – 0.14	<.001	669.65	0.01	358.84 – 980.46	-0.09 – 0.11	<.001
condition_mc	113.16	0.08	16.54 – 209.78	0.01 – 0.14	.022	172.34	0.08	35.77 – 308.90	0.02 – 0.15	.013	181.52	0.07	8.64 – 354.41	0.00 – 0.14	.040
condition_pm	74.24	0.02	-201.84 – 350.32	-0.06 – 0.10	.598	241.98	0.05	-228.43 – 712.40	-0.05 – 0.15	.313	216.87	0.04	-337.54 – 771.28	-0.06 – 0.13	.443
population	138.10	0.08	8.56 – 267.64	0.00 – 0.16	.037	103.92	0.04	-122.05 – 329.90	-0.05 – 0.14	.367	162.77	0.06	-99.06 – 424.60	-0.03 – 0.15	.223
wewd	28.31	0.02	-84.13 – 140.75	-0.05 – 0.08	.622	98.16	0.04	-64.28 – 260.61	-0.03 – 0.11	.236	145.81	0.05	-56.09 – 347.71	-0.02 – 0.12	.157
afternoon	58.32	0.04	-43.33 – 159.96	-0.03 – 0.12	.261	151.47	0.08	7.16 – 295.79	0.00 – 0.15	.040	235.05	0.10	54.77 – 415.32	0.02 – 0.18	.011
evening	-87.02	-0.05	-208.06 – 34.02	-0.13 – 0.02	.159	-162.07	-0.07	-348.33 – 24.20	-0.14 – 0.01	.088	-246.70	-0.08	-492.25 – -1.15	-0.16 – -0.00	.049
Random Effects															
σ^2	450770.06					814737.88					1203517.79				
$\tau_{00\ id}$	13733.47					91955.52					97125.59				
ICC	0.03					0.10					0.07				
N_{id}	80					78					77				
Observations	907					810					739				
Marginal R^2 / Conditional R^2	0.019 / 0.048					0.027 / 0.126					0.035 / 0.107				

Effect of engaging with the just-in-time adaptive intervention on MET count

Within-person effects (Level-1)

Results indicate a significantly higher MET count in the "engaged" condition compared to the "not engaged" condition within-persons. In detail, if a person was triggered and answered the trigger within 60 minutes ("engaged"), he or she had 5.52 more MET counts recorded on average ($\beta = 0.08$, $p = .014$) in the hour following the answering compared to the 60 minutes after sending the trigger if the person did not respond within 60 minutes ("not engaged"). No other significant within-person effects were detected. The results for 90/120 minutes did not reveal any significant differences between "engaged" and "not engaged". Furthermore, evening predicted significantly lower results compared to morning for 90 and 120 minutes.

Between-person effects (Level-2)

Results showed no significant between-person effect for 60/90/120 minutes. The ICC indicated that 6%/14%/7% of the variance in the models was due to between-person differences and 94%/86%/93% due to within-person variance for 60/90/120 minutes respectively.

Table 3

Multilevel model analysis for the influence of the intervention (i.e. just-in-time adaptive intervention) on metabolic equivalents (MET) in the 60/90/120 minutes after the trigger. Displayed are the within-person results of the person mean-centered (mc) variable condition (control (trigger has not been answered) = 0, JITAI (trigger has been answered) =1), the within-person variable weekend/weekday (wewd, weekday = 0, weekend =1), and time (i.e. dummy coded morning (reference), afternoon, and evening). Additionally, the between-person results of the person mean (pm) condition (control (trigger has not been answered) = 0, JITAI (trigger has been answered) =1), and of population (adult = 0, children = 1) are displayed. All results are displayed using the raw estimates (count per timeframe), the standardized Beta (β), 95% confidence intervals (CI), and standardized (std.) 95% CI. Additionally, the within-person variance (σ^2), the between-person variance ($\tau_{00\ id}$), the intraclass correlation coefficient (ICC), the number of participants (N_{id}), the number of observations, and the marginal and conditional R^2 are displayed.

	MET 60					MET 90					MET 120				
Predictors	*Estimates*	*std. Beta*	*95% CI*	*standardized 95% CI*	*p*	*Estimates*	*std. Beta*	*95% CI*	*standardized 95% CI*	*p*	*Estimates*	*std. Beta*	*95% CI*	*standardized 95% CI*	*p*
(Intercept)	86.99	0.03	79.84 – 94.13	-0.06 – 0.12	<.001	131.74	0.04	120.12 – 143.36	-0.08 – 0.16	<.001	173.72	0.01	159.35 – 188.09	-0.09 – 0.11	<.001
condition_mc	5.52	0.08	1.12 – 9.92	0.02 – 0.14	.014	5.26	0.06	-0.65 – 11.18	-0.01 – 0.12	.081	7.36	0.06	-0.79 – 15.52	-0.01 – 0.13	.077
condition_pm	3.36	0.02	-10.84 – 17.57	-0.07 – 0.11	.643	7.31	0.04	-14.88 – 29.50	-0.07 – 0.14	.519	11.10	0.04	-14.45 – 36.65	-0.05 – 0.13	.394
population	4.66	0.06	-1.93 – 11.24	-0.02 – 0.14	.166	1.65	0.02	-9.08 – 12.37	-0.09 – 0.12	.764	4.45	0.03	-7.60 – 16.50	-0.06 – 0.12	.469
wewd	-0.58	-0.01	-5.75 – 4.58	-0.07 – 0.06	.825	1.98	0.02	-5.10 – 9.07	-0.05 – 0.09	.583	2.52	0.02	-6.98 – 12.02	-0.05 – 0.09	.603
afternoon	1.77	0.03	-2.90 – 6.45	-0.05 – 0.10	.457	4.04	0.05	-2.25 – 10.33	-0.03 – 0.12	.208	7.03	0.06	-1.45 – 15.52	-0.01 – 0.14	.104
evening	-4.87	-0.06	-10.45 – 0.71	-0.14 – 0.01	.087	-11.26	-0.10	-19.39 – -3.13	-0.18 – -0.03	.007	-16.62	-0.11	-28.17 – -5.07	-0.19 – -0.03	.005
Random Effects															
σ^2	934.63					1527.87					2677.15				
$\tau_{00\ id}$	58.03					249.11					190.06				
ICC	0.06					0.14					0.07				
N_{id}	80					78					77				
Observations	907					810					739				
Marginal R^2 / Conditional R^2	0.016 / 0.074					0.022 / 0.159					0.030 / 0.094				

Discussion

The present examination showed that engagement with a JITAI triggered by a period of physical inactivity (<100 steps or less than 2 minutes with >2 MET in 60 minutes during waking hours) can enhance device-based measured PA in the subsequent hour. Given that evidence on the momentary effect of engagement with JITAI prompts on free-living PA is yet scarce, an important feature of this examination was the comparison of the PA behavior (steps, MET) in the 60-minute timeframes after the inactivity trigger was answered ("engaged") with the 60-minute timeframes when the inactivity trigger was not responded to ("not engaged"). Overall, results showed that engagement with the basic JITAI implemented in the SMART*FAMILY*2.0 app produced promising results concerning PA enhancement in the subsequent hour after the trigger was answered which needs to be confirmed by future studies.

Results of previous predominantly feasibility studies with small sample sizes indicated the potential of JITAIs to interrupt phases of physical inactivity in individuals with overweight and obesity (Bond et al., 2014; Finkelstein et al., 2015) and individuals with diabetes (Pellegrini et al., 2015). In the above-mentioned studies, the influences of JITAI on accumulated PA outcomes (steps, categories of PA) was investigated either longitudinally on a daily level (Bond et al., 2014), by the comparison of pre- post-intervention (Pellegrini et al., 2015), or in a randomized controlled crossover design (Finkelstein et al., 2015). The current results enhance the understanding of the importance of the engagement with JITAI triggers for subsequent PA behavior directly after the trigger for two different PA measures in a non-clinical sample. Bond et al. (2014) found that a JITAI which was triggered after various periods of inactivity reduced daily values of physical inactivity and enhanced light and moderate PA during the seven day intervention period if compared to a baseline week without a JITAI. Our examination adds that step and MET counts in the 60 minutes directly after the engagement with the trigger are enhanced compared to if the trigger assumedly has not been noticed or has been ignored by the participant. Furthermore, the exploration of 90 and 120 minutes after the trigger indicated a time persistent effect of engagement with the JITAI on step count while the effect vanished for MET count. This points to potential differences for measures related (i.e. MET) and unrelated (i.e. steps) to the intensity of the movement (Silfee et al., 2018).

The time variables referring to daytime (morning, afternoon, and evening) and to the weekday (weekend vs. weekday) had no significant influence on the outcomes for the main model (60 minutes). However, step count was higher in the afternoon for the model using 90 minutes and lower in the evening for 120 minutes while MET count significantly decreased in the evening for both 90- and 120-minute models. This points to dynamic associations and temporal influences of the time in the day to PA measures (e.g., there will be less PA late in the evening but more opportunities for PA directly after work/school in the afternoon) which should be explored in greater detail by future studies to improve the implementation of JITAI-specific features (decision points and rules).

The current examination also provides some exploratory and preliminary indications on the question, if the reason for previous inactivity is associated to the subsequent activity. Here, descriptive results indicate higher variance in step count for the hour following the answering of the trigger if participants stated that they did not have time compared to if they did not want to be active (supplement Figure 1). Furthermore, if participants stated that they were feeling unwell, PA remained low in the following hour. Therefore, future studies should further investigate the reason for inactivity and consider adding it as user-input included in a decision rule at a decision point for the JITAI (e.g. if the participant feels unwell, no trigger will be sent for the rest of the day to reduce participant burden). Additionally, core affect might be an important aspect as it was related to daily PA in a previous examination within the same study (Fiedler et al., 2022). Here, valence and energetic arousal are known to be positively associated with PA while calmness is negatively associated with PA in adults (Forster et al., 2021) and children (Koch et al., 2018). Further contextual factors of inactivity (Giurgiu et al., 2020) like the location, availability and personal preferences of the participant should also be considered to enhance the identification of true opportune moments in future studies.

Strengths and limitations

The primary strength of this examination is that it was conducted in a real-life intervention setting and involved 80 children, adolescents, and adults, two different device-based measured PA outcomes, an extended measurement period of 21 days, and that the design and reporting are guided by a comprehensive JITAI framework (Wunsch et al., 2022). Additionally, the use of multilevel analysis allows for the inclusion of all triggers independently while controlling for the hierarchical structure of the data. This allows for a robust estimation of the effect of engaging with the JITAI.

However, certain limitations need to be considered when interpreting the results of the current investigation. First, included adults and children were already quite active (around 8000 steps and more than 50 minutes of moderate to vigorous PA per day on average while the app was used on 88% of the days see Fiedler et al., 2022). This limited the number of triggers to be analyzed, especially for children. Another aspect was, that the participants had to use provided smartphones instead of their own which can be burdensome and might explain why over 50% of triggers were not answered within 60 minutes. Here, previous research showed that participants who used their own smartphone showed no difference in missed events compared to participants who used an additional smartphone (Ziesemer et al., 2020). One alternative approach for this problem could be to use wearables which can integrate the accelerometer and a small display to respond to JITAIs at the potential cost of the accuracy of the measurement (Feehan et al., 2018). Most importantly, the "not engaged" condition cannot be interpreted as an independent control condition (i.e. no possible influence by the intervention). Furthermore, participants assigned to the "not engaged" condition might have noticed the trigger but simply did not interact with the app and therefore did not create a timestamp. To answer the question regarding the effectiveness of the JITAI (and not the effectiveness of the engagement with the JITAI), micro-randomized trials should be considered in the future to provide a better-controlled comparison and provide insights into causality (Conroy et al., 2020). Finally, it needs to be noted that this examination was a secondary data analysis of a larger study aimed to enhance PA and healthy eating (Wunsch et al., 2020) where the JITAI is only one part of the intervention procedure, which was examined in separation. However, by focusing on the momentary effects, the influences of other interventional aspects (e.g. influence of providing information, and goal setting) are assumed to be limited.

Conclusion

The examination expands previous findings on JITAIs by focusing on the engagement with the JITAI and by considering the temporal associations between the trigger and the outcome in a multilevel approach in children and adults. The results underline the importance of participants' engagement with JITAI triggers to interrupt inactive phases. Here, factors like time of the day and the reason for the inactivity are possibly important influences on PA measures. Future studies should further refine the understanding for opportune moment identification by involving participants in JITAI design and build on existing findings from ecological momentary assessment research (e.g. Giurgiu et al., 2020). These important tailoring variables like the core affective state of the participants and contextual factors like availability and weather should then be used to enhance the adaptation to participants needs and therefore the engagement and effectiveness of JITAIs.

References

Anastasopoulou, P., Härtel, S., & Hey, S. (2013). A Comparison of Two Commercial Activity Monitors for Measuring Step Counts During Different Everyday Life Walking Activities. *International Journal of Sports Science and Engineering*, *7*(1), 31–35.

Anastasopoulou, P., Tubic, M., Schmidt, S., Neumann, R., Woll, A., & Härtel, S. (2014). Validation and comparison of two methods to assess human energy expenditure during free-living activities. *PloS One*, *9*(2), e90606. https://doi.org/10.1371/journal.pone.0090606

Baumann, H., Fiedler, J., Wunsch, K., Wollesen, B., & Woll, A. (2022). mHealth Interventions to Reduce Physical Inactivity and Sedentary Behavior in Children and Adolescents: Systematic Review and Meta-analysis of Randomized Controlled Trials. *JMIR MHealth and UHealth*, *10*(5). https://doi.org/10.2196/35920

Bond, D. S., Thomas, J. G., Raynor, H. A., Moon, J., Sieling, J., Trautvetter, J., Leblond, T., & Wing, R. R. (2014). B-MOBILE—a smartphone-based intervention to reduce sedentary time in overweight/obese individuals: A within-subjects experimental trial. *PloS One*, *9*(6), e100821. https://doi.org/10.1371/journal.pone.0100821

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C., Chaput, J.-P., Chastin, S., Chou, R., Dempsey, P. C., DiPietro, L., Ekelund, U., Firth, J., Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., . . . Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British Journal of Sports Medicine*, *54*(24), 1451–1462. https://doi.org/10.1136/bjsports-2020-102955

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

Carson, V., Wong, S. L., Winkler, E., Healy, G. N., Colley, R. C., & Tremblay, M. S. (2014). Patterns of sedentary time and cardiometabolic risk among Canadian adults. *Preventive Medicine*, *65*, 23–27. https://doi.org/10.1016/j.ypmed.2014.04.005

Conroy, D. E., Lagoa, C. M., Hekler, E., & Rivera, D. E. (2020). Engineering Person-Specific Behavioral Interventions to Promote Physical Activity. *Exercise and Sport Sciences Reviews*, *48*(4), 170–179. https://doi.org/10.1249/JES.0000000000000232

Dunstan, D. W., Kingwell, B. A., Larsen, R., Healy, G. N., Cerin, E., Hamilton, M. T., Shaw, J. E., Bertovic, D. A., Zimmet, P. Z., Salmon, J., & Owen, N. (2012). Breaking up prolonged sitting reduces postprandial glucose and insulin responses. *Diabetes Care*, *35*(5), 976–983. https://doi.org/10.2337/dc11-1931

Feehan, L. M., Geldman, J., Sayre, E. C., Park, C., Ezzat, A. M., Yoo, J. Y., Hamilton, C. B., & Li, L. C. (2018). Accuracy of Fitbit Devices: Systematic Review and Narrative Syntheses of Quantitative Data. *JMIR MHealth and UHealth*, *6*(8), e10527. https://doi.org/10.2196/10527

Fiedler, J., Eckert, T., Wunsch, K., & Woll, A. (2020). Key facets to build up eHealth and mHealth interventions to enhance physical activity, sedentary behavior and nutrition in healthy subjects - an umbrella review. *BMC Public Health*, *20*(1), 1605. https://doi.org/10.1186/s12889-020-09700-7

Fiedler, J., Seiferth, C., Eckert, T., Woll, A., & Wunsch, K. (2022). Sleep quality, valence, energetic arousal, and calmness as predictors of device-based measured physical activity during a three-week mHealth intervention. *German Journal of Exercise and Sport Research.* Advance online publication. https://doi.org/10.1007/s12662-022-00809-y

Finkelstein, J., McKenzie, B., Xuan, L., Jeffrey, W., & Pamela, O. (Eds.) (2015). *Mobile App to Reduce Inactivity in Sedentary Overweight Women*. MedInfo 2015: eHealth-enabled.

Forster, A. K., Richards, E. A., Foli, K. J., McGowan, B., Hass, Z., Becker, M., & Miller, A. (2021). Influence of Affect on Physical Activity: An Integrative Review. *Clinical Nursing Research*, *30*(7), 934–949. https://doi.org/10.1177/1054773820968039

Giurgiu, M., Niermann, C., Ebner-Priemer, U., & 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

Gonul, S., Namli, T., Huisman, S., Laleci Erturkmen, G. B., Toroslu, I. H., & Cosar, A. (2019). An expandable approach for design and personalization of digital, just-in-time adaptive interventions. *Journal of the American Medical Informatics Association: JAMIA*, *26*(3), 198–210. https://doi.org/10.1093/jamia/ocy160

Hardeman, W., Houghton, J., Lane, K., Jones, A., & Naughton, F. (2019). A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. *The International Journal of Behavioral Nutrition and Physical Activity*, *16*(1), 31. https://doi.org/10.1186/s12966-019-0792-7

Härtel, S., Gnam, J.-P., Löffler, S., & Bös, K. (2011). Estimation of energy expenditure using accelerometers and activity-based energy models—validation of a new device. *European Review of Aging and Physical Activity*, *8*(2), 109–114. https://doi.org/10.1007/s11556-010-0074-5

Healy, G. N., Dunstan, D. W., Salmon, J., Cerin, E., Shaw, J. E., Zimmet, P. Z., & Owen, N. (2008). Breaks in sedentary time: Beneficial associations with metabolic risk. *Diabetes Care*, *31*(4), 661–666. https://doi.org/10.2337/dc07-2046

Hoffman, L., & Stawski, R. S. (2009). Persons as Contexts: Evaluating Between-Person and Within-Person Effects in Longitudinal Analysis. *Research in Human Development*, *6*(2-3), 97–120. https://doi.org/10.1080/15427600902911189

Koch, E. D., Tost, H., Braun, U., Gan, G., Giurgiu, M., Reinhard, I., Zipf, A., Meyer-Lindenberg, A., Ebner-Priemer, U., & Reichert, M. (2018). Mood Dimensions Show Distinct Within-Subject Associations With Non-exercise Activity in Adolescents: An Ambulatory Assessment Study. *Frontiers in Psychology*, *9*, 268. https://doi.org/10.3389/fpsyg.2018.00268

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lüdecke, D. (2021). *sjPlot: Data Visualization for Statistics in Social Science* [Computer software]. https://CRAN.R-project.org/package=sjPlot

Lüdecke, D., Mattan, S. B.-S., Indrajeet, P., Philip, W., & Dominique, M. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, *60*(6), 3139. https://doi.org/10.21105/joss.03139

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*, *52*(6), 446–462. https://doi.org/10.1007/s12160-016-9830-8

Owen, N., Healy, G. N., Matthews, C. E., & Dunstan, D. W. (2010). Too much sitting: The population health science of sedentary behavior. *Exercise and Sport Sciences Reviews*, *38*(3), 105–113. https://doi.org/10.1097/JES.0b013e3181e373a2

Pellegrini, C. A., Hoffman, S. A., Daly, E. R., Murillo, M., Iakovlev, G., & Spring, B. (2015). Acceptability of smartphone technology to interrupt sedentary time in adults with diabetes. *Translational Behavioral Medicine*, *5*(3), 307–314. https://doi.org/10.1007/s13142-015-0314-3

R Core Team. (2021). *R: A Language and Environment for Statistical Computing* [Computer software]. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Rabbi, M., Aung, M. H., Zhang, M., & Choudhury, T. (2015). MyBehavior. In K. Mase, M. Langheinrich, D. Gatica-Perez, H. Gellersen, T. Choudhury, & K. Yatani (Eds.), *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15* (pp. 707–718). ACM Press. https://doi.org/10.1145/2750858.2805840

RStudio Team. (2021). *RStudio: Integrated Development Environment for R [*Computer software]. RStudio, PBC. Boston, MA. http://www.rstudio.com/

Silfee, V. J., Haughton, C. F., Jake-Schoffman, D. E., Lopez-Cepero, A., May, C. N., Sreedhara, M., Rosal, M. C., & Lemon, S. C. (2018). Objective measurement of physical activity outcomes in lifestyle interventions among adults: A systematic review. *Preventive Medicine Reports*, *11*, 74–80. https://doi.org/10.1016/j.pmedr.2018.05.003

Statista. (2022, March 7). *Internet users in the world 2021 | Statista*. https://www.statista.com/statistics/617136/digital-population-worldwide/

Thomas, G. J., & Bond, D. S. (2015). Behavioral response to a just-in-time adaptive intervention (JITAI) to reduce sedentary behavior in obese adults: Implications for JITAI optimization. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, *34S*, 1261–1267. https://doi.org/10.1037/hea0000304

Tremblay, Mark S., LeBlanc, A. G., Kho, Michelle E., Saunders, Travis J., Larouche, R., Colley, Rachel, C., Goldfield, G., & Gorber, S. C. (2011). Systematic review of sedentary behaviour and health indicators in school-aged children and youth. *The International Journal of Behavioral Nutrition and Physical Activity*, *8*(98). https://doi.org/10.1186/1479-5868-8-98

Vandelanotte, C., Müller, A. M., Short, C. E., Hingle, M., Nathan, N., Williams, S. L., Lopez, M. L., Parekh, S., & Maher, C. A. (2016). Past, Present, and Future of eHealth and mHealth Research to Improve Physical Activity and Dietary Behaviors. *Journal of Nutrition Education and Behavior*, *48*(3), 219-228.e1. https://doi.org/10.1016/j.jneb.2015.12.006

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* [Computer software]. Springer-Verlag New York. https://ggplot2.tidyverse.org

World Health Organization. (2011). *Mhealth: Second Global Survey on eHealth*. World Health Organization. http://gbv.eblib.com/patron/FullRecord.aspx?p=851142

Wunsch, K., Eckert, T., Fiedler, J., Cleven, L., Niermann, C., Reiterer, H., Renner, B., & Woll, A. (2020). Effects of a Collective Family-Based Mobile Health Intervention Called "SMARTFAMILY" on Promoting Physical Activity and Healthy Eating: Protocol for a Randomized Controlled Trial. *JMIR Research Protocols*, *9*(11), e20534. https://doi.org/10.2196/20534

Wunsch, K., Eckert, T., Fiedler, J., & Woll, A. (2022). Just-in-time adaptive interventions in mobile physical activity interventions - A synthesis of frameworks and future directions. *The European Health Psychologist*, *22*(4), 834–842.

Ziesemer, K., König, L. M., Boushey, C. J., Villinger, K., Wahl, D. R., Butscher, S., Müller, J., Reiterer, H., Schupp, H. T., & Renner, B. (2020). Occurrence of and Reasons for "Missing Events" in Mobile Dietary Assessments: Results From Three Event-Based Ecological Momentary Assessment Studies. *JMIR MHealth and UHealth*, *8*(10), e15430. https://doi.org/10.2196/15430

Chapter 6 Core affect and sleep quality in physical activity interventions

Paper 5: Sleep quality, valence, energetic arousal, and calmness as predictors of device-based measured physical activity during a three-week mHealth intervention.

Slightly modified version of the published manuscript

Fiedler, J., Caroline Seiferth, Eckert, T., Woll, A., & Wunsch, K. Sleep quality, valence, energetic arousal, and calmness as predictors of device-based measured physical activity during a three-week mHealth intervention: An ecological momentary assessment study within the SMART*FAMILY* trial. *German Journal of Exercise and Sport Research (2022).* https://doi.org/10.1007/s12662-022-00809-y

Abstract

Physical inactivity is known to be a risk factor for several non-communicable diseases and has a high prevalence in today's society. Therefore, it is crucial to understand the psychological factors associated with physical activity (PA). Recent developments in the field of ambulatory assessment and technological advances are promising to enhance our understanding of this relationship by analyzing longitudinal data within- and between-persons. These analyses can reveal important factors to design behavior change interventions to enhance PA. Therefore, this study used an ecological momentary assessment during the three-week intervention period in the SMART*FAMILY*2.0 trial and aimed to investigate whether valence, calmness, energetic arousal, and sleep quality predict daily steps and moderate to vigorous PA. Overall, 49 adults (35-60 years) and 40 children (5-19 years) were included in this analysis and self-rated their mental state within our smartphone application while also wearing a hip-worn accelerometer for 21 consecutive days (996 days included) during the intervention period. Multilevel analyses were conducted to predict daily PA while considering covariables (e.g. child/adult and non-wear time) both within- and between-persons. The results indicated that higher than average ratings of a person's valence and energetic arousal on one day predicted increased PA while higher than average calmness predicted decreased PA at the same day within this person. Sleep quality and between-person effects of the affective states showed no clear associations to PA. Overall, these results showed that within-person associations of valence, calmness, and energetic arousal should be considered when designing PA interventions for both children and adults. The influence of sleep quality, as well as between-person effects, should be further explored by future studies.

Keywords

accelerometry, mobile health, ambulatory assessment, affective states

Introduction

Physical inactivity depicts one major risk factor for a variety of non-communicable diseases (Kohl et al., 2012) while sufficient physical activity (PA) represents an effective primary prevention strategy for non-communicable diseases throughout the lifespan (Beaglehole et al., 2011). However, only 32% of the worldwide population reach the PA recommendations of 150 minutes of moderate or 75 minutes of vigorous PA or an equivalent of both for adults (>18 years) and an average of 60 minutes moderate to vigorous PA (MVPA) per day for children (5-17 years) (Bull et al., 2020; Hallal et al., 2012). Hence, effective interventions to reduce physical inactivity and to enhance PA are needed for adults and children to meet their respective guidelines. Today, mobile health (mHealth) interventions are promising tools for health behavior change due to preliminary results for effectiveness, 24/7 availability, large coverage, and their assumed cost-effectiveness (Vandelanotte et al., 2016). Important key facets for effective mHealth interventions are hereby the theoretical foundation, the use of behavior change techniques, interventions‘ embeddedness in a social context, and individual tailoring (Fiedler et al., 2020). Besides these contextual and cognitive factors, there is a further need to investigate affect-related determinants in individuals assigned to a mHealth intervention targeting PA to identify reasons for uptake, or barriers, of subsequent PA (Dunton, 2017).

Ecological Momentary Assessment (EMA) provides an opportunity to not only deliver interventional content but also to gather real-time within- and between-person longitudinal data throughout the intervention period (Trull & Ebner-Priemer, 2013). This allows the detection of dynamic associations between determinants of subsequent PA on an individual level, which can be considered in personalized behavior change interventions (Conroy et al., 2020). Of particular interest are hereby dimensions of affect that are assumed to be linked to an improved health behavior (Trull & Ebner-Priemer, 2013). There is much contradiction and overlap in the conceptualization of affect, mood, and emotion (for a review, see Ekkekakis, 2013). James Russel (2003) proposed a framework that establishes interrelationships between these concepts and defined core affect as a “neurophysiological state consciously accessible as a simplest raw (nonreflective) feeling evident in moods and emotions” (Russell, 2003, p. 148). Building on this, different models and dimensions of core affect have been postulated in recent years. According to the three-dimensional model, core affect includes at least three basic intercorrelated affective dimensions that map the complexity of affective states in daily life: valence (pleasure - displeasure), energetic arousal (wakefulness-tiredness), and calmness (relaxation-tension) (Schimmack & Grob, 2000).

Extensive research has been conducted in the past years investigating the relationship between PA and core affect in adult populations (Forster et al., 2021; Liao et al., 2015). Previous research indicates that valence (Carels et al., 2007; Emerson et al., 2018; Kanning & Schoebi, 2016; Schwerdtfeger et al., 2010) and energetic arousal (Liao et al., 2017; Niermann et al., 2016; Schwerdtfeger et al., 2010) are positively associated with subsequent PA while calmness is negatively associated with PA (Kanning & Schoebi, 2016; Reichert et al., 2016). Although the results seem to be coherent on the affective dimensions, a direct comparison is difficult because the studies analyzed different temporal aspects of subsequent activity (i.e., 24 hours, 15 minutes) and different types of movement (i.e., free-living PA vs. structured exercises) (Forster et al., 2021). For example, Carels et al., (2007) and Emerson et al., (2018) investigated the relationship between affect and PA within a single day and the results indicate that higher ratings of valence in the morning were associated with increased PA over the day. Here, both studies assessed PA by self-report, which may not represent changes within an individual in detail (Reichert et al., 2020) and often differs from device-based measured PA (Fiedler et al., 2021). Comparable results for the relation of PA and energetic arousal alone were also found in children between 9 – 13 years (Dunton et al., 2014), and for all three affective states in children between 12 and 17 years (Koch et al., 2018). Despite these findings, it is important to note that the dynamic relationship between affective states and PA has been studied much less in children and that the existing results are heterogenous (Bourke et al., 2021). Additionally, parameters of sleep (i.e., perceived sleep quality, duration, efficacy) are further important determinants of health-related behavior that are assumed to be linked with PA (Wang & Boros, 2021). However, a recent meta-analysis including adult samples, revealed no direct relationship between sleep on subsequent PA (Atoui et al., 2021) while a longer sleep duration was associated with improved eating behavior and higher levels of PA in children (Khan et al., 2015).

As stated above, the dimensions of core affect and perceived sleep quality can influence PA behavior in both adults and children. Therefore, it is important to investigate these covariates during a theory-based intervention in which key facets of behavior change are implemented. This can help to assess the possible impact of affective states and sleep quality on the main outcome (PA) of the intervention. Here, existing studies have mainly evaluated EMA measured constructs as time-lagged predictors immediately before PA uptake to investigate their momentary effect (Liao et al., 2015). However, in the intervention context day-level peculiarity might also be of interest, as intervention studies usually include time intervals of several days to weeks, and the question if EMA-derived variables have an impact on this time scale is important for designing such interventions. Another important point is to take the PA outcome into account. Here, a study by Reichert et al (2017) found differences in the relationship of PA to affective states for exercise and nonexercise PA which suggests, that there is no uniform relationship between PA and affect. Knowledge of the mechanisms and barriers related to PA uptake during a longer measurement period also will help to anticipate mental health- and sleep quality-related barriers causing physical inactivity which can then be considered for the development of future mHealth interventions (Dunton, 2017).

Hence, the present study aimed to investigate several potential mental health-related covariates of PA including valence, energetic arousal, and calmness as well as perceived sleep quality on a daily level during three weeks to predict same-day PA measured by 1) steps, and 2) MVPA, among children and adults during a PA intervention period. These two PA measures were used to account for possible differences in the relationship between an intensity independent (steps) and intensity related (MVPA) PA measure, and to project two different types of PA guidelines: the step-related guideline of reaching between 7.000 and 10.000 steps per day (e.g. Paluch et al., 2021), which is followed by most people using fitness trackers or smartwatches as a daily goal, and the intensity-related guideline provided by the World Health Organization (Bull et al., 2020).

Following previous findings on the topic, it is hypothesized, that on days where participants report higher than usual valence and energetic arousal, they have greater device-based measured step count and MVPA on the same day while on days where participants report higher than usual calmness, they have lower device-based measured step count and MVPA on the same day (within-persons). Between-person effects of valence, energetic arousal, and calmness on steps and MVPA (e.g. participants who report higher valence on average have higher/lower average device-based measured step count compared to persons who report lower valence on average), and the relationship between sleep quality and PA on a within- and between-person level will be explored.

Methods

Participants and Procedure

Data for the current study was based on the SMART*FAMILY*2.0 trial. For the detailed study protocol of the SMART*FAMILY*2.0 study see elsewhere (Wunsch et al., 2020). Full ethical approval and written informed consent of all participants, children, and legal guardians was obtained (The International Registered Report Identifier (IRRID) for the SMART*FAMILY* study is RR1-10.2196/20534.). The trial was conducted in accordance with the Declaration of Helsinki.

Participants (families) were recruited in schools, school holiday programs, music schools, and sports clubs via personal communication, newspapers, and email distribution lists of the Karlsruhe Institute of Technology. Only families including at least one parent and at least one child who was 10 years of age or older and who were living together in a common household were eligible for the study. Additionally, all siblings were invited to take part in the study if the parent/s vouched for their ability to participate (Wunsch et al., 2020). All participants have been cluster-randomized into an intervention group and a control group. The protocol for both groups included a baseline measurement of one week, followed by a three-week intervention/ waiting period and one week post measurement. The original study aimed to enhance PA and healthy eating with the digital intervention. For this study, only data of the three-week intervention period in the intervention group (n = 98, 52% adults) has been included. During this intervention period, participants used the SMART*FAMILY*2.0 app on provided smartphones and wore an accelerometer. To increase participants‘ health literacy, information about the benefits of PA and healthy eating was provided in the app. Additionally, participants autonomously set activity- and diet-related weekly goals, received feedback on goal achievement, and received a just-in-time adaptive intervention (i.e. a push notification when the participant was inactive during the wake time for at least 60 minutes (neither <2 sensor values at >2 MET nor 100 steps registered on the accelerometer); for an overview of just-in-time adaptive interventions see Wunsch et al., 2022). EMA concerning sleep quality was sent once in the morning (i.e. the first action of the participant on the app each day), and EMA concerning affect after a period of inactivity (following the just-in-time adaptive intervention) and, if no trigger occurred for several hours, in the evening.

Measurements

Accelerometry

PA (i.e. steps and MVPA per day) was continuously recorded by 3-axial accelerometers (Move 3/Move 4, Movisens GmbH, Karlsruhe, Germany). The small-scale (62.3 mm x 38.6 mm x 11.5 mm;) and light-weight accelerometers were worn at the right hip and were attached by a clip or at a belt. Raw data was sampled at an input frequency of 64 Hz and afterward summarized in 60-second epochs. Analyzed raw data were processed by algorithms into steps, time spent during MVPA minutes per day [>3 metabolic equivalents (MET)], inactive time [1-1.5

MET], and non-wear time for this study. The accelerometers have been shown to accurately detect step counts (Anastasopoulou et al., 2013) and to validly estimate energy expenditure (Anastasopoulou et al., 2014).

Participants were instructed to wear the accelerometer during wake time for the whole intervention period of three weeks, with each measurement period starting on a Monday. Participants were told to remove the sensors during showering, swimming, or during contact sports. In this case, the participants were instructed to manually record the duration and intensity of the exercise in the SMART*FAMILY*2.0 app (not included in our study).

Ecological momentary assessment

Several EMAs were assessed within the study. Participants were instructed to use the app throughout the day and only mute it during e.g. meetings or school. With the first action on the app in the morning, every participant rated the perceived sleep quality once a day on a 7-point Likert scale ("How would you rate your sleep quality during the previous night?" 0= very bad, 6 = very good, adapted from Snyder at al. (2018)) once a day. Wilhelm & Schoebi, 2007 previously showed that two bipolar items each provide sensitive and reliable measurements of the three-dimensional model of core affect (Schimmack & Grob, 2000). In this study, only one of those two bipolar items was used asking for affective valence ("How is your current mood?" rated by emojis from 0 = very bad, to 4 = very good), energetic arousal ("Are you feeling awake or tired?" 0 = very tired, 6 = very awake), and calmness ("Are you feeling relaxed/calm or stressed?" 0 = very stressed, 6 = very calm). The items were used based on Bachmann et al. (2015) to keep participant burden low. The use of single items can hereby be beneficial for research focused on a broader perspective of the relationship between affect and PA even though it limits conclusions about discrete affects (Emerson et al., 2018). The EMA concerning affective states was sent following an event-contingent scheme when participants were inactive during the last 60 minutes (neither >2 sensor values at >2 MET nor 100 steps), and when the participant finished their day in the app by pressing the "going to sleep button" (provided no trigger occurred during the past hour). The inactivity triggers were blocked when 1) the app was "asleep", 2) during the night (10 pm till 7 am), 3) less than 50 of 60-minute values have been sent during the past hour by the sensor, and 4) if a participant reached a PA level of 60 minutes of MVPA on a certain day. As EMAs could be sent multiple times a day, daily averages were calculated for valence, energetic arousal, and calmness. Here, it needs to be noted that this study is a secondary data analysis of the intervention period in a free-living study, where participants were not instructed to answer a certain amount of EMA questionnaires. Therefore, the interaction with the app does not represent compliance as in other EMA studies but user engagement with the app (comparable to e.g. Edney et al., 2019).

Statistical Analysis

R (R Core Team, 2021) and RStudio (RStudio Team, 2021) were used for data preparation and analysis. The package 'ggplot2' was used for visualizations (Hadley Wickham, 2016). Due to the hierarchal structure of the data multilevel models were calculated using the package 'nlme' (Jose Pinheiro et al., 2021) with days of the intervention (level 1) nested in participants (level 2) to identify the within- and between-person effects concerning the research question. The result tables of the regression analyses were generated using the package 'sjPlot' (Daniel Lüdecke, 2021). Here, two final models were calculated, one with each PA parameter (steps and MVPA per day) as outcome variables. Intraclass correlation coefficients (ICCs) of the null model indicated that 40% and 54% variances for each of the steps and MVPA respectively were due to within-person differences. Therefore, the influence of the hierarchal data structure on the outcome variables was confirmed and a multilevel approach was used. ICCs for the predictor variables indicated that between 60% and 71% of variance was explained by within-person differences and the variables were therefore disaggregated into within- and between-person variables. Assumptions were checked using the visualization of the 'performance' package (Daniel et al., 2021). If the assumptions seemed to be violated, a robust model was fitted using the package 'robustlmm' (Manuel Koller, 2016) and compared to the non-robust version. Only the non-robust model was reported as no noticeable difference emerged between both versions of the models. The need of controlling for autocorrelation was also checked which improved the model and was therefore included in all models. A hierarchical approach was used for the inclusion of the control variables and the model fit was assessed with –2 restricted log-likelihood and the Akaike information criterion (AIC). A sensitivity test was also performed where participants with less than six measurements (n=24, 27%) were excluded from the analyses which yielded comparing β with similar significances. Therefore, the models including all 89 participants with valid measurements were used.

The predictors sleep quality, valence, energetic arousal, and calmness, and the control variable non-wear time were included at level 1 and centered at the person-mean to estimate within-person effects (Hoffman & Stawski, 2009). Additionally, the control variable weekday or weekend (i.e. weekday = 0, weekend = 1) was included in the models at level 1. Time (i.e. day of the study 0-20) was added as a within-person control variable at level 1 but showed no significant effect and was not included in the final models. The mean scores per person for each level 1 predictor were added as level 2 predictors to unravel the between-person from the within-person results (Hoffman & Stawski, 2009). adult/child (i.e. adult = 0, children = 1) was added as a between-person control variable at level 2. Sex (i.e. female = 0, male = 1) was only added for MVPA as a between-person control variable at level 2 as it did not improve the model for steps. Random slopes were computed for all level 1 predictors which did not improve the models and were therefore excluded in the final models. Random intercepts were used for both models and the level for significance was set a priori to $\alpha < 0.05$.

The equation of the final models (with the only difference that sex was excluded for the steps model) was:

Level 1 equation:

$$Y_{ij} = \beta_{0j} + \beta_{1j} * (sleep\ quality)_{ij} + \beta_{2j} * (valence)_{ij} + \beta_{3j} * (calmness)_{ij} + \beta_{4j} * (energetic\ arousal)_{ij} + \beta_{5j} * (non - wear\ time)_{ij} + \beta_{6j} * (wewd)_{ij} + r_{ij}$$

Level 2 equation:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} * (mean\ sleep\ quality)_j + \gamma_{02} * (mean\ valence)_j + \gamma_{03} * (mean\ calmness)_j + \gamma_{04} * (mean\ energetic\ arousal)_j + \gamma_{05}\ (adult/child) + \gamma_{06}\ (sex) + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

Results

Data availability and participant characteristics

Overall, 98 participants received a total of 2058 sleep quality EMAs over the 21-day collection period. The average number of sleep quality ratings completed by each participant was 17.39 out of 21, equating to 82.85% complete data (represents daily app use). On 1332 out of 2058 days, the participants additionally answered at least one EMA assessing valence, energetic arousal, and calmness (averaged from 2579 triggers). This implied that daily mean values for each affective state and each participant could be calculated for 64.72% of the days. Days (n = 775/2058) with missing values (2531 data points) in either sleep quality or affect ratings were excluded from the final analyses.

Additionally, 656/2058 days indicated either greater non-wear time than 960 minutes (618 data points) or that more than 1200 minutes were classified as an energy-expenditure range of 1.0 to 1.5 METs (34 data points) or that zero step counts were recorded (285 data points). Those days were also excluded from the analysis (some of which overlapped with the excluded days for sleep quality and/or affect).

The exclusion of days due to missing and invalid data points resulted in a final analytic sample of 49 adults (35-60 years) and 40 children (5-19 years) and a total of 996 days (adults = 661; children = 335)), yielding an average of 11.19 valid measurement occasions per participant (affect triggers were summarized from 1-9 measurements per day). Participant characteristics for the final sample are shown in Table 1. A daily overview of all outcomes and predictors divided by adults and children is visualized in Figure 1. Here, the variability of daily data and individual patterns of the variables over time for each participant can be inspected in supplement figures 1 to 7. The mean BMI was 25.38 (SD = 3.91) kg/m^2 in the adult and 17.62 (SD = 2.92) kg/m^2 in the children population.

Table 1 Descriptive data of all participants included in the analyses. Displayed are the means and standard deviations (SD) during three weeks for the parameters age, steps, moderate to vigorous physical activity (MVPA), non-wear time (nwt), self-rated sleep quality (sleep), self-rated valence, self-rated energetic arousal (energetic), and self-rated calmness.

population	adult		child	
sex	female (n=24)	male (n=25)	female (n=23)	male (n=17)
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
age (years)	44.8 (5.4)	46.6 (5.1)	11.2 (2.8)	11.9 (3.9)
steps (count/day)	7880 (2600)	6700 (2540)	8280 (3230)	9890 (2940)
MVPA (min/day)	56.0 (25.7)	59.3 (29.4)	54.7 (45.2)	117 (36.3)
nwt (min/day)	641 (84.5)	630 (98.1)	703 (98.9)	744 (96.3)
sleep (0-6)	4.25 (0.92)	3.82 (0.82)	4.14 (1.50)	4.51 (0.94)
valence (0-4)	2.82 (0.38)	2.89 (0.45)	3.09 (0.47)	3.29 (0.61)
energetic (0-6)	3.89 (0.67)	4.07 (0.74)	4.13 (1.21)	4.82 (0.93)
calmness (0-6)	2.71 (0.844)	2.98 (0.845)	2.73 (1.13)	3.04 (1.22)

Effects of sleep quality and affective states on daily step count

Within-person effects (Level-1)

Results indicate no significant within-person effects between sleep quality ratings and daily step count (see Table 2). As hypothesized, the daily average affective states rated by a person were associated with the number of device-based measured steps per day. In detail, a higher than average rating of a person's valence on one day significantly predicted a higher step count on the same day within this person (β = 0.06, p = .024). In practice, a 1-point increase in valence above the person-mean (original scale 0 – 4) was related to an average increase of 489.63 more steps on the same day. Furthermore, higher than average values of a person's energetic arousal ratings were related to an increase in that person's device-based step count on the same day (β = 0.07, p = .014). As expected, days with higher than average ratings of calmness within a person were associated with significantly lower device-based measured step count on the same day (β = -0.07, p = .007). The daily non-wear time showed a significant effect on steps. This means that on a day when a person wore the accelerometer one minute less than their person-based average, the accelerometer recorded 4.09 fewer steps (p < .001) for the same person. Additionally, the number of recorded steps of a person was significantly higher on weekend days than on weekdays (β = 0.09, p = .001).

Between-person effects (Level-2)

Results showed no significant between-person effects between sleep quality, valence or energetic arousal ratings, and device-based measured step count. However, individuals with higher average calmness ratings had significantly fewer daily steps recorded when compared to individuals with lower averages (β = -0.16, p = .023). Furthermore, significant differences between children and adults in the average number of steps per day (β = 0.16, p = .022) were found insofar as the accelerometers recorded 1452.05 more steps a day on average in children than in adults. Overall, the ICC showed that 35% of the variance in the model was due to between-person and 65% due to within-person variance.

Table 2 Multilevel model analysis for the influences of sleep quality and affective states on daily step count. Displayed are the within-person results (wp) of the person-mean centered variables self-rated sleep quality (sleep) (original range 0-6), self-rated valence (valence) (original range 0-4), self-rated energetic arousal (energetic) (original range 0-6), and self-rated calmness (calmness) (original range 0-6) and the within-person, person-mean centered control variable non-wear time (nwt) and the variable weekend/weekday (wewd, weekday = 0, weekend =1). Additionally, the between-person results (bp) of the affective states and sleep quality, and the influence of adult/child (adult = 0, children = 1) on steps are shown. All results are displayed using the raw Beta (B), the standardized Beta (β), 95% confidence intervals (CI), and standardized (std.) 95% CI. Additionally, the within-person variance (σ^2), the between-person variance (τ_{00} id), the intraclass correlation coefficient (ICC), the number of participants (N_{id}), and the number of observations are displayed.

steps					
Predictors	B	β	CI	std. CI	p
(Intercept)	6157.69	-0.02	2263.20 – 10052.18	-0.16 – 0.11	**.002**
wp_sleep	130.57	0.03	-52.73 – 313.87	-0.01 – 0.08	.165
wp_valence	489.63	0.06	66.15 – 913.10	0.01 – 0.12	**.024**
wp_calmness	-244.11	-0.07	-418.90 – -69.32	-0.11 – -0.02	**.007**
wp_energetic	339.23	0.07	71.45 – 607.01	0.01 – 0.12	**.014**
wp_nwt	-4.09	-0.11	-5.90 – -2.28	-0.16 – -0.06	**<.001**
wewd	848.32	0.09	340.50 – 1356.15	0.03 – 0.14	**.001**
bp_sleep	49.77	0.01	-587.52 – 687.05	-0.12 – 0.15	.878
bp_valence	843.16	0.09	-977.53 – 2663.85	-0.10 – 0.28	.363
bp_calmness	-778.42	-0.16	-1442.62 – -114.21	-0.29 – -0.02	**.023**
bp_energetic	147.51	0.03	-818.79 – 1113.81	-0.15 – 0.20	.764
bp_adult/child	1452.05	0.16	220.21 – 2683.90	0.02 – 0.30	**.022**
Random Effects					
σ^2	10220975.78				

$\tau_{00\ id}$	5595387.95
ICC	0.35
N_{id}	89
Observations	996

Effects of sleep quality and affective states on daily MVPA

Within-person effects (Level-1)

As shown in Table 3, neither perceived sleep quality (p = .434) nor mean energetic arousal (p = .070) ratings of one day were associated with MVPA during the same day within a person. As hypothesized, the daily average valence ratings per day significantly predicted higher MVPA (β = 0.07, p = .006). In practice, a 1-point increase in valence above the person-mean (original scale 0 – 4) was related to an average increase of 6.55 more minutes of MVPA on the same day. Furthermore, days with higher than average ratings of calmness within a person were associated with significantly lower device-based measured time spent in MVPA (β = -0.04, p = .035). Additionally, the analyses revealed that non-wear time significantly predicted lower daily recorded MVPA (β = -0.07, p = .001) and higher MVPA was recorded during weekend days compared to weekdays (β = 0.06, p = .006).

Between-person effects (Level-2)

Differences in person-mean ratings of sleep quality and affective states between-persons did not predict daily MVPA (see Table 3). However, results for the control variables showed that children had recorded significantly more MVPA than adults (β = 0.26, p = .001). In addition, significant sex differences (p < .001) were found for daily MVPA where being male was associated with higher daily MVPA values. Overall, the ICC showed that 47% of the variance in the model was due to between-person and 53% due to within-person variance.

Table 3 Multilevel model analysis for the influences of sleep quality and affective states on daily MVPA. Displayed are the within-person results (wp) of the person-mean centered variables self-rated sleep quality (sleep) (original range 0-6), self-rated valence (valence) (original range 0-4), self-rated energetic arousal (energetic) (original range 0-6), and self-rated calmness (calmness) (original range 0-6) and the within-person, person-mean centered control variable non-wear time (nwt), and the variable weekend/weekday (wewd, weekday = 0, weekend =1). Additionally, the between-person results (bp) of the affective states and sleep quality, and the influence of adult/child (adult = 0, children = 1) and sex (0 = female, 1 = male) on MVPA are shown. All results are displayed using the raw Beta (B), the standardized Beta (β), 95% confidence intervals (CI), and standardized (std.) 95% CI. Additionally, the within-person variance (σ^2), the between-person variance (τ_{00} id), the intraclass correlation coefficient (ICC), the number of participants (N_{id}), and the number of observations are displayed.

MVPA					
Predictors	B	β	CI	std. CI	p
(Intercept)	66.72	-0.03	14.70 – 118.74	-0.18 – 0.12	**.013**
wp_sleep	0.81	0.02	-1.20 – 2.81	-0.03 – 0.06	.434
wp_valence	6.55	0.07	1.91 – 11.19	0.02 – 0.12	**.006**
wp_calmness	-2.07	-0.04	-3.98 – -0.15	-0.09 – -0.00	**.035**
wp_energetic	2.72	0.04	-0.21 – 5.65	-0.00 – 0.09	.070
wp_nwt	-0.03	-0.07	-0.05 – -0.01	-0.12 – -0.03	**.001**
wewd	7.89	0.06	2.29 – 13.49	0.02 – 0.11	**.006**
bp_sleep	0.65	0.01	-7.94 – 9.24	-0.13 – 0.16	.881
bp_valence	-5.85	-0.05	-30.29 – 18.59	-0.25 – 0.15	.637
bp_calmness	-7.41	-0.12	-16.37 – 1.55	-0.26 – 0.03	.106
bp_energetic	2.42	0.04	-10.70 – 15.54	-0.16 – 0.23	.717
bp_adult/child	28.40	0.26	11.69 – 45.12	0.11 – 0.41	**.001**
bp_sex	29.84	0.28	13.64 – 46.04	0.13 – 0.44	**<.001**
Random Effects					
σ^2	1249.24				
$\tau_{00\ id}$	1101.01				
ICC	0.47				
N_{id}	89				
Observations	996				

Discussion

This study used EMA and accelerometry to evaluate the within-person effects of children's and adults' daily self-reported valence, energetic arousal, and calmness on device-based measured MVPA and step count of the same day. Furthermore, the between-person effects of these variables were explored along with the within- and between-person effects of perceived sleep quality on steps and MVPA. The results mainly confirmed the hypotheses that ratings above the person-mean for valence and energetic arousal increased PA while calmness decreased PA on a within-person level. One exception was the relationship between energetic arousal and MVPA which was not significant (p = .07) but showed a standardized estimate in the hypothesized direction. For the exploration of sleep quality as well as between-person effects of the predictors, only calmness showed a significant prediction for steps indicating that participants who rated their calmness one point higher (scale 0-6) had recorded 778.42 fewer steps per day on average. Additionally, the results of the included control variables showed significant effects. Here, being male (only for MVPA) showed the largest effect, followed by being a child, having increased accelerometer wear time, and the measurement being on a weekend day. The results of this study mainly confirm the findings of previous studies using time-lagged predictors, indicating that the relation between affect and PA is solid through different age groups (Cushing et al., 2017; Dunton et al., 2014; Koch et al., 2018; Liao et al., 2017; Niermann et al., 2016; Reichert et al., 2016; Schwerdtfeger et al., 2010), that affect and PA results are related on a day level (Do et al., 2021), and add that these findings also apply to intervention studies.

Valence and physical activity

The use and definitions of mood and affect and their subitems have been used interchangeable in the previous literature (Liao et al., 2015; Niermann et al., 2016), therefore, the results of previous studies related to all mood dimensions and affective states are treated as equal in this paragraph to provide a broader picture even though they often assess different constructs (see Ekkekakis, 2013 for an overview). In our study, valence significantly predicted both steps and MVPA at the same day within-persons which is in accordance with some previous studies (Dunton et al., 2014; Koch et al., 2018; Liao et al., 2017; Niermann et al., 2016; Reichert et al., 2016; Schwerdtfeger et al., 2010) while another study did not find such relation between affect and PA in adolescents (Cushing et al., 2017). This strengthens the view that valence should be considered in building up PA interventions and that valence is a promising target to tailor interventions to a persons’ needs in a randomized controlled trial (Conroy et al., 2020). If for example, an assessment of valence indicates a low rating in a person compared to the usual rating, the most promising intervention might not be to target PA directly, but to improve valence and by doing so increase the probability that the person will engage in PA throughout the day.

Energetic arousal and physical activity

Energetic arousal predicted steps but not MVPA on the same day within-persons while previous studies found an association for different PA measures as time-lagged predictors (Dunton et al., 2014; Koch et al., 2018; Liao et al., 2017; Reichert et al., 2016). Here, two studies, one in adults (Reichert et al., 2016) and one in adolescents (Koch et al., 2018), indicated that the relationship of energetic arousal to nonexercise activity is stable in a timeframe of up to 300 minutes. The relation of energetic arousal to MVPA however, has only been found for shorter (i.e. 30 minutes) timeframes in children (Dunton et al., 2014) and another study found no relationship between energy (measured as a single item on a 1-5 scale) and MVPA 15 or 30 minutes after the EMA assessment in adults (Liao et al., 2017). Therefore, the relation of energetic arousal and MVPA might follow a narrower time pattern while the relationship to steps as a parameter without intensity indication seems to be more time stable. These results should be considered for PA interventions as the influence of energetic arousal on PA seems to depend on the PA outcome and/or intensity. In this case, including or targeting energetic arousal in an intervention seems to be most beneficial for nonexercise activity or overall PA outcomes like step count. Here, digital games could be used to enhance energetic arousal (Collins et al., 2019) which could then be followed by a prompt to engage in PA.

Calmness and physical activity

Self-rated calmness predicted both reduced recorded step count and reduced recorded MVPA per day within-persons which confirms findings of previous studies using it as a time-lagged predictor (Koch et al., 2018; Reichert et al., 2016). This means, that if a person rated their calmness higher on a certain day than their average calmness this person engaged in less PA on that day. Therefore, having calmness included in the context of positive affect or overall mood (e.g. in Liao et al., 2017) might influence the result of the construct. Here, future studies which aim to investigate the relation between affect and PA should measure calmness as a separate construct from positive affect. This indicates the need to explore the relation of different sub-categories of affect and mood. Here, the timeframe in which self-rated calmness predicted less subsequent PA was measurable in up to 130 minutes in one study (Koch et al., 2018) and up to 140 minutes in another study (Reichert et al., 2016) which suggests fairly time stable results for this parameter. Therefore, calmness seems to be another affective state which has to be considered for PA interventions. It is however questionable to reduce calmness to influence PA because calmness is also an important factor for health (Huffziger et al., 2013). In this case, more context about general PA behavior of the person and covariates of health behavior along with a clearer picture of a dose-response relationship between covariates and PA behavior is needed to provide individualized recommendations for behavior change interventions.

Affect and related parameters between-persons

The association between affective states, sleep quality, and PA on a between-person level has been explored by this study. None of the affective states were associated with the average recorded time spent in MVPA per day. Regarding daily step counts, significant between-person variations were found in individuals with higher calmness ratings. These results suggest that individuals who feel calmer on average had recorded fewer steps on average than individuals who feel more stressed. No between-person association between energetic arousal, valence, and daily step count was found. Therefore, the results indicate that days, where a person rated their daily valence or energetic arousal higher than their usual daily valence or energetic arousal (within-person), showed enhanced PA. However, there was no difference for persons who rated their valence or energetic arousal higher on average compared to persons who rated their valence or energetic arousal lower (between-person) concerning PA. Further research is needed to define the link between affective states and PA behavior on the between-person level under consideration of the within-person level to specify if the observed relation is due to individual differences of participants or changes over time within participants, or both (Dunton, 2017).

Sleep quality and physical activity variation within- and between-persons

The exploratory within- and between-person analyses indicated no significant association between the subjectively assessed sleep quality and device-based measured daily step count or daily MVPA. These results fit the findings of previous experimental and cross-sectional studies, which showed inconsistent relations between various sleep characteristics (e.g., efficiency, duration) and PA outcomes in children and adults (Antczak et al., 2020; Semplonius & Willoughby, 2018). In another study, Eythorsdottir et al. (2020) used device-based measured sleep and PA outcomes and found no significant association between children with different sleep duration and sleep efficiency and PA. One explanation for these findings could be the high heterogeneity in the measurements or outcomes (e.g. sleep quality, sleep duration, wake time, bedtime) used. Further studies need to be designed which examine the bidirectional and temporal aspects of the sleep-physical behavior associations. Additionally, to understand the effects of sleep characteristics (e.g., duration, efficiency, latency) on PA in more detail, more long-term studies with objectively measured sleep parameters are needed (e.g. using heart rate variability (Stein & Pu, 2012)), which consider daily schedules and further motivational aspects of activity behavior. Furthermore, as sleep-related measures often differ between adults and children (e.g. earlier bedtime and longer periods of nocturnal sleep for children) even within parent-child dyads (Gau, Susan, Shur-Fen & Merikangas, Kathleen, R., 2004), studies comparing the results of both groups separately would benefit the understanding of the association between sleep quality and PA.

Control variables

The findings of this study suggest significant differences in both recorded steps and MVPA outcomes between children and adults. Here, children showed a higher mean step count per day and higher time in MVPA than adults. Future studies should also investigate if the relationship between affective states, and sleep and PA differs due to e.g. developmental differences throughout the lifetime. Additionally, the results also indicate that time spent in MVPA during one day differed between sexes, suggesting that men and boys spent more time per day in MVPA than women and girls. Descriptive data of this study shows that while boys move more than girls (also illustrated in recent research (McGovern et al., 2020)), women had a higher step count than men which is overlaid by the difference between boys and girls and therefore only visible in sex- and child/adult-disaggregated data. These results suggest that age-related sex differences should be considered when designing, implementing, and evaluating PA interventions for children and adults (Schlund et al., 2021). Furthermore, our study found non-wear time and differences between weekdays and weekend days to influence both steps and MVPA. A higher non-wear time predicted less PA during the day and participants were more active on weekends compared to during the weekdays. Those variables should always be considered when interpreting PA outcomes if data has been measured over several days even if data with a certain wear time (i.e. less than 8 hours per day) were excluded.

Limitations

There are some limitations of this study that have to be considered for the interpretation of the results. First, the data of this study were collected during the ongoing COVID-19 pandemic which might differ from results before the pandemic, as certain restrictions have probably influenced PA patterns (Stockwell et al., 2021) and people's affective states (Panayiotou et al., 2021). However, data collection has only been conducted when schools were open to allow comparability of the data. Secondly, the study focusses on EMA in an intervention design that aimed at increasing PA of the participants (without directly targeting the predictor variables) and included also other factors like health literacy and goal setting. We accounted for this aspect by controlling for days in the study which showed no significant effect and by visualizing the variability within the measurements and the individual development in supplement figures 1-7. Additionally, the EMA triggers for valence, energetic arousal, and calmness were sent after a period of inactivity was detected by the accelerometer in addition to a trigger in the evening when the participants finished the day in the applications (provided no trigger was sent during the previous hour). Therefore, the number of triggers sent per day (which responses were then averaged, range 1-9) varied between the days and persons. The large amount of missing data also needs to be addressed which limits the generalizability of the findings. However, as stated in the method section, sensibility analyses yielded comparing results and multilevel approaches are fairly robust to missing data. Moreover, daily mean values were used in this study instead of time-lagged predictors which are important for the interpretation as PA and affect can have a bidirectional relationship (Liao et al., 2015). As it is not known if the participants were active before, during, or after the assessments in this study, the multilevel modeling results concern the overall association of the measures on a certain day but are not related to the question of the time-related direction of the effect. Furthermore, it is unclear if the daily mean values for valence, energetic arousal and calmness are representative of the person's average as they were answered once to multiple times per day and the different parameters of affect are known to change throughout the day (Reichert et al., 2020). Finally, the selection of epoch lengths is important to consider for PA estimations by accelerometers, especially if both adults and children are included in the study (Fiedler et al., 2021). The choice of another epoch length (e.g. 10-second epochs instead of 60-second epochs) might have led to differing findings for intensity-related parameters (i.e. MVPA).

Conclusion

The study expands previous findings from studies examining the dynamic relations of PA, sleep quality and affective states by considering the whole day instead of shorter timeframes, focusing on multiple outcome parameters and predictors during an intervention period, and by including both adults and children as participants of the study. The results confirm that every one unit increase in self-rated energetic arousal and valence was associated with 336.23 and

489.64 higher step count per day respectively, while every one unit increase in valence was associated with 6.55 more minutes MVPA while energetic arousal was not associated with MVPA. Additionally, an one unit increase in calmness was associated with 244.11 fewer steps and 2.07 fewer minutes MVPA per day. The additional exploration found sex, age, non-wear time, and the differentiation between weekday and weekend as important covariates and control variables for PA. Overall, this study shows that affective states are important predictors for PA and should be included in the development of effective mHealth interventions to facilitate health behavior change. Future EMA studies should explore the dose-response relationship for predictors and covariates of PA while future intervention studies should consider the known associations between predictors and PA as possible targets for individual tailoring of the interventions. In doing so, barriers for PA uptake can be identified and targeted by including the individual needs for each person under a variety of circumstances into the equation and form the basis for highly individualized just-in-time adaptive interventions.

References

Anastasopoulou, P., Härtel, S., & Hey, S. (2013). A Comparison of Two Commercial Activity Monitors for Measuring Step Counts During Different Everyday Life Walking Activities. *International Journal of Sports Science and Engineering*, *7*(1), 31–35.

Anastasopoulou, P., Tubic, M., Schmidt, S., Neumann, R., Woll, A., & Härtel, S. (2014). Validation and comparison of two methods to assess human energy expenditure during free-living activities. *PloS One*, *9*(2), e90606. https://doi.org/10.1371/journal.pone.0090606

Antczak, D., Lonsdale, C., Lee, J., Hilland, T., Duncan, M. J., Del Pozo Cruz, B., Hulteen, R. M., Parker, P. D., & Sanders, T. (2020). Physical activity and sleep are inconsistently related in healthy children: A systematic review and meta-analysis. *Sleep Medicine Reviews*, *51*, 101278. https://doi.org/10.1016/j.smrv.2020.101278

Atoui, S., Chevance, G., Romain, A.-J., Kingsbury, C., Lachance, J.-P., & Bernard, P. (2021). Daily associations between sleep and physical activity: A systematic review and meta-analysis. *Sleep Medicine Reviews*, *57*, 101426. https://doi.org/10.1016/j.smrv.2021.101426

Bachmann, A., Klebsattel, C., Budde, M., Riedel, T., Beigl, M., Reichert, M., Santangelo, P., & Ebner-Priemer, U. (2015). How to use smartphones for less obtrusive ambulatory mood assessment and mood recognition. In K. Mase, M. Langheinrich, D. Gatica-Perez, H. Gellersen, T. Choudhury, & K. Yatani (Eds.), *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers - UbiComp '15* (pp. 693–702). ACM Press. https://doi.org/10.1145/2800835.2804394

Beaglehole, R., Bonita, R., Horton, R., Adams, C., Alleyne, G., Asaria, P., Baugh, V., Bekedam, H., Billo, N., Casswell, S., Cecchini, M., Colagiuri, R., Colagiuri, S., Collins, T., Ebrahim, S., Engelgau, M., Galea, G., Gaziano, T., Geneau, R., . . . Watt, J. (2011). Priority actions for the non-communicable disease crisis. *The Lancet*, *377*(9775), 1438–1447. https://doi.org/10.1016/S0140-6736(11)60393-0

Bourke, M., Hilland, T. A., & Craike, M. (2021). A systematic review of the within-person association between physical activity and affect in children's and adolescents' daily lives. *Psychology of Sport and Exercise*, *52*, 101825. https://doi.org/10.1016/j.psychsport.2020.101825

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C., Chaput, J.-P., Chastin, S., Chou, R., Dempsey, P. C., DiPietro, L., Ekelund, U., Firth, J., Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., . . . Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British Journal of Sports Medicine*, *54*(24), 1451–1462. https://doi.org/10.1136/bjsports-2020-102955

Carels, R. A., Coit, C., Young, K., & Berg, A. (2007). Exercise Makes You Feel Good, But Does Feeling Good Make You Exercise? An Examination of Obese Dieters. *Journal of Sport and Exercise Psychology*, *29*(6), 706–722. https://doi.org/10.1123/jsep.29.6.706

Collins, E., Cox, A., Wilcock, C., & Sethu-Jones, G. (2019). Digital Games and Mindfulness Apps: Comparison of Effects on Post Work Recovery. *JMIR Mental Health*, *6*(7), e12853. https://doi.org/10.2196/12853

Conroy, D. E., Lagoa, C. M., Hekler, E., & Rivera, D. E. (2020). Engineering Person-Specific Behavioral Interventions to Promote Physical Activity. *Exercise and Sport Sciences Reviews*, *48*(4), 170–179. https://doi.org/10.1249/JES.0000000000000232

Cushing, C. C., Mitchell, T. B., Bejarano, C. M., Walters, R. W., Crick, C. J., & Noser, A. E. (2017). Bidirectional Associations Between Psychological States and Physical Activity in Adolescents: A mHealth Pilot Study. *Journal of Pediatric Psychology*, *42*(5), 559–568. https://doi.org/10.1093/jpepsy/jsw099

Daniel, L., Mattan, S. B.-S., Indrajeet, P., Philip, W., & Dominique, M. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, *60*(6), 3139. https://doi.org/10.21105/joss.03139

Daniel Lüdecke. (2021). *sjPlot: Data Visualization for Statistics in Social Science* [Computer software]. https://CRAN.R-project.org/package=sjPlot

Do, B., Wang, S. D., Courtney, J. B., & Dunton, G. F. (2021). Examining the day-level impact of physical activity on affect during the early months of the COVID-19 pandemic: An ecological momentary assessment study. *Psychology of Sport and Exercise*, *56*. https://doi.org/10.1016/j.psychsport.2021.102010

Dunton, G. F. (2017). Ecological Momentary Assessment in Physical Activity Research. *Exercise and Sport Sciences Reviews*, *45*(1), 48–54. https://doi.org/10.1249/JES.0000000000000092

Dunton, G. F., Huh, J., Leventhal, A. M., Riggs, N., Hedeker, D., Spruijt-Metz, D., & Pentz, M. A. (2014). Momentary assessment of affect, physical feeling states, and physical activity in children. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, *33*(3), 255–263. https://doi.org/10.1037/a0032640

Edney, S., Ryan, J. C., Olds, T., Monroe, C., Fraysse, F., Vandelanotte, C., Plotnikoff, R., Curtis, R., & Maher, C. (2019). User Engagement and Attrition in an App-Based Physical Activity Intervention: Secondary Analysis of a Randomized Controlled Trial. *Journal of Medical Internet Research*, *21*(11), e14645. https://doi.org/10.2196/14645

Ekkekakis, P. (2013). *The Measurement of Affect, Mood, and Emotion: A Guide for Health-Behavioral Research*. Cambridge University Press.

Emerson, J. A., Dunsiger, S., & Williams, D. M. (2018). Reciprocal within-day associations between incidental affect and exercise: An EMA study. *Psychology & Health*, *33*(1), 130–143. https://doi.org/10.1080/08870446.2017.1341515

Eythorsdottir, D. Y., Frederiksen, P., Larsen, S. C., Olsen, N. J., & Heitmann, B. L. (2020). Associations between objective measures of physical activity, sleep and stress levels among preschool children. *BMC Pediatrics*, *20*(1), 258. https://doi.org/10.1186/s12887-020-02108-7

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 (Basel, Switzerland)*, *21*(8). https://doi.org/10.3390/s21082672

Fiedler, J., Eckert, T., Wunsch, K., & Woll, A. (2020). Key facets to build up eHealth and mHealth interventions to enhance physical activity, sedentary behavior and nutrition in healthy subjects - an umbrella review. *BMC Public Health*, *20*(1), 1605. https://doi.org/10.1186/s12889-020-09700-7

Forster, A. K., Richards, E. A., Foli, K. J., McGowan, B., Hass, Z., Becker, M., & Miller, A. (2021). Influence of Affect on Physical Activity: An Integrative Review. *Clinical Nursing Research*, *30*(7), 934–949. https://doi.org/10.1177/1054773820968039

Gau, Susan, Shur-Fen, & Merikangas, Kathleen, R. (2004). Similarities and differences in sleep-wake patterns among adults and their children. *Sleep*, *27*(2), 299–304.

Hadley Wickham. (2016). *ggplot2: Elegant Graphics for Data Analysis* [Computer software]. Springer-Verlag New York. https://ggplot2.tidyverse.org

Hallal, P. C., Andersen, L. B., Bull, F. C., Guthold, R., Haskell, W., & Ekelund, U. (2012). Global physical activity levels: surveillance progress, pitfalls, and prospects. *The Lancet*, *380*(9838), 247–257. https://doi.org/10.1016/S0140-6736(12)60646-1

Hoffman, L., & Stawski, R. S. (2009). Persons as Contexts: Evaluating Between-Person and Within-Person Effects in Longitudinal Analysis. *Research in Human Development*, *6*(2-3), 97–120. https://doi.org/10.1080/15427600902911189

Huffziger, S., Ebner-Priemer, U., Zamoscik, V., Reinhard, I., Kirsch, P., & Kuehner, C. (2013). Effects of mood and rumination on cortisol levels in daily life: An ambulatory assessment study in remitted depressed patients and healthy controls. *Psychoneuroendocrinology*, *38*(10), 2258–2267. https://doi.org/10.1016/j.psyneuen.2013.04.014

Jose Pinheiro, Douglas Bates, Saikat DebRoy, Deepayan Sarkar, & R Core Team. (2021). *nlme: Linear and Nonlinear Mixed Effects Models* [Computer software]. https://CRAN.R-project.org/package=nlme

Kanning, M. K., & Schoebi, D. (2016). Momentary Affective States Are Associated with Momentary Volume, Prospective Trends, and Fluctuation of Daily Physical Activity. *Frontiers in Psychology*, *7*, 744. https://doi.org/10.3389/fpsyg.2016.00744

Khan, M. K. A., Chu, Y. L., Kirk, S. F. L., & Veugelers, P. J. (2015). Are sleep duration and sleep quality associated with diet quality, physical activity, and body weight status? A population-based study of Canadian children. *Canadian Journal of Public Health = Revue Canadienne De Sante Publique*, *106*(5), e277-82. https://doi.org/10.17269/cjph.106.4892

Koch, E. D., Tost, H., Braun, U., Gan, G., Giurgiu, M., Reinhard, I., Zipf, A., Meyer-Lindenberg, A., Ebner-Priemer, U., & Reichert, M. (2018). Mood Dimensions Show Distinct Within-Subject Associations With Non-exercise Activity in Adolescents: An Ambulatory Assessment Study. *Frontiers in Psychology*, *9*, 268. https://doi.org/10.3389/fpsyg.2018.00268

Kohl, H. W., Craig, C. L., Lambert, E. V., Inoue, S., Alkandari, J. R., Leetongin, G., & Kahlmeier, S. (2012). The pandemic of physical inactivity: global action for public health. *The Lancet*, *380*(9838), 294–305. https://doi.org/10.1016/S0140-6736(12)60898-8

Liao, Y., Chou, C.-P., Huh, J., Leventhal, A., & Dunton, G. F. (2017). Examining acute bi-directional relationships between affect, physical feeling states, and physical activity in free-living situations using electronic ecological momentary assessment. *Journal of Behavioral Medicine*, *40*(3), 445–457. https://doi.org/10.1007/s10865-016-9808-9

Liao, Y., Shonkoff, E. T., & Dunton, G. F. (2015). The Acute Relationships Between Affect, Physical Feeling States, and Physical Activity in Daily Life: A Review of Current Evidence. *Frontiers in Psychology*, *6*, 1975. https://doi.org/10.3389/fpsyg.2015.01975

Manuel Koller (2016). robustlmm: An R Package for Robust Estimation of Linear Mixed-Effects Models. *Journal of Statistical Software, 75*(6), 1–24. https://doi.org/10.18637/jss.v075.i06

McGovern, J., Drewson, S. R., Hope, A., & Konopack, J. F. (2020). Gender Differences in a Youth Physical Activity Intervention: Movement Levels and Children's Perceptions. *American Journal of Health Education*, *51*(2), 109–119. https://doi.org/10.1080/19325037.2020.1712667

Niermann, C., Herrmann, C., Haaren, B. von, van Kann, D., & Woll, A. (2016). Affect and Subsequent Physical Activity: An Ambulatory Assessment Study Examining the Affect-Activity Association in a Real-Life Context. *Frontiers in Psychology*, *7*, 677. https://doi.org/10.3389/fpsyg.2016.00677

Paluch, A. E., Gabriel, K. P., Fulton, J. E., Lewis, C. E., Schreiner, P. J., Sternfeld, B., Sidney, S., Siddique, J., Whitaker, K. M., & Carnethon, M. R. (2021). Steps per Day and All-Cause Mortality in Middle-aged Adults in the Coronary Artery Risk Development in Young Adults Study. *JAMA Network Open*, *4*(9), e2124516. https://doi.org/10.1001/jamanetworkopen.2021.24516

Panayiotou, G., Panteli, M., & Leonidou, C. (2021). Coping with the invisible enemy: The role of emotion regulation and awareness in quality of life during the COVID-19 pandemic. *Journal of Contextual Behavioral Science*, *19*, 17–27. https://doi.org/10.1016/j.jcbs.2020.11.002

R Core Team. (2021). *R: A Language and Environment for Statistical Computing* [Computer software]. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

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., Brüβler, 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

Reichert, M., Tost, H., Reinhard, I., Schlotz, W., Zipf, A., Salize, H.-J., Meyer-Lindenberg, A., & Ebner-Priemer, U. W. (2017). Exercise versus Nonexercise Activity: E-diaries Unravel Distinct Effects on Mood. *Medicine and Science in Sports and Exercise*, *49*(4), 763–773. https://doi.org/10.1249/MSS.0000000000001149

Reichert, M., Tost, H., Reinhard, I., Zipf, A., Salize, H.-J., Meyer-Lindenberg, A., & Ebner-Priemer, U. W. (2016). Within-Subject Associations between Mood Dimensions and Non-exercise Activity: An Ambulatory Assessment Approach Using Repeated Real-Time and Objective Data. *Frontiers in Psychology*, *7*, 918. https://doi.org/10.3389/fpsyg.2016.00918

RStudio Team. (2021). *RStudio: Integrated Development Environment for R [*Computer software]. RStudio, PBC. Boston, MA. http://www.rstudio.com/

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145–172. https://doi.org/10.1037/0033-295x.110.1.145

Schimmack, U., & Grob, A. (2000). Dimensional models of core affect: a quantitative comparison by means of structural equation modeling. *European Journal of Personality*, *14*(4), 325–345. https://doi.org/10.1002/1099-0984(200007/08)14:4<325::AID-PER380>3.3.CO;2-9

Schlund, A., Reimers, A. K., Bucksch, J., Brindley, C., Schulze, C., Puil, L., Coen, S. E., Phillips, S. P., Knapp, G., & Demetriou, Y. (2021). Do Intervention Studies to Promote Physical Activity and Reduce Sedentary Behavior in Children and Adolescents Take Sex/Gender Into Account? A

Systematic Review. *Journal of Physical Activity & Health*, *18*(4), 461–468. https://doi.org/10.1123/jpah.2020-0666

Schwerdtfeger, A., Eberhardt, R., Chmitorz, A., & Schaller, E. (2010). Momentary Affect Predicts Bodily Movement in Daily Life: An Ambulatory Monitoring Study. *Journal of Sport and Exercise Psychology*, *32*(5), 674–693.

Semplonius, T., & Willoughby, T. (2018). Long-Term Links between Physical Activity and Sleep Quality. *Medicine and Science in Sports and Exercise*, *50*(12), 2418–2424. https://doi.org/10.1249/MSS.0000000000001706

Snyder, E., Cai, B., DeMuro, C., Morrison, M. F., & Ball, W. (2018). A New Single-Item Sleep Quality Scale: Results of Psychometric Evaluation in Patients With Chronic Primary Insomnia and Depression. *Journal of Clinical Sleep Medicine : JCSM : Official Publication of the American Academy of Sleep Medicine*, *14*(11), 1849–1857. https://doi.org/10.5664/jcsm.7478

Stein, P. K., & Pu, Y. (2012). Heart rate variability, sleep and sleep disorders. *Sleep Medicine Reviews*, *16*(1), 47–66. https://doi.org/10.1016/j.smrv.2011.02.005

Stockwell, S., Trott, M., Tully, M., Shin, J., Barnett, Y., Butler, L., McDermott, D., Schuch, F., & Smith, L. (2021). Changes in physical activity and sedentary behaviours from before to during the COVID-19 pandemic lockdown: A systematic review. *BMJ Open Sport & Exercise Medicine*, *7*(1), e000960. https://doi.org/10.1136/bmjsem-2020-000960

Trull, T. J., & Ebner-Priemer, U. W. (2013). Ambulatory assessment. *Annual Review of Clinical Psychology*, *9*, 151–176. https://doi.org/10.1146/annurev-clinpsy050212-185510

Vandelanotte, C., Müller, A. M., Short, C. E., Hingle, M., Nathan, N., Williams, S. L., Lopez, M. L., Parekh, S., & Maher, C. A. (2016). Past, Present, and Future of eHealth and mHealth Research to Improve Physical Activity and Dietary Behaviors. *Journal of Nutrition Education and Behavior*, *48*(3), 219-228.e1. https://doi.org/10.1016/j.jneb.2015.12.006

Wang, F., & Boros, S. (2021). The effect of physical activity on sleep quality: a systematic review. *European Journal of Physiotherapy*, *23*(1), 11–18. https://doi.org/10.1080/21679169.2019.1623314

Wilhelm, P., & Schoebi, D. (2007). Assessing Mood in Daily Life. *European Journal of Psychological Assessment*, *23*(4), 258–267. https://doi.org/10.1027/1015-5759.23.4.258

Wunsch, K., Eckert, T., Fiedler, J., Cleven, L., Niermann, C., Reiterer, H., Renner, B., & Woll, A. (2020). Effects of a Collective Family-Based Mobile Health Intervention Called "SMARTFAMILY" on Promoting Physical Activity and Healthy Eating: Protocol for a Randomized Controlled Trial. *JMIR Research Protocols*, *9*(11), e20534. https://doi.org/10.2196/20534

Wunsch, K., Eckert, T., Fiedler, J., & Woll, A. (2022). Just-in-time adaptive interventions in mobile physical activity interventions – A synthesis of frameworks and future directions. *The European Health Psychologist*.

Chapter 7 General discussion

In our work, we delved into a variety of questions around mobile health interventions for physical activity behavior change for healthy participants. Overall, mobile health interventions for physical activity are a large and growing research field with many still unanswered questions. There are so many different aspects that have to be considered for mobile health interventions that it is of uttermost importance to gain an overview of the most potent influences for a certain behavior, environment, and participant group. Our work aimed to refine the focus on promising and yet understudied aspects of mobile health interventions for healthy participants under consideration of previous findings. Our five articles found

1. mobile health interventions to be potentially effective which was facilitated by theoretical foundation and behavior change techniques (Fiedler et al., 2020), while the influence of just-in-time adaptive interventions was underreported and methodological issues limited the comparability of studies;
2. differences between device-based measured and self-reported physical activity are seldom stable over time (Fiedler et al., 2021), limiting comparability between different measures;
3. just-in-time adaptive interventions for physical activity show great potential for behavior change (Wunsch et al., 2022) and could be designed and reported using our adapted framework;
4. engaging with just-in-time adaptive interventions is associated with enhanced physical activity up to two hours after the trigger (Fiedler et al., submitted) while distinctive effects appear for step vs. metabolic equivalent count in the longer timeframes;
5. core affect is associated with physical activity during intervention studies (Fiedler et al., 2022) and could be a valuable addition to intervention designs.

To gain a better impression of our work, it is important to consider some limitations of the articles included in this thesis.

Our first overview article (Fiedler et al., 2020) strongly depended on the detailed reporting of our selected key facets in the included reviews. While the key facets of choice were selected based on previous literature (Glanz et al., 2008; Hardeman et al., 2019; Heron & Smyth, 2010; Morrison et al., 2012; Prestwich et al., 2014; Schembre et al., 2018; Umberson et al., 2010; Viner et al., 2012; Webb et al., 2010), there might be several additional important factors which are related to intervention effectiveness but have not been evaluated in greater detail within our review. Additionally, important published studies might not have been included as they were not yet part of any review. This might be one reason why the most recent developments like just-in-time adaptive interventions were not found by our umbrella review.

For our second article (Fiedler et al., 2021) concerning reliability, comparability, and stability of different physical activity measures, it needs to be considered that the true amount of physical activity or energy expenditure of the participants remains unknown as no gold standard measurement like direct observation or indirect calorimetry as criterion measures have been used in this examination (Burchartz et al., 2020; Keadle et al., 2019). Additionally, the generalizability of this explorative study is limited by the rather small sample size and the use of the spearman correlation, due to the distribution of the data and the number of comparisons, instead of intraclass correlation coefficients (Koo & Li, 2016; Liu et al., 2016) or Bland Altman plots (Giavarina, 2015).

Our synthesis of just-in-time adaptive intervention frameworks (Wunsch et al., 2022) is mainly limited by the novelty of the research field, limited studies on effectiveness, and a non-systematic approach to the literature search in this position paper.

The main limitation of our fourth article (Fiedler et al., submitted) regarding the engagement with just-in-time adaptive interventions was the lack of a true control condition that would have allowed us to draw conclusions about causality. Furthermore, participants of this examination were already quite active which limited the occasions when the trigger has been sent and therefore limited the number of level 1 assessments, especially in children. Since this was a secondary examination of a larger study, additional intervention aspects like goal setting and the social context might have influenced the results.

This limitation is also true for our fifth article (Fiedler et al., 2022). Furthermore, the results of this article are limited by the accumulation of ecological momentary assessments based on a limited number of questions per day which were mostly triggered after a period of inactivity. A large amount of missing data also needs to be recognized.

Under consideration of these limitations and the fact that mobile health interventions are a relatively new and fast-evolving field, our work contributed to I) an overview of the current evidence on electronic and mobile health effectiveness for healthy participants, II) evaluating methodological issues for the measurement of physical activity, and III) the conceptualization, evaluation, and possible future directions for an understudied potential key facet for mobile health interventions (i.e. just-in-time adaptive interventions).

Nevertheless, there are plenty of open questions about the effectiveness of mobile health interventions which have not been included in our work and should be considered by future research. We will evaluate some of them in the following chapters.

Question 1

What are important considerations to progress future mobile health studies for physical activity promotion?

Our umbrella review (Fiedler et al., 2020) confirmed that theoretical foundation and behavior change techniques are important facets of effective mobile health interventions. What remains unclear is which theories or behavior change techniques are especially promising for different settings and participants and if dynamically changing theories are a better fit than traditional theories for the fast-evolving field (Martín et al., 2014; Navarro-Barrientos et al., 2011; Riley et al., 2011; Riley et al., 2016). Additionally, the impact of social context on effectiveness remains unclear and should be examined more thoroughly e.g. regarding workplace, or family interventions and the role of social support or competition regarding specific participant samples (Buckingham et al., 2019; Tong & Laranjo, 2018; Wunsch et al., 2020). Further important issues not included in our review regard for example differences in the socioeconomic status of the target group where Western and colleagues (2021) found no evidence for mobile health effectiveness in participants with low socioeconomic status. This is especially troublesome as noncommunicable diseases have a higher prevalence and more severe health consequences in this group (Lago-Peñas et al., 2021) and effective interventions for behavior change could be beneficial to prevent noncommunicable disease development in this target group (Beaglehole et al., 2011; Bull et al., 2020). Furthermore, while evidence for small to medium effect sizes in adults is readily available, the effectiveness of mobile health interventions for children and adolescents is still limited (Baumann et al., 2022; Böhm et al., 2019; Domin et al., 2021; Mönninghoff et al., 2021; Sporrel et al., 2021). As physical activity and health behavior during youth are important determinants for health behaviors later in life (Baird et al., 2017), the benefit of effective interventions during the early years could magnify the impact. Since the digital natives are already using digital devices on a daily basis (Naszay et al., 2018), barriers to the use of mobile health interventions are probably low in this group. Another important point is to consider if participants have access to areas where they can be physically active (Giles-Corti et al., 2022). While this is something mobile health interventions cannot change directly, it would be important to include this as a covariate for the evaluation to understand barriers to physical activity uptake (Dunton, 2017) and to indicate opportune moments for physical activity (Gonul et al., 2019; Rabbi et al., 2015). A big issue is also the lack of evaluation of the long-term effectiveness of mobile health interventions (Mönninghoff et al., 2021) which would be crucial for sustainable health behavior change. Finally, the evaluation of the variety of parameters linked to intervention effectiveness requires a precise and transparent reporting of all details in studies (Domin et al., 2021; Kwasnicka et al., 2022; Norris et al., 2022). This is especially important for the main outcomes (i.e. physical activity) where we found the selection of the outcome and assessment methodology to have a large impact on total physical activity estimations and a lack of comparability between the assessment methods in our second article (Fiedler et al., 2021). Only then can clear conclusions about the effectiveness for different participant groups and settings be detangled, and progress in the field be accelerated. In a practical sense, it remains important to consider the relevant aspects when recommending electronic or mobile health devices to the broader public. It is crucial to distinguish between commercial claims and scientific evidence.

As our thesis had a particular focus on just-in-time adaptive interventions for physical activity promotion, we will narrow the following chapter down to this topic. Just-in-time adaptive interventions are especially promising as technological advances allow for continuous measurement of parameters of interest to unravel within-person differences over time (Dunton, 2017; Reichert et al., 2020; Trull & Ebner-Priemer, 2013) which can then be potentially used to adapt the interventions to each participant (Conroy et al., 2020; Gonul et al., 2019; Kwasnicka & Naughton, 2020; Nahum-Shani et al., 2015; Nahum-Shani et al., 2018; Tong et al., 2021). Yet, research on just-in-time adaptive interventions as well as continuously assessing outcomes and tailoring variables are still in their infancy (Reichert et al., 2020; Tong et al., 2021).

Question 2

What has to be considered in future just-in-time adaptive interventions for physical activity promotion?

As discussed earlier, the main challenge for a successful implementation of just-in-time adaptive interventions remains the opportune moment identification (Gonul et al., 2019). Here, the choice of sensor input to decide if a moment is a truly opportune moment or not has to be considered with great care and the decision has to rely on reliable, valid, and feasible measurements. Thankfully, there is a lot to learn from previous ambulatory assessment studies as those studies also rely on real-time assessments using e.g. e-diaries and require a good sample strategy to assess ecological valid data (Ebner-Priemer et al., 2013; Reichert et al., 2020; Trull & Ebner-Priemer, 2013). While these assessments can be triggered randomly or during fixed times throughout the day to access e.g. affect (Trull & Ebner-Priemer, 2013), previous studies also used sensor input for specific questions. Here, the triggering of diaries has been adapted by surpassing thresholds for physical activity (Ebner-Priemer et al., 2013), distance covered via Global Positioning System (Tost et al., 2019), thresholds for sedentary behavior (Giurgiu et al., 2020), physically inactive phases (Fiedler et al., 2022), and elevated heart-rate indicating emotional events (Ebner-Priemer et al., 2007). This research is crucial to lay the foundation for just-in-time adaptive interventions by exploring opportune moments for behavior change and the feasibility of sensor input as triggers. Additionally, as e.g., core affect is associated with physical activity on the same day (Do et al., 2021; Fiedler et al., 2022) and especially with the minutes following the ecological momentary assessments (Y. Liao et al., 2015), core affect could be included in just-in-time adaptive interventions by tailoring e.g. the message sent at a decision point based on core affect of the participant. Another possibility would be to aim to increase e.g. valence as a proximal outcome to enhance physical activity in the following time (if the association between valence and physical activity proves to be causal). While there are many additional promising variables like blood sugar (Clavel et al., 2022), oxygen consumption (Düking et al., 2022), and blood pressure (Moon et al., 2020) which can be assessed continuously using smartwatches, smart patches, or smart clothing, many of those still lack the validity, especially under everyday life conditions (Shei et al., 2022). One main reason for the abundance of wearable technology paired with limited quality criteria is that most devices are commercial devices that lack transparency of algorithms and high-quality validation studies (Shei et al., 2022). Therefore, current just-in-time adaptive interventions are still limited in the choice of parameters which will most likely extend within the next years.

It is therefore no surprise that previous just-in-time adaptive intervention studies for physical activity promotion or sedentary behavior reduction focused on parameters that already meet quality criteria like Global Positioning System, accelerometers, time of day, weather, and digital diaries for opportune moment identification (Hardeman et al., 2019). The review of Hardeman and colleagues (2019) found that those studies were mainly feasibility studies with a lack of randomized controlled trials, small sample sizes, within-person perspectives, and transparent and uniform reporting. We addressed that gap with our third article (Wunsch et al., 2022) by providing a framework for future just-in-time adaptive interventions. Additionally, in our fourth article (Fiedler et al., submitted) we evaluated the importance of engagement with a just-in-time adaptive intervention within-persons in a pre-registered study while providing all data and analysis code for transparency. Furthermore, we found core affect to be related to daily physical activity during our intervention study within our fifth article (Fiedler et al., 2022), which confirmed findings from previous ecological momentary assessment studies (Cushing et al., 2017; Do et al., 2021; Dunton et al., 2014; Y. Liao et al., 2017; Niermann et al., 2016; Reichert et al., 2016; Schwerdtfeger et al., 2010). This points to core affect being an important moderator for behavior change and just-in-time adaptive intervention effectiveness in particular which could be targeted by intervention as a proximal variable if the relation proofs to be causal. Further important considerations for just-in-time adaptive interventions and tailored interventions, in general, are to explore the person-specific dose for the interventions (Conroy et al., 2020; Goldstein et al., 2020; Hojjatinia et al., 2021; P. Liao et al., 2018) to avoid overburdening participants. That can be achieved using machine learning processes adapted to personal preferences and sensor input to minimize untimely triggers (Gonul et al., 2019). Rabbi and colleagues (2015) used a sequential decision-making algorithm that delivered automatic suggestions for physical activity and dietary behavior. They report promising results but also important lessons learned. Here, they point out that user input is important to correct for changes in the environment (e.g. by relocating), social circle (e.g. an exercise peer being on vacation), or adding new habits (e.g. starting to go to the gym). As the algorithm can only learn from the past, it is slow to adapt to new circumstances without user input.

To advance the knowledge about such highly individualized interventions and their moderators, within-subject perspectives in longitudinal data analyses are needed to assess the time-varying effects. One common statistical method to examine those effects is general linear mixed models, also called multilevel models or hierarchical linear models (Hoffman, 2015), which we applied in our fourth (Fiedler et al., submitted) and fifth article (Fiedler et al., 2022). To evaluate causality and improve just-in-time adaptive interventions, N-of-1 trials (Kwasnicka & Naughton, 2020; McDonald et al., 2017) and micro randomized trials (Klasnja et al., 2015; Qian et al., 2022) are recommended designs. N-of-1 trials can be used to explore within-person associations and treatment effects where participants are their own control group. These designs are extremely helpful to identify patterns of behavior, within-person differences, and antecedents or consequences of behavior (Dunton, 2017; Kwasnicka & Naughton, 2020; McDonald et al., 2017). In micro randomized trials, each participant is randomly assigned to be triggered or not at decision points many times throughout the trial and within-person differences can then be estimated to evaluate the effectiveness of the trigger (Klasnja et al., 2015; Qian et al., 2022). This allows not only to evaluate the effect on a proximal target (e.g. being active in the minutes after the trigger) but also the effect as a moderator for a distal target (e.g. physical activity behavior change in a pre-post design) and time lagged effects (e.g. following a recommendation not immediately but remembering it at the next occasion), and is especially useful for just-in-time adaptive interventions (Klasnja et al., 2015; Qian et al., 2022). Additionally, qualitative research evaluating stakeholder and participant feedback and including participants during the design of such interventions using intervention mapping approaches can have great benefits (Direito et al., 2018; Fernandez et al., 2019). By doing so, barriers for physical activity uptake in the participant group of choice can be detected a priori, their needs be addressed, and appropriate theoretical foundation and behavior change techniques selected for the most promising intervention design (Direito et al., 2018).

After discussing the variety of requirements to successfully design and evaluate mobile health interventions, just-in-time adaptive interventions for physical activity, and related topics, it is a natural conclusion that interdisciplinary research throughout (but not limited to) the areas of psychology, engineering, sports science, statistics, and informatics are needed to enhance and evaluate sensors, develop behavior change theories to guide interventions, provide timely feedback, consider different aspects of physical activity or inactivity measures and pool all of it together into sophisticated yet feasible and accepted applications (Molina Recio et al., 2016; Nahum-Shani et al., 2018; Nilsen et al., 2012). While this sounds reasonable in theory, interdisciplinarity also has its challenges and requires thorough organization and communication to climb the ladder from creating a stable application to achieving a clinically effective intervention (Blandford et al., 2018).

Question 3

Which participants could particularly profit from future just-in-time adaptive interventions?

Future just-in-time adaptive interventions are promising to promote physical activity in groups of patients who have additional benefits from the highly individualized approach like knee osteoarthritis (Esser & Bailey, 2011). Osteoarthritis in the knee joint has a high prevalence and severe risk for disability and comorbidities like cardiovascular diseases in today's society (Cui et al., 2020; Palazzo et al., 2016). Here, physical activity has been found to be beneficial for pain reduction and improvement of physical function while patients have a high prevalence of physical inactivity, and exercise interventions are an effective but underused tool for health improvements (Bosomworth, Neil, J., 2009; Cronström et al., 2019; Esser & Bailey, 2011; Gay et al., 2016; Kraus et al., 2019). In this special case, it is important to consider an optimal amount and type of physical activity (e.g. avoiding jumps) to reduce the pain, and disability and enhance the physical function, and quality of life of the patients (Kraus et al., 2019). Previous digital health studies on this topic showed promising results and good acceptance of the intervention by participants (Berry et al., 2018; Bossen et al., 2013; Nero et al., 2017; Safari et al., 2020). Here, just-in-time adaptive interventions could advance web-based approaches by adapting the intervention according to sensor input which indicates the daily load on the knee joint (Inan et al., 2018; van der Straaten et al., 2018) and e.g. the assessment of pain (Stone et al., 2021). This could provide the participant with the optimal daily activity dose which includes personal experiences (i.e. pain) and recommend feasible physical activities. By receiving highly individualized feedback and recommendations for physical activity uptake, participants' fears to engage in too much or inappropriate exercises could be reduced while keeping them active and improving their health.

To conclude, physical activity promotion in healthy participants remains a tedious but essential topic to enhance health behavior throughout the lifespan and support the prevention of noncommunicable diseases. Here, digital behavior change interventions are especially promising due to their high acceptance and availability of devices in the population (chapter 1). Our work included in this thesis contributes to the development of effective mobile health interventions in many ways. We highlighted key facets for effective interventions and indicate understudied or problematic topics in digital health promotion (chapter 2). Based on these shortcomings, we contributed to the understanding of discrepancies between different self-reported and device-based physical activity measures (chapter 3). As we found just-in-time adaptive interventions to be understudied in healthy participants, we combined previous frameworks and highlighted challenges and opportunities for these highly individualized interventions (chapter 4). In the next step, we advanced previous knowledge about the importance of engagement with just-in-time adaptive interventions and opportune moment identification (chapter 5), and the association of core affect with physical activity during interventions (chapter 6). Finally, we refined and expanded our findings to provide important considerations for future mobile health interventions for physical activity promotion in general, and just-in-time adaptive interventions in particular, including a knowledge transfer to the promising field of just-in-time adaptive interventions for patients with knee osteoarthritis (chapter 7). Overall, we are convinced that the field of digital interventions can develop into an important addition to conventional intervention methods if previous shortcomings highlighted by this thesis are addressed in future research, and different branches of science cooperate effectively.

References

Baird, J., Jacob, C., Barker, M., Fall, C. H. D., Hanson, M., Harvey, N. C., Inskip, H. M., Kumaran, K., & Cooper, C. (2017). Developmental Origins of Health and Disease: A Lifecourse Approach to the Prevention of Non-Communicable Diseases. *Healthcare (Basel, Switzerland)*, *5*(1). https://doi.org/10.3390/healthcare5010014

Baumann, H., Fiedler, J., Wunsch, K., Woll, A., & Wollesen, B. (2022). mHealth Interventions to Reduce Physical Inactivity and Sedentary Behavior in Children and Adolescents: Systematic Review and Meta-analysis of Randomized Controlled Trials. *JMIR mHealth and uHealth*, *10*(5), e35920. https://doi.org/10.2196/35920

Beaglehole, R., Bonita, R., Horton, R., Adams, C., Alleyne, G., Asaria, P., Baugh, V., Bekedam, H., Billo, N., Casswell, S., Cecchini, M., Colagiuri, R., Colagiuri, S., Collins, T., Ebrahim, S., Engelgau, M., Galea, G., Gaziano, T., Geneau, R., . . . Watt, J. (2011). Priority actions for the non-communicable disease crisis. *The Lancet*, *377*(9775), 1438–1447. https://doi.org/10.1016/S0140-6736(11)60393-0

Berry, A., McCabe, C. S., Muir, S., & Walsh, N. (2018). Digital behaviour change interventions to facilitate physical activity in osteoarthritis: a systematic review. *Physical Therapy Reviews*, *23*(3), 197–206. https://doi.org/10.1080/10833196.2018.1470747

Blandford, A., Gibbs, J., Newhouse, N., Perski, O., Singh, A., & Murray, E. (2018). Seven lessons for interdisciplinary research on interactive digital health interventions. *Digital health*, *4*, 2055207618770325. https://doi.org/10.1177/2055207618770325

Böhm, B., Karwiese, S. D., Böhm, H., & Oberhoffer, R. (2019). Effects of mobile health including wearable activity trackers to increase physical activity outcomes among healthy children and adolescents: Systematic review. *JMIR mHealth and uHealth*, *7*(4), e8298. https://doi.org/10.2196/mhealth.8298

Bosomworth, Neil, J. (2009). Exercise and knee osteoarthritis: benefit or hazard? *Canadian Family Physician*, *55*(9), 871–878.

Bossen, D., Veenhof, C., van Beek, K. E., Spreeuwenberg, P. M., Dekker, J., & Bakker, D. H. de (2013). Effectiveness of a web-based physical activity intervention in patients with knee and/or hip osteoarthritis: randomized controlled trial. *Journal of medical Internet research*, *15*(11), e257. https://doi.org/10.2196/jmir.2662

Buckingham, S. A., Williams, A. J., Morrissey, K., Price, L., & Harrison, J. (2019). Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: A systematic review. *Digital health*, *5*, 1-50. https://doi.org/10.1177/2055207619839883

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P, Cardon, G., Carty, C., Chaput, J.-P., Chastin, S., Chou, R., Dempsey, P. C., DiPietro, L., Ekelund, U., Firth, J., Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., . . . Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British journal of sports medicine*, *54*(24), 1451–1462. https://doi.org/10.1136/bjsports-2020-102955

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

Clavel, P., Tiollier, E., Leduc, C., Fabre, M., Lacome, M., & Buchheit, M. (2022). Concurrent Validity of a Continuous Glucose-Monitoring System at Rest and During and Following a High-Intensity Interval Training Session. *International Journal of Sports Physiology and Performance*, *17*(4), 627–633. https://doi.org/10.1123/ijspp.2021-0222

Conroy, D. E., Lagoa, C. M., Hekler, E. B., & Rivera, D. E (2020). Engineering Person-Specific Behavioral Interventions to Promote Physical Activity. *Exercise and sport sciences reviews*, *48*(4), 170–179. https://doi.org/10.1249/JES.0000000000000232

Cronström, A., Dahlberg, L. E., Nero, H., Ericson, J., & Hammarlund, C. S. (2019). 'I would never have done it if it hadn't been digital': a qualitative study on patients' experiences of a digital management programme for hip and knee osteoarthritis in Sweden. *BMJ open*, *9*(5).

Cui, A., Li, H., Wang, D., Zhong, J., Chen, Y., & Lu, H. (2020). Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies. *EClinicalMedicine*, *29-30*, 100587. https://doi.org/10.1016/j.eclinm.2020.100587

Cushing, C. C., Mitchell, T. B., Bejarano, C. M., Walters, R. W., Crick, C. J., & Noser, A. E. (2017). Bidirectional Associations Between Psychological States and Physical Activity in Adolescents: A mHealth Pilot Study. *Journal of pediatric psychology*, *42*(5), 559–568. https://doi.org/10.1093/jpepsy/jsw099

Direito, A., Walsh, D., Hinbarji, M., Albatal, R., Tooley, M., Whittaker, R., & Maddison, R. (2018). Using the Intervention Mapping and Behavioral Intervention Technology Frameworks: Development of an mHealth Intervention for Physical Activity and Sedentary Behavior Change. *Health education & behavior: the official publication of the Society for Public Health Education*, *45*(3), 331–348. https://doi.org/10.1177/1090198117742438

Do, B., Wang, S. D., Courtney, J. B., & Dunton, G. F. (2021). Examining the day-level impact of physical activity on affect during the early months of the COVID-19 pandemic: An ecological momentary assessment study. *Psychology of Sport and Exercise*, *56.* https://doi.org/10.1016/j.psychsport.2021.102010

Domin, A., Spruijt-Metz, D., Theisen, D., Ouzzahra, Y., & Vögele, C. (2021). Smartphone-Based Interventions for Physical Activity Promotion: Scoping Review of the Evidence Over the Last 10 Years. *JMIR mHealth and uHealth*, *9*(7), e24308. https://doi.org/10.2196/24308

Düking, P., van Hooren, B., & Sperlich, B. (2022). Assessment of Peak Oxygen Uptake with a Smartwatch and its Usefulness for Training of Runners. *International journal of sports medicine.* Vorab-Onlinepublikation. https://doi.org/10.1055/a-1686-9068

Dunton, G. F. (2017). Ecological Momentary Assessment in Physical Activity Research. *Exercise and sport sciences reviews*, *45*(1), 48–54. https://doi.org/10.1249/JES.0000000000000092

Dunton, G. F., Huh, J., Leventhal, A. M., Riggs, N., Hedeker, D., Spruijt-Metz, D., & Pentz, M. A. (2014). Momentary assessment of affect, physical feeling states, and physical activity in children. *Health psychology: official journal of the Division of Health Psychology, American Psychological Association*, *33*(3), 255–263. https://doi.org/10.1037/a0032640

Ebner-Priemer, U. W., Koudela, S., Mutz, G., & Kanning, M. (2013). Interactive Multimodal Ambulatory Monitoring to Investigate the Association between Physical Activity and Affect. *Frontiers in psychology*, *3*, 596. https://doi.org/10.3389/fpsyg.2012.00596

Ebner-Priemer, U. W., Welch, S. S., Grossman, P., Reisch, T., Linehan, M. M., & Bohus, M. (2007). Psychophysiological ambulatory assessment of affective dysregulation in borderline personality disorder. *Psychiatry research*, *150*(3), 265–275. https://doi.org/10.1016/j.psychres.2006.04.014

Esser, S., & Bailey, A. (2011). Effects of exercise and physical activity on knee osteoarthritis. *Current pain and headache reports*, *15*(6), 423–430. https://doi.org/10.1007/s11916-011-0225-z

Fernandez, M. E., Ruiter, R. A. C., Markham, C. M., & Kok, G. (2019). Intervention Mapping: Theory- and Evidence-Based Health Promotion Program Planning: Perspective and Examples. *Frontiers in public health*, *7*, 209. https://doi.org/10.3389/fpubh.2019.00209

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 (Basel, Switzerland)*, *21*(8). https://doi.org/10.3390/s21082672

Fiedler, J., Eckert, T., Wunsch, K., & Woll, A. (2020). Key facets to build up eHealth and mHealth interventions to enhance physical activity, sedentary behavior and nutrition in healthy subjects - an umbrella review. *BMC public health*, *20*(1), 1605.

Fiedler, J., Seifert, C., Eckert, T., Woll, A., & Wunsch, K. (submitted). A just-in-time adaptive intervention to enhance physical activity in the SMARTFAMILY2.0 trial. *Sport, Exercise, and Performance Psychology*.

Fiedler, J., Seiferth, C., Eckert, T., Woll, A., & Wunsch, K. (2022). Sleep quality, valence, energetic arousal, and calmness as predictors of device-based measured physical activity during a three-week mHealth intervention. *German Journal of Exercise and Sport Research.* Vorab-Onlinepublikation. https://doi.org/10.1007/s12662-022-00809-y

Gay, C., Chabaud, A., Guilley, E., & Coudeyre, E. (2016). Educating patients about the benefits of physical activity and exercise for their hip and knee osteoarthritis. Systematic literature review. *Annals of physical and rehabilitation medicine*, *59*(3), 174–183. https://doi.org/10.1016/j.rehab.2016.02.005

Giavarina, D. (2015). Understanding Bland Altman analysis. *Biochemia medica*, *25*(2), 141–151. https://doi.org/10.11613/BM.2015.015

Giles-Corti, B., Moudon, A. V., Lowe, M., Cerin, E., Boeing, G., Frumkin, H., Salvo, D., Foster, S., Kleeman, A., Bekessy, S., Sá, T. H. de, Nieuwenhuijsen, M., Higgs, C., Hinckson, E., Adlakha, D., Arundel, J., Liu, S., Oyeyemi, A. L., Nitvimol, K., & Sallis, J. F. (2022). What next? Expanding our view of city planning and global health, and implementing and monitoring evidence-informed policy. *The Lancet Global Health*, *10*(6), e919-e926. https://doi.org/10.1016/S2214-109X(22)00066-3

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

Glanz, K., Rimer, B. K., & Viswanath, K. (Hrsg.). (2008). *Health behavior and health education: theory, research, and practice*. John Wiley & Sons.

Goldstein, S. P., Thomas, J. G., Foster, G. D., Turner-McGrievy, G., Butryn, M. L., Herbert, J. D., Martin, G. J., & Forman, E. M. (2020). Refining an algorithm-powered just-in-time adaptive weight control intervention: A randomized controlled trial evaluating model performance and behavioral outcomes. *Health informatics journal*, 1460458220902330. https://doi.org/10.1177/1460458220902330

Gonul, S., Namli, T., Huisman, S., Laleci Erturkmen, G. B., Toroslu, I. H., & Cosar, A. (2019). An expandable approach for design and personalization of digital, just-in-time adaptive interventions. *Journal of the American Medical Informatics Association: JAMIA*, *26*(3), 198–210. https://doi.org/10.1093/jamia/ocy160

Hardeman, W., Houghton, J., Lane, K., Jones, A., & Naughton, F. (2019). A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. *The international journal of behavioral nutrition and physical activity*, *16*(1), 31. https://doi.org/10.1186/s12966-019-0792-7

Heron, K. E., & Smyth, J. M. (2010). Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. *British journal of health psychology*, *15*(Pt 1), 1–39. https://doi.org/10.1348/135910709X466063

Hoffman, L. (2015). *Longitudinal analysis: Modeling within-person fluctuation and change*. Routledge.

Hojjatinia, S [Sarah], Hojjatinia, S [Sahar], Lagoa, C. M., Brunke-Reese, D., & Conroy, D. E. (2021). Person-specific dose-finding for a digital messaging intervention to promote physical activity. *Health psychology: official journal of the Division of Health Psychology, American Psychological Association*, *40*(8), 502–512. https://doi.org/10.1037/hea0001117

Inan, O. T., Whittingslow, D. C., Teague, C. N., Hersek, S., Pouyan, M. B., Millard-Stafford, M., Kogler, G. F., & Sawka, M. N. (2018). Wearable knee health system employing novel physiological biomarkers. *Journal of applied physiology (Bethesda, Md. : 1985)*, *124*(3), 537–547. https://doi.org/10.1152/japplphysiol.00366.2017

Keadle, S. K., Lyden, K. A., Strath, S. J., Staudenmayer, J. W., & Freedson, P. S. (2019). A Framework to Evaluate Devices That Assess Physical Behavior. *Exercise and sport sciences reviews*, *47*(4), 206–214. https://doi.org/10.1249/JES.0000000000000206

Klasnja, P., Hekler, E. B., Shiffman, S., Boruvka, A., Almirall, D., Tewari, A., & Murphy, S. A. (2015). Microrandomized trials: An experimental design for developing just-in-time adaptive interventions. *Health psychology: official journal of the Division of Health Psychology, American Psychological Association*, *34S*, 1220–1228. https://doi.org/10.1037/hea0000305

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of chiropractic medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kraus, V. B., Sprow, K., Powell, K. E., Buchner, D., Bloodgood, B., Piercy, K., George, S. M., & Kraus, W. E. (2019). Effects of Physical Activity in Knee and Hip Osteoarthritis: A Systematic Umbrella Review. *Medicine and science in sports and exercise*, *51*(6), 1324–1339. https://doi.org/10.1249/MSS.0000000000001944

Kwasnicka, D., Keller, J., Perski, O., Potthoff, S., Hoor, G. A. ten, Ainsworth, B., Crutzen, R., Dohle, S., van Dongen, A., Heino, M., Henrich, J. F., Knox, L., König, L. M., Maltinsky, W., McCallum, C., Nalukwago, J., Neter, E., Nurmi, J., Spitschan, M., . . . Sanderman, R. (2022). White

Paper: Open Digital Health - accelerating transparent and scalable health promotion and treatment. *Health psychology review*, 1–17. https://doi.org/10.1080/17437199.2022.2046482

Kwasnicka, D., & Naughton, F. (2020). N-of-1 methods: A practical guide to exploring trajectories of behaviour change and designing precision behaviour change interventions. *Psychology of Sport and Exercise*, *47*, 101570. https://doi.org/10.1016/j.psychsport.2019.101570

Lago-Peñas, S., Rivera, B., Cantarero, D., Casal, B., Pascual, M., Blázquez-Fernández, C., & Reyes, F. (2021). The impact of socioeconomic position on non-communicable diseases: what do we know about it? *Perspectives in public health*, *141*(3), 158–176. https://doi.org/10.1177/1757913920914952

Liao, P., Dempsey, W., Sarker, H., Hossain, S. M., Al'absi, M., Klasnja, P., & Murphy, S. A. (2018). Just-in-Time but Not Too Much: Determining Treatment Timing in Mobile Health. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, *2*(4). https://doi.org/10.1145/3287057.

Liao, Y., Chou, C.-P., Huh, J., Leventhal, A., & Dunton, G. F. (2017). Examining acute bi-directional relationships between affect, physical feeling states, and physical activity in free-living situations using electronic ecological momentary assessment. *Journal of behavioral medicine*, *40*(3), 445–457. https://doi.org/10.1007/s10865-016-9808-9

Liao, Y., Shonkoff, E. T., & Dunton, G. F. (2015). The Acute Relationships Between Affect, Physical Feeling States, and Physical Activity in Daily Life: A Review of Current Evidence. *Frontiers in psychology*, *6*, 1975. https://doi.org/10.3389/fpsyg.2015.01975

Liu, J., Tang, W., Chen, G., Lu, Y., Feng, C., & Tu, X. M. (2016). Correlation and agreement: overview and clarification of competing concepts and measures. *Shanghai archives of psychiatry*, *28*(2), 115–120. https://doi.org/10.11919/j.issn.1002-0829.216045

Martín, C. A., Rivera, D. E, Riley, W. T, Hekler, E. B., Buman, M. P, Adams, M. A, & King, A. C (2014). A dynamical systems model of social cognitive theory. *American Control Conference*, 2407–2412.

McDonald, S., Quinn, F., Vieira, R., O'Brien, N., White, M., Johnston, D. W., & Sniehotta, F. F. (2017). The state of the art and future opportunities for using longitudinal n-of-1 methods in health behaviour research: a systematic literature overview. *Health psychology review*, *11*(4), 307–323. https://doi.org/10.1080/17437199.2017.1316672

Molina Recio, G., García-Hernández, L., Molina, Luque, Rafael, & Salas-Morera, L. (2016). The role of interdisciplinary research team in the impact of health apps in health and computer science publications: a systematic review. *BioMedical Engineering OnLine*, *15 Suppl 1*(1), 77. https://doi.org/10.1186/s12938-016-0185-y

Mönninghoff, A., Kramer, J. N., Hess, A. J., Ismailova, K., Teepe, G. W., Tudor Car, L., Müller-Riemenschneider, F., & Kowatsch, T. (2021). Long-term Effectiveness of mHealth Physical Activity Interventions: Systematic Review and Meta-analysis of Randomized Controlled Trials. *Journal of medical Internet research*, *23*(4), e26699. https://doi.org/10.2196/26699

Moon, J. H., Kang, M.-K., Choi, C.-E., Min, J., Lee, H.-Y., & Lim, S. (2020). Validation of a wearable cuff-less wristwatch-type blood pressure monitoring device. *Scientific reports*, *10*(1), 19015. https://doi.org/10.1038/s41598-020-75892-y

Morrison, L. G., Yardley, L., Powell, J., & Michie, S. (2012). What design features are used in effective e-health interventions? A review using techniques from Critical Interpretive Synthesis. *Telemedicine journal and e-health: the official journal of the American Telemedicine Association*, *18*(2), 137–144. https://doi.org/10.1089/tmj.2011.0062

Nahum-Shani, I., Hekler, E. B., & Spruijt-Metz, D. (2015). Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. *Health Psychology*, *34S*, 1209–1219. https://doi.org/10.1037/hea0000306

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Annals of Behavioral Medicine*, *52*(6), 446–462. https://doi.org/10.1007/s12160-016-9830-8

Naszay, M., Stockinger, A., Jungwirth, D., & Haluza, D. (2018). Digital age and the Public eHealth perspective: Prevailing health app use among Austrian Internet users. *Informatics for health & social care*, *43*(4), 390–400. https://doi.org/10.1080/17538157.2017.1399131

Navarro-Barrientos, J.-E., Rivera, D. E, & Collins, L. M. (2011). A dynamical model for describing behavioural interventions for weight loss and body composition change. *Mathematical and computer modelling of dynamical systems*, *17*(2), 183–203. https://doi.org/10.1080/13873954.2010.520409

Nero, H., Dahlberg, J., & Dahlberg, L. E. (2017). A 6-Week Web-Based Osteoarthritis Treatment Program: Observational Quasi-Experimental Study. *Journal of medical Internet research*, *19*(12), e422. https://doi.org/10.2196/jmir.9255

Niermann, C., Herrmann, C., Haaren, B. von, van Kann, D., & Woll, A. (2016). Affect and Subsequent Physical Activity: An Ambulatory Assessment Study Examining the Affect-Activity Association in a Real-Life Context. *Frontiers in psychology*, *7*, 677. https://doi.org/10.3389/fpsyg.2016.00677

Nilsen, W., Kumar, S., Shar, A., Varoquiers, C., Wiley, T., Riley, W. T, Pavel, M., & Atienza, A. A. (2012). Advancing the science of mHealth. *Journal of Health Communication*, *17 Suppl 1*(sup1), 5–10. https://doi.org/10.1080/10810730.2012.677394

Norris, E., Sulevani, I., Finnerty, A. N., & Castro, O. (2022). Assessing Open Science practices in physical activity behaviour change intervention evaluations. *BMJ Open Sport & Exercise Medicine*, *8*(2), e001282. https://doi.org/10.1136/bmjsem-2021-001282

Palazzo, C., Nguyen, C., Lefevre-Colau, M.-M., Rannou, F., & Poiraudeau, S. (2016). Risk factors and burden of osteoarthritis. *Annals of physical and rehabilitation medicine*, *59*(3), 134–138. https://doi.org/10.1016/j.rehab.2016.01.006

Prestwich, A., Sniehotta, F. F., Whittington, C., Dombrowski, S. U., Rogers, L., & Michie, S. (2014). Does theory influence the effectiveness of health behavior interventions? Meta-analysis. *Health psychology: official journal of the Division of Health Psychology, American Psychological Association*, *33*(5), 465–474. https://doi.org/10.1037/a0032853

Qian, T., Walton, A. E., Collins, L. M., Klasnja, P., Lanza, S,T, Nahum-Shani, I., Rabbi, M., Russell, M,A, Walton, M,A, Yoo, H., & Murphy S, A. (2022). The microrandomized trial for developing digital interventions: Experimental design and data analysis considerations. *Psychological Methods.* Vorab-Onlinepublikation. https://doi.org/10.1037/met0000283

Rabbi, M., Aung, M. H., Zhang, M., & Choudhury, T. (2015). MyBehavior. In K. Mase, M. Langheinrich, D. Gatica-Perez, H. Gellersen, T. Choudhury & K. Yatani (Hrsg.), *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15* (S. 707-718). ACM Press. https://doi.org/10.1145/2750858.2805840

Reichert, M., Giurgiu, M., Koch, E. D., 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*, 101742. https://doi.org/10.1016/j.psychsport.2020.101742

Reichert, M., Tost, H., Reinhard, I., Zipf, A., Salize, H.-J., Meyer-Lindenberg, A., & Ebner-Priemer, U. W. (2016). Within-Subject Associations between Mood Dimensions and Non-exercise Activity: An Ambulatory Assessment Approach Using Repeated Real-Time and Objective Data. *Frontiers in psychology*, *7*, 918. https://doi.org/10.3389/fpsyg.2016.00918

Riley, W. T, Martin, C. A., Rivera, D. E, Hekler, E. B., Adams, M. A, Buman, M. P, Pavel, M., & King, A. C (2016). Development of a dynamic computational model of social cognitive theory. *Translational behavioral medicine*, *6*(4), 483–495. https://doi.org/10.1007/s13142-015-0356-6

Riley, W. T, Rivera, D. E, Atienza, A. A., Nilsen, W., Allison, S. M., & Mermelstein, R. (2011). Health behavior models in the age of mobile interventions: are our theories up to the task? *Translational behavioral medicine*, *1*(1), 53–71. https://doi.org/10.1007/s13142-011-0021-7

Safari, R., Jackson, J., & Sheffield, D. (2020). Digital Self-Management Interventions for People With Osteoarthritis: Systematic Review With Meta-Analysis. *Journal of medical Internet research*, *22*(7), e15365. https://doi.org/10.2196/15365

Schembre, S. M., Liao, Y., Robertson, M. C., Dunton, G. F., Kerr, J., Haffey, M. E., Burnett, T., Basen-Engquist, K., & Hicklen, R. S. (2018). Just-in-Time Feedback in Diet and Physical Activity Interventions: Systematic Review and Practical Design Framework. *Journal of medical Internet research*, *20*(3), e106. https://doi.org/10.2196/jmir.8701

Schwerdtfeger, A., Eberhardt, R., Chmitorz, A., & Schaller, E. (2010). Momentary Affect Predicts Bodily Movement in Daily Life: An Ambulatory Monitoring Study. *Journal of Sport and Exercise Psychology*, *32*(5), 674–693.

Shei, R.-J., Holder, I. G., Oumsang, A. S., Paris, B. A., & Paris, H. L. (2022). Wearable activity trackers-advanced technology or advanced marketing? *European journal of applied physiology*. Vorab-Onlinepublikation. https://doi.org/10.1007/s00421-022-04951-1

Sporrel, K., Nibbeling, N., Wang, S., Ettema, D., & Simons, M. (2021). Unraveling Mobile Health Exercise Interventions for Adults: Scoping Review on the Implementations and Designs of Persuasive Strategies. *JMIR mHealth and uHealth*, *9*(1), e16282. https://doi.org/10.2196/16282

Stone, A. A., Obbarius, A., Junghaenel, D. U., Wen, C. K. F., & Schneider, S. (2021). High-resolution, field approaches for assessing pain: Ecological Momentary Assessment. *Pain*, *162*(1), 4–9. https://doi.org/10.1097/j.pain.0000000000002049

Tong, H. L., & Laranjo, L. (2018). The use of social features in mobile health interventions to promote physical activity: a systematic review. *NPJ digital medicine*, *1*, 43. https://doi.org/10.1038/s41746-018-0051-3

Tong, H. L., Quiroz, J. C., Kocaballi, A. B., Fat, S. C. M., Dao, K. P., Gehringer, H., Chow, C. K., & Laranjo, L. (2021). Personalized mobile technologies for lifestyle behavior change: A systematic review, meta-analysis, and meta-regression. *Preventive medicine*, *148*, 106532. https://doi.org/10.1016/j.ypmed.2021.106532

Tost, H., Reichert, M., Braun, U., Reinhard, I., Peters, R., Lautenbach, S., Hoell, A., Schwarz, E., Ebner-Priemer, U. W., 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

Trull, T. J., & Ebner-Priemer, U. W. (2013). Ambulatory assessment. *Annual Review of Clinical Psychology*, *9*, 151–176. https://doi.org/10.1146/annurev-clinpsy050212-185510

Umberson, D., Crosnoe, R., & Reczek, C. (2010). Social Relationships and Health Behavior Across Life Course. *Annual review of sociology*, *36*, 139–157. https://doi.org/10.1146/annurev-soc-070308-120011

van der Straaten, R., Baets, L. de, Jonkers, I., & Timmermans, A. (2018). Mobile assessment of the lower limb kinematics in healthy persons and in persons with degenerative knee disorders: A systematic review. *Gait & Posture*, *59*, 229–241. https://doi.org/10.1016/j.gaitpost.2017.10.005

Viner, R. M., Ozer, E. M., Denny, S., Marmot, M., Resnick, M., Fatusi, A., & Currie, C. (2012). Adolescence and the social determinants of health. *The Lancet*, *379*(9826), 1641–1652. https://doi.org/10.1016/S0140-6736(12)60149-4

Webb, T. L., Joseph, J., Yardley, L., & Michie, S. (2010). Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *Journal of medical Internet research*, *12*(1), e4. https://doi.org/10.2196/jmir.1376

Western, M. J., Armstrong, M. E. G., Islam, I., Morgan, K., Jones, U. F., & Kelson, M. J. (2021). The effectiveness of digital interventions for increasing physical activity in individuals of low socioeconomic status: a systematic review and meta-analysis. *The international journal of behavioral nutrition and physical activity*, *18*(1), 148. https://doi.org/10.1186/s12966-021-01218-4

Wunsch, K., Eckert, T., Fiedler, J., Cleven, L., Niermann, C., Reiterer, H., Renner, B., & Woll, A. (2020). Effects of a Collective Family-Based Mobile Health Intervention Called "SMARTFAMILY" on Promoting Physical Activity and Healthy Eating: Protocol for a Randomized Controlled Trial. *JMIR Research Protocols*, *9*(11), e20534. https://doi.org/10.2196/20534

Wunsch, K., Eckert, T., Fiedler, J., & Woll, A. (2022). Just-in-time adaptive interventions in mobile physical activity interventions - A synthesis of frameworks and future directions. *The European Health Psychologist*, *22*(4), 834–842.