DIPFdocs
Institutionelles Open Access Repositorium

DIPF
Leibniz-Institut für Bildungsforschung und Bildungsinformation

Reichert, Markus; Giurgiu, Marco; Koch, Elena; ...

Ambulatory assessment for physical activity research. State of the science, best practices and future directions

formal und inhaltlich überarbeitete Version der Originalveröffentlichung in:
formally and content revised edition of the original source in:
Psychology of sport and exercise 50 (2020) 101742, 49 S., 10.1016/j.psychsport.2020.101742

Bitte verwenden Sie in der Quellenangabe folgende URN oder DOI /
Please use the following URN or DOI for reference:
urn:nbn:de:0111-dipfdocs-228896
10.25657/02:22889

https://nbn-resolving.org/urn:nbn:de:0111-dipfdocs-228896
https://doi.org/10.25657/02:22889

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: http://creativecommons.org/licenses/by-nc-nd/4.0/deed.de - Sie dürfen das Werk bzw. den Inhalt unter folgenden Bedingungen vervielfältigen, verbreiten und öffentlich zugänglich machen: Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen. Dieses Werk bzw. dieser Inhalt darf nicht für kommerzielle Zwecke verwendet werden und es darf nicht bearbeitet, abgewandelt oder in anderer Weise verändert werden.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-License: http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en - You may copy, distribute and transmit, adapt or exhibit the work in the public as long as you attribute the work in the manner specified by the author or licensor. You are not allowed to make commercial use of the work or its contents. You are not allowed to alter, transform, or change this work in any other way.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

DIPF | Leibniz-Institut für
Bildungsforschung und Bildungsinformation
Frankfurter Forschungsbibliothek
publikationen@dipf.de
www.dipfdocs.de

Mitglied der Leibniz-Gemeinschaft

Ambulatory Assessment for Physical Activity Research: State of the Science, Best Practices and Future Directions

Highlights

Ambulatory Assessment enables to

- study physical activity's correlates (e.g., behavioral, biological) in REAL-LIFE
- assess data near REAL-TIME, WITHIN-PERSONS across multiple assessments
- minimize retrospective biases for ECOLOGICAL VALID FINDINGS
- unravel within-subject ANTECEDENTS, CONCOMITANTS, and CONSEQUENCES of PA
- design and deliver impactful INTERVENTIONS in REAL-WORLD contexts

Abstract

Technological and digital progress benefits physical activity (PA) research. Here we compiled expert knowledge on how Ambulatory Assessment (AA) is utilized to advance PA research, i.e., we present results of the 2nd International CAPA Workshop 2019 “Physical Activity Assessment – State of the Science, Best Practices, Future Directions” where invited researchers with experience in PA assessment, evaluation, technology and application participated. First, we provide readers with the state of the AA science, then we give best practice recommendations on how to measure PA via AA and shed light on methodological frontiers, and we furthermore discuss future directions. AA encompasses a class of methods that allows the study of PA and its behavioral, biological and physiological correlates as they unfold in everyday life. AA includes monitoring of movement (e.g., via accelerometry), physiological function (e.g., via mobile electrocardiogram), contextual information (e.g., via geolocation-tracking), and ecological momentary assessment (EMA; e.g., electronic diaries) to capture self-reported information. The strengths of AA are data assessment that near real-time, which minimizes retrospective biases in real-world settings, consequentially enabling ecological valid findings. Importantly, AA enables multiple assessments across time within subjects resulting in intensive longitudinal data (ILD), which allows unraveling within-person determinants of PA in everyday life. In this paper, we show how AA methods such as triggered e-diaries and geolocation-tracking can be used to measure PA and its correlates, and furthermore how these findings may translate into real-life interventions. In sum, AA provides numerous possibilities for PA research, especially the opportunity to tackle within-subject antecedents, concomitants, and consequences of PA as they unfold in everyday life. In-depth insights on determinants of PA could help us design and deliver impactful interventions in real-world contexts, thus enabling us to solve critical health issues in the 21st century such as insufficient PA and high levels of sedentary behavior.

Keywords: ambulatory assessment, physical activity, ecological momentary assessment, experience sampling, best practices, future directions

Ambulatory Assessment for Physical Activity Research: State of the Science, Best Practices and Future Directions

Markus Reichert[1,2], Marco Giurgiu[1,2], Elena Koch[1], Lena M. Wieland[1], Sven Lautenbach[3], Andreas B. Neubauer[4,5], Birte von Haaren-Mack[6], Renè Schilling[7], Irina Timm[1], Nanna Notthoff[8], Isabel Marzi[9], Holger Hill[1], Sarah Brüßler[1], Tobias Eckert[11], Janis Fiedler[10], Alexander Burchartz[10], Bastian Anedda[10], Kathrin Wunsch[10], Markus Gerber[7], Darko Jekauc[10], Alexander Woll[10], Genevieve F. Dunton[11], Martina Kanning[12], Claudio R. Nigg[10], Ulrich Ebner-Priemer[1], and Yue Liao[13]

[1] Mental mHealth Lab, Institute of Sports and Sports Science, Karlsruhe Institute of Technology (KIT); postal address: Engler-Bunte-Ring 15, 76131 Karlsruhe, Baden-Wuerttemberg, Germany

[2] Department of Psychiatry and Psychotherapy, Central Institute of Mental Health (CIMH), Medical Faculty Mannheim, Heidelberg University; postal address: Central Institute of Mental Health (CIMH), Square J5, 68159 Mannheim, Baden-Wuerttemberg, Germany

[3] Heidelberg Institute for Geoinformation Technology (HeiGIT) at Heidelberg University; postal address: Heidelberg Institute for Geoinformation Technology, Schloss-Wolfsbrunnenweg 33, 69118 Heidelberg, Baden-Wuerttemberg, Germany

[4] Department of Education and Human Development, DIPF | Leibniz Institute for Research and Information in Education; postal address: Rostocker Straße 6, 60323 Frankfurt am Main, Hessen, Germany

[5] Center for Research on Individual Development and Adaptive Education of Children at Risk (IDeA); postal address: Rostocker Straße 6, 60323 Frankfurt am Main, Hessen, Germany

[6] Department of Health and Social Psychology, Institute of Psychology, German Sport University Cologne; postal address: Am Sportpark Muengersdorf 6, 50933 Cologne, North Rhine-Westphalia, Germany

[7] Department of Sport, Exercise and Health, University of Basel, Postal address: Birsstrasse 320 B, 4052 Basel, Switzerland

[8] Institute of Sport and Exercise Psychology, Department of Sports Science, Leipzig University; postal address: Jahnallee 59, 04109 Leipzig, Saxony, Germany

[9] Department of Sports Science and Sport, Friedrich-Alexander-University Erlangen-Nürnberg (FAU); postal address: Gebbertstrasse 123b, 91058 Erlangen, Bavaria, Germany

[10] Institute of Sports and Sports Science, Karlsruhe Institute of Technology (KIT); postal address: Engler-Bunte-Ring 15, 76131 Karlsruhe, Baden-Wuerttemberg, Germany

[11] Departments of Preventive Medicine and Psychology, University of Southern California; postal address: 2001 N. Soto St., Los Angeles, CA, 90032, USA

[12] Department of Sport Science, Social and Health Sciences, University of Konstanz; postal address: Universitätsstraße 10, 78464 Konstanz, Baden-Wuerttemberg, Germany

[13] Department of Behavioral Science, University of Texas MD Anderson Cancer Center, postal address: 1515 Holcombe Blvd., Houston, Texas 77030, USA

Author Note

This work was supported by the National Institutes of Health [grant numbers U01HL146327, R01CA240713]; the European Union's Horizon 2020 research and innovation programme [grant number 667302]; a faculty fellowship from the Duncan Family Institute for Cancer Prevention and Risk Assessment at The University of Texas MD Anderson Cancer Center. S.L. acknowledges support by the Klaus Tschira Stiftung. This paper was developed from the 2nd International Physical Activity Assessment Workshop of the Center of the

Assessment of Physical Activity (CAPA), Institute of Sports and Sports Science, Karlsruhe Institute of Technology, Germany.

Correspondence concerning this article should be addressed to Markus Reichert, Karlsruhe Institute of Technology (KIT), Institute of Sports and Sports Science, Mental mHealth Lab / Chair of Applied Psychology, Hertzstr. 16, building 06.31, 76187 Karlsruhe, Germany. Email: markus.reichert@partner.kit.edu

Ambulatory Assessment for Physical Activity Research: State of the Science, Best Practices and Future Directions

In the past decades, physical activity (PA) research benefitted tremendously from technological and digital progress that facilitated objective PA measurement via pedometers, accelerometers, and other wearable devices (Troiano, 2005). While past epidemiological research heavily relied on retrospective questionnaire assessments to address critical issues such as the prevalence and economic burden of physical inactivity worldwide (e.g., Ding et al., 2016; Lee et al., 2012), recent research utilizes smartphone data to shed light on PA inequalities around globe (Althoff et al., 2017). The methodological advantages of digital technologies offer a large potential for scientific endeavors above and beyond associative epidemiological research questions.

PA measurement and surveillance studies mainly use cross-sectional study designs and gather between-subject data. For example, between-subject data can reveal how the health status differs between people with higher or lower levels of PA or sedentary behavior, for example. These information are important in order to better understand the negative impact of sedentary behavior on human health (Ekelund et al., 2019). However, if one wants to know why a person engages in PA at a certain time of the day or in a certain context (e.g. what makes an individual exercise on Thursday at 18:00 or at the gym compared to what makes her or him skip the exercise session), between-subject data is limited in gaining insights into what determines PA engagement within-persons. To answer these questions, a promising approach is intensive longitudinal data (ILD), which can be obtained from frequently repeated within-subject measurements with a high timely density (e.g., daily or sub-daily) in real-life settings. The methodological, statistical, and empirical differences between within- and between-subject effects are described as the “ecological fallacy” (Robinson, 1950; Zawadzki et al., 2017). The fact that between- and within-subject relationships may be different becomes obvious when stressing the example of the relationship of PA with blood-pressure: more physically active persons show habitually lower blood-pressure (negative between-subject correlation), but within-persons PA increases blood-pressure (e.g., when a person climbs stairs; positive within-subject correlation (Kamarck et al., 2003)). Novel within-subject insights about what drives PA promise critical insights for future PA interventions. Put to the extreme and drawing from examples used above: only if one knows about the reasons that hindered a person from exercising at the gym on Thursday at 18:00, this issue can be treated appropriately.

In order to collect such intensive longitudinal datasets in real-time, a group of methods often called Ambulatory Assessment (AA) is very well suited. Following the definition from the Society for AA (*Society for Ambulatory Assessment*, n.d.), AA is described as an umbrella term that encompasses a wide range of methods to study individuals' real-life processes gathering ecological valid data on the behavioral, biological and physiological level (Fahrenberg et al., 2007; Trull & Ebner-Priemer, 2013). Importantly, AA collects intensive longitudinal datasets near or in real-time, providing the data necessary to model dynamic within-person processes and minimizes retrospective and heuristic biases. Thus, the usage of AA methods allows to bypass several challenges of traditional methods, such as interviews, questionnaires, the artificial (laboratory) setting or the reliance on retrospective self-reports. AA can provide the method to unravel within-subject antecedents and consequences of PA as they unfold in everyday life in naturalistic settings. Consequentially, it may be used to inform about the physiological and psychological drivers of PA, which in turn can better inform public health and policymakers on strategies to promote PA.

AA comprises the monitoring of PA, but is not just limited to ambulatory movement assessments (e.g., walking), since this group of methods also includes the monitoring of physiological function and environmental parameters using various sensors (e.g., accelerometers, geolocation tracking, or mobile electrocardiogram), experience sampling method (ESM; e.g., paper-and-pencil diaries) and ecological momentary assessment (EMA; e.g., electronic diaries (e-diaries); Stone & Shiffman, 1994). In particular, AA enables: 1) the capture of multiple assessments across time and consequentially allows for the model and analysis of within-subject processes such as effects of physical behavior on mood (Dunton, Huh et al., 2014; Giurgiu, Koch, Ottenbacher et al., 2019; Kanning et al., 2015; Koch et al., 2018; Reichert et al., 2017), (2) the assessment of data near real-time that comes with minimized retrospective biases, (3) the assessment of data in real-life thereby enabling researchers to collect contextual information that can influence PA and behavior (e.g., daytime, social contact, location characteristics), and (4) the capture of additional information on participants' daily life using a rich set of methods.

The benefits of AA are well highlighted in the literature (Trull & Ebner-Priemer, 2013) and there are several examples that emphasize the importance of studying human behavior such as PA in everyday life. For instance, the "white coat effect" means that the assessment of blood pressure measured in-hospital versus in everyday life does not strongly correlate and often leads to missed diagnoses and consequently poor medical treatment (Salles et al., 2008). The reason for a high ecological validity of real-world assessments is that human behavior follows unique characteristics. Accordingly, humans‘ ability to recall emotions, thoughts, and behaviors is limited (Mehl & Conner, 2012; Trull & Ebner-Priemer, 2013). For example, people remember events more accurately if theyoccurred recently (recency effect), are personally relevant (affective valence effect), and are significant and unusual (memory effect) (Mehl & Conner, 2012; Trull & Ebner-Priemer, 2013). Potentially due to these effects, meta-analyses in adults (Prince et al., 2008) and children (Adamo et al., 2009) have shown a shared variance between self-reported and device-based measured PA of less than 16 percent, indicating that both approaches do indeed measure distinct features of PA. However, even though device-based measures of PA provide more consistent results, it is important to mention that self-reports can add unique information such as type of activity (e.g., walking or playing tennis) or movement quality (walking with pain or painless walk) (Skender et al., 2016).

In practice, many AA studies use smartphones(Trull & Ebner-Priemer, 2013), additionally offering access to a large variety of sensor data (e.g., on geolocations, social contacts and interactions, weather conditions, phone calls, messages). These tools allow researchers to ask participants repeatedly in real-time and in everyday life about their experiences (e.g., mood, stress) and behavior (e.g., PA). More sophisticated methods are the triggered e-diaries (i.e., event-contingent EMA or context-sensitive EMA (Intille, 2007), which ask participants to report their current behavior, thoughts and feelings at predefined moments or contexts of interest, manually-initiated or sensor-informed automatically-initiated. For example, Ebner-Priemer et al. (2013) developed activity-triggered e-diaries: a real-time algorithm running on an accelerometer triggered e-diary prompts every time participants' PA was heightened or reduced compared to predefined thresholds (Ebner-Priemer et al., 2013). This approach has been shown to prevent missing episodes that are of interest, which may happen when using traditional time-based designs (such as triggering every hour). Moreover, it has been found to maximize the assessment of within-subject variance of interest (Ebner-Priemer et al., 2013) and has been shown to be feasible in several studies (Dunton et al., 2016; Dunton, Dzubur et al., 2014).

In this consensus paper, we compile knowledge from various experts in the field of PA and AA research to provide a summary of the current state of the AA science for the measurement of PA, in addition to best practice recommendations including methodological frontiers, and future directions. In particular, this consensus paper will expand on triggered diaries, psychometric properties, latest technological developments, geolocation tracking, and the usage as well as the translation of AA to real-life interventions. In each of these sections, readers are given information about what to consider when starting a PA research endeavor using AA and about what is already feasible with these different AA tools. Moreover, readers are also provided with existing frontiers and progress to be expected in the upcoming years.

This paper is a result of the 2nd International CAPA Workshop 2019 "Physical Activity Assessment – State of the Science, Best Practices, Future Directions", where invited researchers with experience in PA assessment, evaluation, technology and application participated in topics of either PA questionnaires, accelerometers, or ambulatory assessment. The goals of the workshops were to discuss and create consensus statements on best practices for application and future directions for research of each methodology. Specifically, the workshop participants were informed that the aim of the resultant special issue is to motivate practitioners to use state-of-the art PA assessments. The second aim is to motivate the researchers to innovatively progress PA and SB assessment research. The current consensus paper is one of three papers in this special issue derived from the workshop and is focusing on the discussion of AA.

Best Practice Recommendations and State of the Science

Overall, AA methods to assess PA and related variables of interest encompass self-report, observational, and physiological/behavioral (including device-based) methods. Self-report AA methods comprise increasingly digitized ESM (historically using paper-and-pencil diaries) and EMA (via e-diaries) (Trull & Ebner-Priemer, 2013). Observational methods comprise, for example, data from video cameras or geolocation-tracking. Physiological/behavioral AA encompasses device-based methods that continuously monitor current PA behavior, e.g., via accelerometry or heart rate assessment.

There are manifold challenges when planning AA studies and landmarks are not yet categorized clearly. To provide support for researchers in planning AA investigations, landmark decisions occurring are summarized in Figure 1.

Subjective Measures of Ambulatory Assessment: E-Diary Self-Reports

Recent EMA approaches most often use smartphones, which entail e-diaries to query people for their PA behavior and related variables of interest. While traditional self-report questionnaires on PA rather represent a remembering or “believing self”(that is people report on their past PA as they recall it), ESM and EMA are less prone to recall biases and rather represent an “experiencing self” (Reis, 2012), because subjects are asked during or immediately after the behavior of interest. EMA can be combined with device-based methods (e.g., accelerometers and geolocation tracking) to assess the frequency, duration, intensity, and context of PA (Von Haaren-Mack et al., in press). Moreover, EMA and ESM can capture additional self-report information about PA, which is difficult to capture through device-based measures, for example, the type of PA conducted and the perceived exertion, or how much weight is carried. In sum, ESM and EMA are less prone to response biases than traditional self-report measures such as paper and pencil questionnaires, but ESM and EMA can be more burdensome for the participants.

Capturing Different Variables of Interest Using e-Diaries in Real-Life

EMA self-reports on PA can refer to PA itself and to experiences associated with PA (Shiffman et al., 2008). For PA, its occurrence, type (e.g., swimming, weight lifting, yoga) (Knell et al., 2017), duration, and contexts/domain (such as leisure time, work, commute) are features that can be assessed (Knell et al., 2017). Further variables related to PA and of interest for PA researchers can refer to health behaviors (e.g., nutrition or drug consumption) (Bruening et al., 2016; Morgenstern et al., 2014), context (e.g., environment, social interactions), cues and barriers (e.g., affect, mood, intentions and self-efficacy) (Dunton et al., 2007) (Niermann et al., 2016). Beyond these aspects, variables with important implications for clinical research, e.g., related to symptoms such as stress, anxiety (Wilhelm & Roth, 1998), pain (Spenkelink et al., 2002), or fatigue (Kop et al., 2005) may also be of interest. Capturing all these information is a strength of self-report measures.

Item Selection. EMA studies are characterized by multiple repeated assessments across time, thus standardized PA questionnaires in their entirety are not feasibly to be used due to their length and wording (this refers to retrospective PA events, e.g., “In the last two weeks, …”) (Mehl & Conner, 2012; Trull & Ebner-Priemer, 2013). When assessing PA in EMA studies, it is useful to apply items/questions that are based on existing theoretical concepts of PA (Caspersen et al., 1985; Pettee Gabriel et al., 2012) and to search for items/questions which have been tested for psychometric properties in previous EMA studies (Knell et al., 2017). Importantly, one should ask for the subject´s current state (or very recent state) instead of asking to recall their state over an extended period of time, especially if one is interested in variables known to change rather quickly such as mood (Mehl & Conner, 2012; Trull & Ebner-Priemer, 2013). Furthermore, it is reasonable to conduct pre-tests prior to the final decision for an instrument regarding items of each questionnaire and each prompt, e.g., determine whether participants can easily understand the item (Bowden et al., 2002). For general questionnaire selection guidance please refer to Nigg et al. (2020, within this issue).

Psychometric Properties. Reliability and validity are two central requirements by which the psychometric properties of virtually all psychological assessments are typically evaluated. Reliability refers to the degree by which the obtained measurements are free of measurement error. In classical test theory (CTT), reliability is defined as the proportion of true score variance to observed variance, which is the sum of true score variance plus error variance (e.g., McDonald, 1999). Validity can be examined from multiple perspectives (see terms such as content validity, construct validity, or criterion-oriented validity; Cronbach & Meehl, 1955), but in its most basic form, validity refers to what degree the measurement is a good approximation of the construct it attempts to measure. To that end, theoretical considerations play a central role in determining the validity of measurement, and statistical indicators (e.g., correlations with external criteria to which the construct should be related) are often used to support these considerations.

AA data typically has a multilevel structure with repeated measurements nested within individuals. This allows to examine relations among constructs on at least two levels: between-persons (e.g., are persons who engage in more PA happier on average?) and within-persons (e.g., is a person happier when (s)he engages in more PA than (s)he usually is?). Importantly, between-person and within-person relations of constructs are statistically independent, showing that because an association has been found on the between-person level, this does not guarantee that it will also hold true in the within-person level or vice versa (see e.g., (Hamaker, 2012; Molenaar, 2004; Voelkle et al., 2014). From the statistical independence of these two levels of analyses it follows that reliability (which is often examined as internal consistency, i.e. a measure depending on inter-item correlations) and validity also need to be examined separately for the within-person and between-person level. There are several introductory articles on this topic (e.g., Brose et al., 2019; Cranford et al., 2006; Geldhof et al., 2014; Nezlek, 2017; Wilhelm & Schoebi, 2007).

Here, we aim to stress that it is paramount importance to examine reliability and validity, specifically for the level of interest to the research question (see also Brose et al., 2019). For example, if a researcher is interested in the question of whether there is a within-person effect of PA on mood (is a person happier when (s)he engages in more PA than when (s)he usually does?), (s)he needs to ensure that both measures of PA and mood reliability and validly capture within-person changes in these two constructs. Establishing reliability and validity on the between-person level does not provide evidence for reliability and validity on the within-person level. For constructs that are assessed using several indicators (e.g., a questionnaire that assesses PA with three items at each measurement occasion), estimators of reliability can be obtained using for instance multilevel extensions of the widely popular internal consistency estimates Cronbach's α and McDonald's ω (see Geldhof et al., 2014). These within-person reliability estimators (as well as comparable estimators; e.g., Cranford et al., 2006; Nezlek, 2017) assess the extent to which items supposedly assessing the same construct (e.g., "How physically active have you been today?"; "How do you rate the intensity of your physical activity today?"; "For how long have you been physically active today?") and co-vary on the within-person level. With increasing within-person correlations among these items (i.e., on days when a participant indicates higher scores on item 1, (s)he is also more likely to indicate a higher response on item 2 and item 3), internal consistency will become larger (closer to the theoretical maximum of 1), and the measure would be deemed reliable on the within-person level from the perspective of CTT. To the extent that items are uncorrelated, the reliability estimator will tend to zero.

This demonstrates that multilevel α and ω (as well as the alternative approaches listed above) share a crucial assumption regarding the underlying measurement model: only if the measurement follows a reflective measurement model (i.e., a latent variable "predicts" responses on the single indicators), measures of internal consistency are reasonable estimates of measurement's reliability. In case the measurement model would be better described by a formative measurement model (in which the indicators "predict" the latent variable), internal consistency estimators are no valid approximations for the reliability of the measurement (see Edwards & Bagozzi, 2000) for a comparison of reflective and formative measurement models). Considering the example of intensity and duration of PA: it seems reasonable that these two items do not necessarily positively co-vary on the within-person level (e.g., a person might not necessarily exercise longer on days when (s)he exercises more intensely). In this case, the within-person correlation of these two items tends to zero confirming low reliability for the measure, meaning the reliability estimate has been derived from an inappropriate measurement model. Rhemtulla et al. (2019) illustrate potential biases that arise when reflective measurement models are applied as default in cases when a formative measurement model might be more appropriate.

In sum, the reliability and validity of psychological assessments have been discussed in the literature for decades, but the implications of these discussions for AA data have only recently received thorough attention. In these designs, it is important to examine reliability and validity separately for the levels of interest (within-person/between-person) and to report results on these measurement properties accordingly. Geldhof et al. (2014) provide annotated Mplus code to estimate multilevel α and ω; an alternative estimation approach in the framework of generalizability theory using the psych package in R illustrated in a tutorial by (Revelle & Wilt, 2019). If applicable to the research context, these multilevel reliability estimates should be reported for all relevant constructs. However, researchers should be aware of the assumptions underlying widely used reliability estimators (positive inter-item correlations resulting from a reflective measurement model) and should carefully consider whether these assumptions are reasonably tenable in the context of the respective work and discuss resulting reliability estimators accordingly.

Device-based Measures of Ambulatory Assessment

Accelerometry. Accelerometry is based on continuous and real-time measurement and recording of movement-induced raw acceleration signals over a specific period of time. Accelerometers register intensity and duration of single- or multi-axial accelerations and convert this raw data into manufacturer- and model-specific outcome metrics. For detailed information on accelerometry for PA research, please refer to Burchartz et al. (2020, within this issue) and to Von Haaren-Mack et al. (in press).Burchartz et al. (2020, within this issue) reflects the authors' expert consensus on the topic of accelerometry for movement behavior assessment include highly relevant discussions on recording movement behaviors, while using different metrics and the impact of study design and technical decision on data collection with accelerometers. EMA approaches have widely been combined with accelerometers (e.g., Giurgiu, Koch, Ottenbacher et al., 2019; Koch et al., 2018; Niermann et al., 2016; Reichert et al., 2016, Wunsch et al., 2019). For example, combining PA measurement in real-life with e-diaries repeatedly assessing psychological variables such as mood enables an ecologically valid investigation of within-subject processes across time. A combination of accelerometers with e-diaries can also have very practical implications; for example, it can help to instruct participants to wear the sensor (thus fostering adherence) and to record non-wear time via self-reports (on the e-diary).

Contextual Measures via Geolocation Tracking. Besides other methods for the assessment of contextual variables such as (body-) cameras or self-reports with e-diaries, geo-location tracking offers new and interesting possibilities specifically for PA measurement and for PA research in general. In combination with geo-informatic methods, geo-location data allows for the characterization of PA types in more detail: PA can be located, for example, workouts at the gym versus gardening (Reichert et al., 2017). Moreover, geolocation tracking can help investigate contextual influences, such as to address how different city environments may influence PA (Althoff et al., 2017). It can be used to trigger EMA assessments in certain locations (Törnros et al., 2016), for example, every time a participant enters the gym. It can also help to guide PA interventions, such as routing people through green areas in cities thus potentially increasing PA and enhancing well-being (Tost et al., 2019).

One important opportunity may be to incorporate contextual influences into PA research. Human behavior is critically influenced by context variables (Reichert et al., 2020). People will behave differently if they are together with peers compared to being alone, during the summer holiday compared to being at work in winter, or when exposed to traffic noise compared to a situation without traffic noise for example. All kinds of context variables potentially influence human PA. One also needs to consider that people may choose to go to certain environments that enable them to perform intended behaviors, referred to as selective mobility bias (Fong et al., 2018). For example, one might only want to exercise in a gym environment, despite having a park nearby their home. One other related methodological issue is the uncertain geographic context problem (Kwan, 2012), which highlights the concern that being in close proximity to an environment does not always equal to actually spending time in that environment. Fortunately, technological advances allow us to assess contextual characteristics through the use of geolocated measures. Existing data sets on all kinds of contextual parameters, e.g., on land use, average noise level, and potentially on average noise pollution can be incorporated simply by intersecting tracked positions by available spatial layers. While measurements from sensors attached to the participant would provide a better indicator for the current exposure to environmental effects, information on average environmental conditions could act as a proxy if sensors are not available or temporarily dysfunctional.

Furthermore, measurements from, e.g., accelerometers can be interpreted more thoroughly if the position can be mapped to the type of building or land use. Accordingly, accelerometer signals from inside gyms or sports pitches could thereby be distinguished, e.g., from signals at home, in shopping malls or in green spaces. Here, freely available datasets such as OpenStreetMap provide a rich set of context information. A well-established set of geoinformation algorithms allows to derive additional layers of information, e.g., the distance to locations such as green spaces, gyms, or noisy streets, visibility of locations such as green spaces, water bodies or scenic outlooks, and the steepness of the terrain or the shadiness of the current position. Spatial explicit information from social media provides additional information. Examples are the crowdedness or emptiness of a location estimated by the number of tweets at this position, and the presence of sign of physical disorder such as abandoned cars or graffiti based on an analysis of georeferenced imagery from data sources such as Google Street View®, Mapillary® or Flickr® (e,g., Quinn et al., 2016; Rzotkiewicz et al., 2018). Such kind of spatial data in AA has been used for estimation of travel mode or activity (e.g., Brondeel et al., 2015; Jansen et al., 2017; Stewart et al., 2016; Stewart et al., 2018).

However, several other options exist to incorporate geoinformation and geoinformatics in AA for PA research. For example, survey triggering could be implemented under consideration of the current position (e.g., distance to the position of last survey, land use, noise level, indoor/outdoor). Törnros et al. (2016) and Dorn et al. (2015) have utilized land use information to trigger surveys, ensuringthat land use classes of interest (e.g., urban green spaces) are represented sufficiently in the sample (Dorn et al., 2015; Törnros et al., 2016). Indeed, intervention strategies could profit from information on the context of the current location: is the current location suitable for a specific type of PA or is there a suitable location nearby? Routing offers further potential to be explored: in addition to proprietary services such as Google Maps®, Apple Maps® or Bing Maps® specialized open source services such as the OpenRouteService are available which offer specialized options for pedestrian, bicycle or even wheelchair routing (Mobasheri et al., 2017; Zipf et al., 2016). Routing algorithms could be used to suggest a route to the closest location suitable for an intended exercise session. The suggested route could potentially also consider factors such as noisy streets or greenness along the route (Novack et al. 2018). Routing algorithms could also be used to expose participants to environmental stressors or distresses –by routing them along a noisy street, a green route or through a quiet neighborhood – and measure the effectiveness of such treatments by triggering surveys prior to and after this intervention.

Physiological Methods

AA allows for the objective measurement of physiological parameters in everyday life. For example, Electrocardiography (ECG) can be a useful instrument in assessing different parameters concerning the heart. One of the most commonly used ECG parameters is heart rate (HR), which can yield information about physical or psychological stress and is often used as an immediate feedback method during PA (Dunton et al., 2005). Besides HR, heart-rate-variability (HRV) is a parameter which has increasingly become popular during the last decade among sport and exercise scientists (Ludyga et al., 2019). Using HRV makes it possible to assess the activation of the parasympathetic nervous system in a non-invasive way. This information can be used to detect symptoms of overtraining early or to detect the point in time where the physiologic systems is most adaptive for training stimuli (Brosschot et al., 2007; Dunton, Huh et al., 2014; Hynynen et al., 2006). In addition, electrodermal skin conductance (EDA) can also be included in PA assessment to measure changes in physiological arousal (Poh, Swenson, Picard, 2010), as well as picture the sympathetic activity of the nervous system, and furthermore for long term measurements (Poh, Loddenkemper et al., 2010). However, data quality of such recordings has been discussed controversially. Moreover, blood pressure (BP) can be assessed in real-life settings and is a valuable parameter. For instance, BP is utilized among hypertension patients to monitor and adjust PA duration and intensity as is best for the current intra-individual condition. The measurement of 24 hours ambulatory BP has become a common parameter for both diagnosis and intervention (Wilhelm et al., 2012). Moreover, ambulatory cortisol assessments can be linked to PA to study the link between daily stress and PA (for details refer to Dunton et al., 2015). Nevertheless, it needs to be mentioned that all of these physiological methods require additional specialist equipment and produce large amounts of data that needs to be pre-processed and analyzed by experts or expert-developed algorithms to properly interpret the data.

Combination and Integration of Distinct Measures

Combining the above-mentioned measures can help to unravel correlates of PA and their complex interrelatedness in real-life. In a feasibility study, Ben-Zeev et al. (2015) provided participants with smartphones and a wide range of sensors and software that enabled the continuous tracking of geospatial activity (via geolocation-tracking), PA (via accelerometry), sleep duration (via device-usage information, for example accelerometer inferences, ambient sound features, and ambient light levels) and speech duration (via microphone and speech detection algorithms. The researchers found that sensor-derived sleep duration and variability in geospatial activity was associated with daily stress levels and showed associations between changes in symptoms of depression and sensor-derived speech duration, geospatial activity, and sleep duration.

One further advantage of AA for PA research is the capability to combine the strengths of different assessment methods of PA and minimize the inaccuracies of a single assessment method. For example, while accelerometers have been shown to be an inaccurate assessment of specific activities (e.g. cycling, swimming), self-report measurements suffer from low reliability and validity according to recall difficulties (Jekauc, 2009; Jekauc et al., 2014). The inaccuracies of both methods could be overcome by combining them using AA. However, only few studies used this method combination to improve measurement properties of PA assessment. However, triangulation of multiple measures is complex and could lead to misinterpretation of PA as many of the independent measures have noted limitations. A simple combination might not eliminate these inaccuracies, therefore raising the need to control for them (e.g. reporting bias, reactivity, non-wear-time, etc.).

Finally, many of the above-described methods are integrated in commercial smartphones or commercially available sensors that combine different types of measures (e.g., the movisens LightMove 4[1] or the ActiGraph wGT3X-BT[2], combining measurements of light intensity and composition and triaxial movement acceleration, which enables the assessment of wake- and sleep-times; or the EcgMove 4[3], which combines cardiovascular measures with accelerometry) and can be readily used in AA studies.

Sampling Design and Triggered E-Diaries

The sampling design is a highly critical part of AA studies since it directly influences the data and can impact results. For example, if a researcher is interested in within-subject associations of PA and mood, and if the sampling design leads to queries for mood only in phases of inactivity (which may happen by chance if participants are prompted every hour), this results in restricted PA variance in the data. Thus, the researcher will not be able to

[1] https://www.movisens.com/en/products/light-and-activity-sensor/
[2] https://www.actigraphcorp.com/actigraph-wgt3x-bt/
[3] https://www.movisens.com/en/products/ecg-sensor/

detect an association between PA and mood just because of an inappropriate sampling scheme leading to a false-negative finding.

If designing a sampling strategy, first, the *wave duration* (i.e., the number of waves or measurement bursts, for example, three monitoring periods/year) and the *monitoring period* (i.e., the number of days/wave) should be defined. Second, one has to choose whether individuals should respond to queries that are self-initiated (event-based) or prompted (fixed or random time intervals) or triggered by sensor data (Ebner-Priemer & Trull, 2009; Trull & Ebner-Priemer, 2013). The moments of assessment should be representative of an individual´s experience. Thus, the aims of the study will guide the decision for the *sampling strategy* (Shiffman et al., 2008). While event-based sampling focuses on particular events or episodes (e.g., discrete exercise sessions), time-based sampling with fixed and/or random time intervals schemes aim to capture the entire experience (e.g., how exercise intentions vary over time) (Liao et al., 2016; Shiffman et al., 2008). Researchers can either use one sampling or combine. For instance, if a high outcome variance is of particular interest, a combined sampling scheme may be a promising approach. (i.e., event-based sampling with random intervals), whereas a pure event-based design might be the sampling of choice if researchers are interested to gain more information about a specific event (e.g., social context of PA or emotions during PA). Third, the *prompt frequency*, which is the number of prompts per day and wave depends on the time frame, during which one expects the outcome to change. In previous studies, the time-frame varied from 15 minutes to 4 hours in PA studies (Liao et al., 2016; Romanzini et al., 2019). The prompt frequency could also just be once per day (in daily diary studies) and depends heavily upon what burden can be tolerated by participants, which varies across sub-populations. Since there are various types of PA, the specific research question guides the sampling design: it makes a difference whether one is interested in PA that are rather sporadic (e.g., climbing stairs), longer enduring episodes (e.g., exercise sessions) or both. In addition, characteristics of the target group (e.g., children and youths, working adults, older adults, patients, etc.) may force researchers to adapt the schedule (Maher et al., 2018; Wen et al., 2018). Moreover, the act of answering an e-diary can be difficult during exercise sessions itself (i.e., individuals may not carry phones with them or be able to stop and answer while exercising). This can be a major source of missing data and thus constitutes a limitation of this kind of assessment. Accordingly, researches should consider prompting e-diary questions, for example directly prior to and after exercise sessions.

Since sampling strategy has a significant impact on the data, Ebner-Priemer et al. (2013), developed and applied triggered e-diaries (e.g., activity- and GPS-triggered e-diaries can be utilized to improve the number of assessments during episodes of interest). In comparison to non-triggered time- and event-based designs, triggered e-diaries are connected to external devices such as accelerometers or geolocation-tracking-systems and trigger participants in situations of interest. This approach improves the assessment of within-subject variance of interest and minimizes participants' burden (Dorn et al., 2015; Ebner-Priemer et al., 2013; Törnros et al., 2016).

To illustrate this, triggered e-diaries have been successfully used to investigate whether prolonged sedentary behavior is negatively associated with health outcomes (triggering in phases of sedentary behavior using accelerometer data; Giurgiu, Koch, Plotnikoff et al., 2019) and can be applied to study whether the environmental context moderates the association between PA and well-being (e.g., triggering if participants access green spaces using geolocation data; Törnros et al., 2016).

Given the rapid digital progress, researchers now have the opportunity to collect data in ways that were inconceivable one or two decades ago, which enables them to assess more parameters objectively, continuously, and apply them to triggered e-diaries (e.g., glucose values in the blood). Moreover, a combination of random and triggered diaries is a promising approach to assess constructs of interest in special (triggered) situations and in general (random) situations to analyze differences. Finally, triggered e-diaries provide a sophisticated approach for just in time interventions in real-life. For instance, if a situation of interest is associated with an unhealthy behavior (e.g., sitting longer than 30 minutes), this situation can be defined as a situation of interest, which triggers the intervention (Myin-Germeys et al., 2016).

Reactivity, Compliance and Participant Burden

Researchers need to consider that measurement reactivity, defined as the potential of a behavior or experience to be affected by the assessment (Shiffman et al., 2008). This may be driven by the motivation and opportunity for subjects to take control of the behavior (Korotitsch & Nelson-Gray, 1999). Regarding reactivity to the measurement of PA, the work of Clemes et al. (2008; Clemes & Parker, 2009) revealed that wearing a sensor can indeed influence PA behaviour. In particular, Clemes et al. (2008) compared the number of steps measured in a week where participants were not aware that their steps were recorded to a week where the same participants were informed that their steps were counted. The number of steps increased significantly in the second condition (M = 9541, SD = 3186 steps/day vs. M = 11385, SD = 3763 steps/day) (Clemes et al., 2008). However, in a subsequent study, Clemes and Deans (2012) found that measurement reactivity does not persist over time; in their sample of adults, pedometer recorded steps returned to normal after one week of measurement, even if participants recorded the steps in a daily diary. Scott et al. (2014) specifically examined different methods for dealing with reactivity to tracking PA with accelerometers and pedometers in adolescents. Their findings suggested that hiding the results of the measurements from participants could combat reactivity. Importantly, hiding measurement results for the day only was not enough; rather, a longer time span (e.g., week) was necessary. Accordingly, feedback on PA (via smartphone or on the accelerometer) may cause reactivity, thus researchers may consider not giving feedback to limit reactivity.

Compliance, defined as participants' adherence to the assessment protocol (Shiffman et al., 2008) and usually parameterized as accelerometer wear time and percentage of EMA prompts answered, is another source of potential bias that can be influenced by the sampling design (Liao et al., 2016). EMA can be applied to increase accelerometer wear-time and to improve the reporting of non-wear times as described in Burchartz et al. (2020, within this issue). Previous EMA studies investigating PA have shown relatively high compliance rates, e.g., ranging up to 96% (Liao et al., 2016; Maher et al., 2018). Reasons for missing data and non-compliance should be reported in publications (for reporting guidelines see Liao et al., 2016; Trull et al., 2019). To reduce this non-compliance, researchers should consider how rapidly the target phenomenon is expected to change, use the fewest number of prompted surveys possible to answer the questions or interests (for EMA; Dunton, 2017), and aim to use sensors that are as convenient as possible (e.g., small accelerometers attached to the wrist). If participants are able to install EMA software on their own smartphone this usually comes with increased compliance (carrying a second study-smartphone is often problematic), but researchers have to be aware of potential technical incompatibilities, especially if smartphone and sensors are thought to communicate with each other.

Technological Advancement and its Potential Use in Ambulatory Assessment

Extending the Assessment Battery: Acoustic Signals, Ambient Light Detection and Biomarkers

Ongoing technological and digital progress offer numerous possibilities for future real-life assessment of parameters potentially impacting human PA. In the following, selected parameters and possibilities for their real-life assessment will be reviewed. First, to parameterize social contact and interaction, nearby Bluetooth device detection, RFID are already in use, while so-called electronically activated recorders (EAR) to sample acoustic information of participants in their daily life have been already applied two decades ago (Lammers et al., 2008). Nowadays, voice recording via smartphones is no problem at all from the hardware perspective (Xu et al., 2013) and new algorithms for voice recognition are being developed. However, in this context data privacy is an open issue. Second, daylight exposure is positively related to overall health and sleep quality (Boubekri et al., 2014), and thus a parameter of interest to researchers. The usage of ambient light sensors allows to objectively measure daylight exposure of participants and to gain knowledge on their wakeand sleep-times (Gaston et al., 2012). Third, continuous measurement of biomarkers started with invasive glucose sensors in 1954 when Leland C. Clark invented the “Clark electrode”, which was able to measure oxygen/ glucose in different solutions (Wang & Lee, 2015). Today there are a variety of non-invasive sensors categorized as transdermal and optical sensors (Wang & Lee, 2015). Fourth, cortisol is an important biomarker related to PA. While a continuous measurement is thus far not feasible, participants can already be triggered via EMA to measure cortisol at multiple times a day and is possible to provide participants with real-time feedback of the cortisol level on their smartphone application (Choi et al., 2014).

AA Research Contributing to Ecological Momentary Interventions (EMI) Fostering PA

As described above, the question of what drives PA everyday life can only be answered using intensive longitudinal within-subject data. Once a set of the most important momentary antecedents of PA has been established, the critical step will be to translate these findings into ecological momentary interventions (e.g., Heron & Smyth, 2010) fostering PA. If successful, AA research might enable the development of tailored behavioral interventions that are automatically provided over mobile devices at the time and in the place where needed. In this way, AA research may contribute to advances in future ecological momentary interventions that fosters everyday PA, and thereby support human somatic and mental health, and solve one critical issue of humankind in the 21st century namely the high levels of physical inactivity. In the following section, we approach the issue on how to translate findings from AA research on PA into ecological momentary interventions (EMI) from a psychological perspective of behavior change and provide an example of an already existing EMI.

It has been stated in numerous well-established theories that the initiation and execution of any purposive (or functional) behavior requires the initial existence of an intention to act, or at least some kind of positive attitude towards the aspired behavioral outcomes (e.g., Ajzen, 1991; Bandura, 1986; Carver & Scheier, 1982; Noar & Zimmerman, 2005; Prochaska & DiClemente, 1983; Webb & Sheeran, 2006). However, while many people strive to maintain a healthy lifestyle (e.g., being physically active, exercising regularly, or to sustaining a healthy diet), they often fail to act in line with their intentions (e.g., Armitage & Conner, 2001; Rhodes & Bruijn, 2013; Webb & Sheeran, 2006). Research has often investigated the intention-behavior relationship, comparing the behavior of individuals with different strengths of intention whereby intentions and behavioral outcomes are usually measured at once (prospectively and/or retrospectively) (Webb & Sheeran, 2006). This kind of research demonstrated a strong association between the intensity of intentions and the execution of corresponding behavior (e.g., Sheeran, 2002).

However, meta-analytical exploration revealed that the temporal stability of an intention has a considerable influence on whether the desired behavior is realized (Cooke & Sheeran, 2004; Noar & Zimmerman, 2005; Sheeran & Abraham, 2003). Only a few studies examined the intention-behavior relationship at the within-person level considering time-dependent changes in participant's intentional strength. Those studies reveal that time-dependent within-person variability of intentions predict the realization of health behavior in daily life, with higher accuracy than between-person differences in intentions (e.g., Conroy et al., 2013; Inauen et al., 2016; Kiene et al., 2008; Maher et al., 2016; Maher et al., 2017).

In addition, when testing the effectiveness of interventions aimed at increasing PA, it is necessary to not only consider whether participants' activity levels are increasing (e.g., via accelerometer measurements), but to also differentiate whether the behavior recorded in everyday life was actually intended. To assess potential intervention effects accurately, it is necessary to ensure that observed behavioral changes (the occurrence of PA) occurred intentionally. For this purpose, the combination of accelerometer data and e-diaries appears suitable to capture participants' behavior depending on their individual training schedule or their individual intentions for active exercises.

Mechanisms involved in the realization of health behavior (such as PA) are theoretically conceptualized as within-person processes. It is presumed that peoples' momentary (cognitive, affective, motivational) capacities can be influenced by situational determinants and context-variables, which may in turn influence behavioral intentions as well as their actual realization (Brand & Ekkekakis, 2018; Brand & Schweizer, 2015; Ekkekakis, 2017;

Jekauc et al., 2015; Jones et al., 2018; Sheeran et al., 2005). Moreover, desired and undesired behavior can be automatically triggered by environmental stimuli eliciting specific stimulus-response patterns that are learned or based on previous experience (Aarts & Dijksterhuis, 2000; Bargh & Ferguson, 2000; Strack & Deutsch, 2004; Wood et al., 2002). Increasing evidence suggests that unconscious cognitive, affective, and motivational processes have a significant influence on health-related behavior and PA (Calitri et al., 2009; Conroy & Berry, 2017; Jones et al., 2018). Little research has been conducted on the conditions under which PA (or health behavior in general) is regulated by intentional or automatically operating cognitive-affective processes (e.g., implicit attitudes or habits) in everyday life. However, there is evidence that under critical conditions (e.g. short-term limitations of cognitive capacity or low self-control), automatic or unconscious processes can gain predominance and trigger behavior that is against an individuals' intention (e.g., Conner et al., 2007; Friese & Hofmann, 2009; Hofmann et al., 2009). Future studies using AA methodology could provide additional insight into the conditions under which intentional and non-intentional processes determine the occurrence or absence of PA (or health behavior in general) in everyday life.

Once these processes and interdependencies are understood, effective real-time interventions can be developed. Such interventions can be deployed in the form of micro-randomized trials (Klasnja et al., 2015) or within-person encouragement designs (Schmiedek & Neubauer, 2019). The former might be relevant when the core focus of an intervention is on increasing a concrete behavior (e.g., PA), whereas within-person encouragement designs can be helpful when the goal of an intervention is not (only) the change in behavior, but also an improvement in a distal outcome (e.g., momentary fatigue).

The ongoing real-life exercise intervention of the H2020-funded European project "Comorbid Conditions of Attention-deficit/hyperactive disorders (CoCA)" to prevent comorbid obesity and depression in patients with Attention-deficit/hyperactive disorders (ADHD) (Mayer et al., 2018) is one of the few outstanding examples that demonstrates current technological possibilities for real-life exercise interventions and provides us with frontiers to be undertaken in the future. In CoCA project, participants use a mobile-Health (mHealth) system' comprising a smartphone and an accelerometer on the wrist, both connected via Bluetooth low energy and a commercial software tool[4] including questionnaires, exercise videos, reminders, and motivational phrases to increase adherence with the exercise intervention.

The highlights of the CoCA exercise intervention include aerobic and strengthening exercise instructions provided through custom-made exercise videos on the smartphone. They are individually and automatically adapted every one and a half weeks in terms of the difficulty and intensity of the exercise program. As shown in Figure 2, automatic feedback is realized via WIFI connection and real-time analyses on the server providing participants with their training progress such as the percentage of watching and executing the strengthening videos and with distinct PA outcomes as quantified by the accelerometer (see Mayer et al., 2018).

Discussion and Recommendations

This consensus paper summarizes the current state of AA for PA research and provides recommendations for best practices. AA offers enormous opportunities for researchers to capture a rich amount of data – from behaviors, physiological, psychosocial, and cognitive states, to social interactions and built environments, all in the context of a person's daily life.

While associations between a physically active lifestyle and health outcomes (Kirchner & Shiffman, 2016; Schlicht et al., 2013) can be researched using AA (an issue tackled in many epidemiological studies), the strength of this group of methods is that it enables researchers to examine the dynamic patterns of change in PA together with contexts, emotional states, beliefs, attitudes, and perceptions in real-life (Dunton, 2017). Since PA and the latter aspects vary intra-individually, within-subject investigations may examine temporal trends (Wilhelm et al., 2006), dose-response relations, and reciprocal associations between PA and

[4] https://www.movisens.com/en/products/movisensxs/

psychological variables (e.g., stress, emotions, self-concept, motives and barriers to being physically active) (Niermann et al., 2016), and research questions may target within-person intervention effects of PA.

When using AA such as triggered e-diaries, researchers need to:

- assess the variable of interest (e.g., PA, sedentary behavior, environmental factors) objectively and continuously and with the possibility of feedback (i.e., a technical interface such as Bluetooth Low Energy (BLE)) in daily life,
- Define, in which situations they want to assess the constructs of interest (e.g., special activity threshold, time of sedentary behavior, special kind of environmental areas like green spaces), and
- develop a trigger algorithm or use a developed trigger software to facilitate assessment during these defined situations.

Given the many advantages in AA, the number of PA research studies that utilize this type of assessment has been increasing exponentially in the past few years. While multilevel modeling is now widely used in data collected from EMA to disaggregate the within- and between-person effects, less attention has been given to the psychometric properties when designing and reporting the EMA survey items. In this consensus paper, the importance of examining the reliability and validity separately for within- and between-person levels was given special attention. All authors highly encourage future EMA studies to include results on these measurement properties when reporting findings.

Limitations of AA

The technology for AA is rapidly developing. Integrating data from multiple assessment sources allows researchers to tackle different predictors and consequences of PA related behaviors, and it offers a unique opportunity to tailor interventions on those behaviors. The current consensus paper presents some examples of linking EMA and accelerometry to enhance both behavioral assessment and intervention. Of course, there are also challenges that merit discussion. First, researchers should balance participant burden. To achieve a high level of adherence, researchers may tailor the sampling scheme by reducing the number of items or the number of study days (Trull et al., 2019) or make use of unobtrusive assessments via sensors (Exler et al., 2015). Translated into practice, rerearchers should rather ask few EMA questions on few assessment points, gather valid data and use unobtrusive real-life measures instead of overstraining participants. Put simply, if many questions are being asked, the frequency of prompts should be reduced and vice versa. For example, eight to ten prompts per day in a wave with up to seven days have been employed successfully in several studies with healthy samples and these studies showed good e-diary compliance on average > 80% (Koch et al., 2018; Reichert et al., 2016; Reichert et al., 2017). Second, as more researchers are relying on technologies in their studies, the potential privacy concerns and data security need special attention, especially when studying vulnerable populations such as children and clinical populations. Most research institutions have their own guidelines and regulations for participant privacy and data security. In addition to following the institutional guidelines, it is also important to assess and evaluate participants' concerns regarding technology and privacy. This information would be very useful to inform the development of future studies, especially when scaling up to larger populations. Third, in comparison to traditional methods such as paper-pencil questionnaires, the usage of technical devices (e.g., accelerometers and smartphones) is relatively expensive and associated with considerable data management and pre-processing issues, which are time-consuming.

Lastly, researchers need to be mindful about technological challenges that may arise while utilizing AA. For example, technological challenges for connections issues such as the limited stability of Bluetooth connections (distance between the smartphone and the accelerometer currently needs to be less than ten meters), and limited mobile network coverage in rural places. There are also challenges of a limited battery life of the accelerometer and smartphone. However, ongoing technological and digital progress will improve these limitations concurrently leading to a decline in sensors and Smartphone costs.

Implications for sport and exercise psychology research

The use of AA opens promising avenues for sport and exercise psychology research. For example, AA studies can foster insights how sport, exercise and psychological variables (such as motivation, mood, rumination, self-esteem) are associated with each other across time in everyday life. Put simple, AA can help to answer questions such as 'How do people feel before/during/after sports compared to before/during/after exercise in their everyday life?', 'Which factors drive sport and exercise engagement?'. For the latter example, AA is perfectly suited since it allows to capture contextual influences (such as nearby green spaces, social interactions) and therewith enables to take into account individual perceptions of the context when researching sport and exercise. Moreover, AA enables to research the congruency of objective to subjective perception. Researchers can make use of this, if interested in exercise addiction for example, to compare the subjectively perceived amount or intensity of sport and exercise with objectively measured parameters such as duration or energy expenditure of sport and exercise in real life.

Conclusion

In sum, the ongoing development of innovative technical features in AA provides numerous possibilities for PA research, especially the chance to tackle within-subject antecedents and consequences of PA as they unfold in everyday life. In therapy settings, very close monitoring, communication, and interaction between physicians, therapists, and patients are already feasible and the integration of interventions into the everyday life of patients comes with the large advantage of independence of sparse appointments. In future, in-depth insights about the drivers of PA can inform within-subject real-life interventions. Such interventions may be able to increase PA levels sustainably, for example by triggering at the time and in the place where needed. Therewith, one critical issue of humankind in the 21st century, namely the high levels of physical inactivity, which comes with a high individual and societal health burden, may be tackled.

REFERENCES

Aarts, H., & Dijksterhuis, A. (2000). Habits as knowledge structures: Automaticity in goal-directed behavior. *Journal of Personality and Social Psychology*, *78*(1), 53–63. https://doi.org/10.1037//0022-3514.78.1.53

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

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Althoff, T., Sosič, R., Hicks, J. L., King, A. C., Delp, S. L., & Leskovec, J. (2017). Large-scale physical activity data reveal worldwide activity inequality. *Nature*, *547*(7663), 336–339. https://doi.org/10.1038/nature23018

Armitage, C. J., & Conner, M. (2001). Efficacy of the Theory of Planned Behaviour: A meta-analytic review. *British Journal of Social Psychology*, *40*(4), 471–499. https://doi.org/10.1348/014466601164939

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. *Prentice-Hall series in social learning theory*. Prentice-Hall.

Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin*, *126*(6), 925–945. https://doi.org/10.1037//0033-2909.126.6.925

Ben-Zeev, D., Scherer, E. A., Wang, R., Xie, H., & Campbell, A. T. (2015). Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health. *Psychiatric Rehabilitation Journal*, *38*(3), 218–226. https://doi.org/10.1037/prj0000130

713 Boubekri, M., Cheung, I. N., Reid, K. J., Wang, C. H., & Zee, P. C. (2014). Impact of
714 windows and daylight exposure on overall health and sleep quality of office workers: A
715 case-control pilot study. *Journal of Clinical Sleep Medicine*, *10*(6), 603–611.
716 https://doi.org/10.5664/jcsm.3780

717 Bowden, A., Fox-Rushby, J. A., Nyandieka, L., & Wanjau, J. (2002). Methods for pre-testing
718 and piloting survey questions: Illustrations from the KENQOL survey of health-related
719 quality of life. *Health Policy and Planning*, *17*(3), 322–330.
720 https://doi.org/10.1093/heapol/17.3.322

721 Brand, R., & Ekkekakis, P. (2018). Affective–Reflective Theory of physical inactivity and
722 exercise. *German Journal of Exercise and Sport Research*, *48*(1), 48–58.
723 https://doi.org/10.1007/s12662-017-0477-9

724 Brand, R., & Schweizer, G. (2015). Going to the gym or to the movies? Situated decisions as
725 a functional link connecting automatic and reflective evaluations of exercise with
726 exercising behavior. *Journal of Sport & Exercise Psychology*, *37*(1), 63–73.
727 https://doi.org/10.1123/ jsep.2014–0018.

728 Brondeel, R., Pannier, B., & Chaix, B. (2015). Using GPS, GIS, and Accelerometer Data to
729 Predict Transportation Modes. *Medicine and Science in Sports and Exercise*, *47*(12),
730 2669–2675. https://doi.org/10.1249/MSS.0000000000000704

731 Brose, A., Schmiedek, F., Gerstorf, D., & Voelkle, M. C. (2019). The measurement of within-
732 person affect variation. *Emotion.* Advance online publication.
733 https://doi.org/10.1037/emo0000583

734 Brosschot, J. F., van Dijk, E., & Thayer, J. F. (2007). Daily worry is related to low heart rate
735 variability during waking and the subsequent nocturnal sleep period. *International Journal*
736 *of Psychophysiology*, *63*(1), 39–47. https://doi.org/10.1016/j.ijpsycho.2006.07.016

737 Bruening, M., van Woerden, I., Todd, M., Brennhofer, S., Laska, M. N., & Dunton, G. (2016).
738 A Mobile Ecological Momentary Assessment Tool (devilSPARC) for Nutrition and

Physical Activity Behaviors in College Students: A Validation Study. *Journal of Medical Internet Research*, *18*(7), Article e209. https://doi.org/10.2196/jmir.5969

Calitri, R., Lowe, R., Eves, F. F., & Bennett, P. (2009). Associations between visual attention, implicit and explicit attitude and behaviour for physical activity. *Psychology & Health*, *24*(9), 1105–1123. https://doi.org/10.1080/08870440802245306

Carver, C. S., & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality–social, clinical, and health psychology. *Psychological Bulletin*, *92*(1), 111–135. https://doi.org/10.1037/0033-2909.92.1.111

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: Definitions and distinctions for health-related research. *Public Health Reports*, *100*(2), 126–131. https://www.ncbi.nlm.nih.gov/pubmed/3920711

Choi, S., Kim, S., Yang, J. S., Lee, J. H., Joo, C., & Jung, H. I. (2014). Real-time measurement of human salivary cortisol for the assessment of psychological stress using a smartphone. *Sensing and Bio-Sensing Research*, *2*(2), 8-11. https://doi.org/10.1016/j.sbsr.2014.08.001

Clemes, S. A., & Deans, N. K. (2012). Presence and duration of reactivity to pedometers in adults. *Medicine and Science in Sports and Exercise*, *44*(6), 1097–1101. https://doi.org/10.1249/MSS.0b013e318242a377

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

Clemes, S. A., & Parker, R. A. (2009). Increasing our understanding of reactivity to pedometers in adults. *Medicine and Science in Sports and Exercise*, *41*(3), 674–680. https://doi.org/10.1249/MSS.0b013e31818cae32

Conner, M. T., Perugini, M., O'Gorman, R., Ayres, K., & Prestwich, A. (2007). Relations between implicit and explicit measures of attitudes and measures of behavior: Evidence of moderation by individual difference variables. *Personality & Social Psychology Bulletin*, *33*(12), 1727–1740. https://doi.org/10.1177/0146167207309194

Conner, T. S., & Lehman, B. J. (2012). Getting started: Launching a study in daily life. In Mehl, M. R., & Conner T. S. (Eds.), *Handbook of research methods for studying daily life* (pp. 89–107). Guilford Press.

Conroy, D. E., & Berry, T. R. (2017). Automatic Affective Evaluations of Physical Activity. *Exercise and Sport Sciences Reviews*, *45*(4), 230–237. https://doi.org/10.1249/JES.0000000000000120

Conroy, D. E., Maher, J. P., Elavsky, S., Hyde, A. L., & Doerksen, S. E. (2013). Sedentary behavior as a daily process regulated by habits and intentions. *Health Psychology, 32*(11), 1149–1157. https://doi.org/10.1037/a0031629

Cooke, R., & Sheeran, P. (2004). Moderation of cognition-intention and cognition-behaviour relations: A meta-analysis of properties of variables from the theory of planned behaviour. *British Journal of Social Psychology*, *43*(2), 159–186. https://doi.org/10.1348/0144666041501688

Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably?. *Personality & Social Psychology Bulletin*, *32*(7), 917–929. https://doi.org/10.1177/0146167206287721

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. https://doi.org/10.1037/h0040957

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

Dorn, H., Törnros, T., Reichert, M., Salize, H. J., Tost, H., Ebner-Priemer, U. W., Meyer-Lindenberg, A., & Zipf, A. (2015). Incorporating land use in a spatiotemporal trigger for ecological momentary assessments. *Journal of Geographic Information System 1*, 13–1165. https://doi.org/10.1552/giscience2015s113

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., Dzubur, E., & Intille, S. (2016). Feasibility and Performance Test of a Real-Time Sensor-Informed Context-Sensitive Ecological Momentary Assessment to Capture Physical Activity. *Journal of Medical Internet Research*, *18*(6), Article e106. https://doi.org/10.2196/jmir.5398

Dunton, G. F., Dzubur, E., Kawabata, K., Yanez, B., Bo, B., & Intille, S. (2014). Development of a smartphone application to measure physical activity using sensor-assisted self-report. *Frontiers in Public Health*, *2*Article 12. https://doi.org/10.3389/fpubh.2014.00012

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

Dunton, G. F., Liao, Y., Dzubur, E., Leventhal, A. M., Huh, J., Gruenewald, T., Margolin, G., Koprowski, C., Tate, E., & Intille, S. (2015). Investigating within-day and longitudinal effects of maternal stress on children's physical activity, dietary intake, and body composition: Protocol for the MATCH study. *Contemporary Clinical Trials*, *43*, 142–154. https://doi.org/10.1016/j.cct.2015.05.007

Dunton, G. F., Whalen, C. K., Jamner, L. D., & Floro, J. N. (2007). Mapping the social and physical contexts of physical activity across adolescence using ecological momentary assessment. *Annals of Behavioral Medicine*, *34*(2), 144–153. https://doi.org/10.1007/BF02872669

Dunton, G. F., Whalen, C. K., Jamner, L. D., Henker, B., & Floro, J. N. (2005). Using ecologic momentary assessment to measure physical activity during adolescence. *American Journal of Preventive Medicine*, *29*(4), 281–287. https://doi.org/10.1016/j.amepre.2005.07.020

Ebner-Priemer, U. W., Koudela, S., Mutz, G., & Kanning, M. (2013). Interactive multimodal ambulatory monitoring to investigate the association between physical activity and mood. *Frontiers in Movement Science and Sport Psychology*, *3*, 596. https://doi.org/10.3389/fpsyg.2012.00596

Ebner-Priemer, U. W., & Trull, T. J. (2009). Ambulatory Assessment. *European Psychologist*, *14*(2), 109–119. https://doi.org/10.1027/1016-9040.14.2.109

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, *5*(2), 155–174. https://doi.org/10.1037/1082-989X.5.2.155

Ekelund, U., Brown, W. J., Steene-Johannessen, J., Fagerland, M. W., Owen, N., Powell, K. E., Bauman, A. E., & Lee, I. M. (2019). Do the associations of sedentary behaviour with cardiovascular disease mortality and cancer mortality differ by physical activity level? A systematic review and harmonised meta-analysis of data from 850 060 participants. *British Journal of Sports Medicine*, *53*(14), 886. https://doi.org/10.1136/bjsports-2017-098963

Ekkekakis, P. (2017). People have feelings! Exercise psychology in paradigmatic transition. *Current Opinion in Psychology*, *16*, 84–88. https://doi.org/10.1016/j.copsyc.2017.03.018.

Exler, A., Klebsattel, C., Schankin, A., Riedel, T., Beigl, M., Reichert, M., Santangelo, P., & Ebner-Priemer, U. W. (2015, September 7-11). *Leveraging Smartwatches for Unobtrusive Mobile Ambulatory Mood Assessment.* [Paper presentation]. ACM International Joint Conference on Pervasive and Ubiquitous Computing, Osaka, Japan. https://doi.org/10.1145/2800835.2800960

Fahrenberg, J., Myrtek, M., Pawlik, K., & Perrez, M. (2007). Ambulantes Assessment - Verhalten im Alltagskontext erfassen [Ambulatory Assessment - Assessing behaviour in the context of daily life]. *Psychologische Rundschau*, *58*(1), 12–23. https://doi.org/10.1026/0033-3042.58.1.12

Fong, K. C., Hart, J. E., & James, P. (2018). A Review of Epidemiologic Studies on Greenness and Health: Updated Literature Through 2017. *Current Environmental Health Reports*, *5*(1), 77–87. https://doi.org/10.1007/s40572-018-0179-y

Friese, M., & Hofmann, W. (2009). Control me or I will control you: Impulses, trait self-control, and the guidance of behavior. *Journal of Research in Personality*, *43*(5), 795–805. https://doi.org/10.1016/j.jrp.2009.07.004

Gaston, K. J., Davies, T. W., Bennie, J., & Hopkins, J. (2012). Reducing the ecological consequences of night-time light pollution: Options and developments. *The Journal of Applied Ecology*, *49*(6), 1256–1266. https://doi.org/10.1111/j.1365-2664.2012.02212.x

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*(1), 72–91. https://doi.org/10.1037/a0032138

Giurgiu, M., Koch, E. D., Ottenbacher, J., Plotnikoff, R. C., Ebner-Priemer, U. W., & Reichert, M. (2019). Sedentary behavior in everyday life relates negatively to mood: An ambulatory assessment study. *Scandinavian Journal of Medicine & Science in Sports*, *29*(9), 1340–1351. https://doi.org/10.1111/sms.13448

Giurgiu, M., Koch, E. D., Plotnikoff, R. C., Ebner-Priemer, U. W., & Reichert, M. (2019). Breaking Up Sedentary Behavior Optimally to Enhance Mood. *Medicine and Science in Sports and Exercise, 52*(2), 457–465. https://doi.org/10.1249/MSS.0000000000002132

Hamaker, E. L. (2012). Why researchers should think “within-person”: A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of Research Methods for Studying Daily Life* (pp. 43–61). Guilford Press.

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*(1), 1–39. https://doi.org/10.1348/135910709X466063

Hofmann, W., Friese, M., & Roefs, A. (2009). Three ways to resist temptation: The independent contributions of executive attention, inhibitory control, and affect regulation to the impulse control of eating behavior. *Journal of Experimental Social Psychology*, *45*(2), 431–435. https://doi.org/10.1016/j.jesp.2008.09.013

Hynynen, E., Uusitalo, A., Konttinen, N., & Rusko, H. (2006). Heart rate variability during night sleep and after awakening in overtrained athletes. *Medicine and Science in Sports and Exercise*, *38*(2), 313–317. https://doi.org/10.1249/01.mss.0000184631.27641.b5

Inauen, J., Shrout, P. E., Bolger, N., Stadler, G., & Scholz, U. (2016). Mind the Gap? An Intensive Longitudinal Study of Between-Person and Within-Person Intention-Behavior Relations. *Annals of Behavioral Medicine : A Publication of the Society of Behavioral Medicine*, *50*(4), 516–522. https://doi.org/10.1007/s12160-016-9776-x

Intille, S. (2007). Technological innovations enabling automatic, context-sensitive ecological momentary assessment. In A. A. Stone (Ed.), *The science of real-time data capture: Self-reports in health research* (pp. 308–337). Oxford University Press.

Jansen, F. M., Ettema, D. F., Kamphuis, C.B.M., Pierik, F. H., & Dijst, M. J. (2017). How do type and size of natural environments relate to physical activity behavior? *Health & Place*, *46*, 73–81. https://doi.org/10.1016/j.healthplace.2017.05.005

Jekauc, D. (2009). *Entwicklung und Stabilität der körperlich-sportlichen Aktivität im mittleren Erwachsenenalter: eine prospektive Längsschnittstudie* [Development and stability of physical-sporting activities in middle age: a prospective longitudinal study]. Logos Verlag.

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

Jekauc, D., Völkle, M., Wagner, M. O., Mess, F., Reiner, M., & Renner, B. (2015). Prediction of attendance at fitness center: A comparison between the theory of planned behavior, the social cognitive theory, and the physical activity maintenance theory. *Frontiers in Psychology*, *6* Article121. https://doi.org/10.3389/fpsyg.2015.00121

Jones, A., Tiplady, B., Houben, K., Nederkoorn, C., & Field, M. (2018). Do daily fluctuations in inhibitory control predict alcohol consumption? An ecological momentary assessment study. *Psychopharmacology*, *235*(5), 1487–1496. https://doi.org/10.1007/s00213-018-4860-5

Kamarck, T. W., Schwartz, J. E., Janicki, D. L., Shiffman, S., & Raynor, D. A. (2003). Correspondence between laboratory and ambulatory measures of cardiovascular reactivity: A multilevel modeling approach. *Psychophysiology*, *40*(5), 675–683. https://doi.org/10.1111/1469-8986.00069

Kanning, M., Ebner-Priemer, U. W., & Schlicht, W. (2015). Using activity triggered e-diaries to reveal the associations between physical activity and affective states in older adult's daily living. *The International Journal of Behavioral Nutrition and Physical Activity*, *12*, Article 111. https://doi.org/10.1186/s12966-015-0272-7

Kiene, S. M., Tennen, H., & Armeli, S. (2008). Today I'll use a condom, but who knows about tomorrow: A daily process study of variability in predictors of condom use. *Health Psychology, 27*(4), 463–472. https://doi.org/10.1037/0278-6133.27.4.463

Kirchner, T. R., & Shiffman, S. (2016). Spatio-temporal determinants of mental health and well-being: Advances in geographically-explicit ecological momentary assessment (GEMA). *Social Psychiatry and Psychiatric Epidemiology*, *51*(9), 1211–1223. https://doi.org/10.1007/s00127-016-1277-5

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*, *34S*, 1220–1228. https://doi.org/10.1037/hea0000305

Knell, G., Gabriel, K. P., Businelle, M. S., Shuval, K., Wetter, D. W., & Kendzor, D. E. (2017). Ecological Momentary Assessment of Physical Activity: Validation Study. *Journal of Medical Internet Research*, *19*(7), Article e253. https://doi.org/10.2196/jmir.7602

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*Article 268. https://doi.org/10.3389/fpsyg.2018.00268

Kop, W. J., Lyden, A., Berlin, A. A., Ambrose, K., Olsen, C., Gracely, R. H., Williams, D. A., & Clauw, D. J. (2005). Ambulatory monitoring of physical activity and symptoms in fibromyalgia and chronic fatigue syndrome. *Arthritis and Rheumatism*, *52*(1), 296–303. https://doi.org/10.1002/art.20779

Korotitsch, W. J., & Nelson-Gray, R. O. (1999). An overview of self-monitoring research in assessment and treatment. *Psychological Assessment*, *11*(4), 415–425. https://doi.org/10.1037/1040-3590.11.4.415

Kwan, M. P. (2012). The Uncertain Geographic Context Problem. *Annals of the Association of American Geographers*, *102*(5), 958–968. https://doi.org/10.1080/00045608.2012.687349

Lammers, M. O., Brainard, R. E., Au, W. W. L., Mooney, T. A., & Wong, K. B. (2008). An ecological acoustic recorder (EAR) for long-term monitoring of biological and anthropogenic sounds on coral reefs and other marine habitats. *The Journal of the Acoustical Society of America*, *123*(3), 1720–1728. https://doi.org/10.1121/1.2836780

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

Liao, Y., Skelton, K., Dunton, G., & Bruening, M. (2016). A Systematic Review of Methods and Procedures Used in Ecological Momentary Assessments of Diet and Physical Activity Research in Youth: An Adapted STROBE Checklist for Reporting EMA Studies (CREMAS). *Journal of Medical Internet Research*, *18*(6), Article e151. https://doi.org/10.2196/jmir.4954.

Ludyga, S., Pühse, U., Lucchi, S., Marti, J., & Gerber, M. (2019). Immediate and sustained effects of intermittent exercise on inhibitory control and task-related heart rate variability in adolescents. *Journal of Science and Medicine in Sport*, *22*(1), 96–100. https://doi.org/10.1016/j.jsams.2018.05.027

Maher, J. P., Dzubur, E., Huh, J., Intille, S., & Dunton, G. F. (2016). Within-Day Time-Varying Associations Between Behavioral Cognitions and Physical Activity in Adults. *Journal of Sport & Exercise Psychology*, *38*(4), 423–434. https://doi.org/10.1123/jsep.2016-0058

Maher, J. P., Rebar, A. L., & Dunton, G. F. (2018). Ecological Momentary Assessment Is a Feasible and Valid Methodological Tool to Measure Older Adults' Physical Activity and Sedentary Behavior. *Frontiers in Psychology*, *9*, 1485. https://doi.org/10.3389/fpsyg.2018.01485

Maher, J. P., Rhodes, R. E., Dzubur, E., Huh, J., Intille, S., & Dunton, G. F. (2017). Momentary assessment of physical activity intention-behavior coupling in adults. *Translational Behavioral Medicine*, *7*(4), 709–718. https://doi.org/10.1007/s13142-017-0472-6

Mayer, J. S., Hees, K., Medda, J., Grimm, O., Asherson, P., Bellina, M., Colla, M., Ibáñez, P., Koch, E., Martinez-Nicolas, A., Muntaner-Mas, A., Rommel, A., Rommelse, N., Ruiter, S. de, Ebner-Priemer, U. W., Kieser, M., Ortega, F. B., Thome, J., Buitelaar, J. K., Kuntsi, J., . . . Freitag, C. M. (2018). Bright light therapy versus physical exercise to prevent co-morbid depression and obesity in adolescents and young adults with attention-deficit / hyperactivity disorder: Study protocol for a randomized controlled trial. *Trials*, *19*(1), 140. https://doi.org/10.1186/s13063-017-2426-1

McDonald, R. P. (1999). Test theory: A unified treatment. *Lawrence Earlbaum Associates*, 142–145.

Mehl, M. R., & Conner, T. S. (Eds.). (2012). *Handbook of Research Methods for Studying Daily Life*. Guilford Press.

Mobasheri, A., Sun, Y., Loos, L., & Ali, L. A. (2017). *Are Crowdsourced Datasets Suitable for Specialized Routing Services? Case Study of OpenStreetMap for Routing of People with Limited Mobility*. *Sustainability, 9*(6), Article 997. https://doi.org/10.3390/su9060997

Molenaar, P. C. M. (2004). A Manifesto on Psychology as Idiographic Science: Bringing the Person Back Into Scientific Psychology, This Time Forever. *Measurement: Interdisciplinary Research & Perspective*, *2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1

Morgenstern, J., Kuerbis, A., & Muench, F. (2014). Ecological Momentary Assessment and Alcohol Use Disorder Treatment. *Alcohol Research : Current Reviews*, *36*(1), 101–110.

Myin-Germeys, I., Klippel, A., Steinhart, H., & Reininghaus, U. (2016). Ecological momentary interventions in psychiatry. *Current Opinion in Psychiatry*, *29*(4), 258–263. https://doi.org/10.1097/YCO.0000000000000255

Nezlek, J. B. (2017). A practical guide to understanding reliability in studies of within-person variability. *Journal of Research in Personality*, *69*, 149–155. https://doi.org/10.1016/j.jrp.2016.06.020

Niermann, C. Y. N., 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*, Article 677. https://doi.org/10.3389/fpsyg.2016.00677

Nigg, C. R., Fuchs, R., Gerber, M., Jekauc, D., Koch, T., Krell-Roesch, J., Lippke, S., Mnich, C., Novak, B., Ju, Q., Sattler, M. C., Schmidt, S. C. E., van Poppel, M., Reimers, A. K., Wagner, P., Woods, C. & Woll, A. (2020). Assessing Physical Activity through Questionnaires – A Consensus of Best Practices and Future Directions. *Psychology of Sport & Exercise.*

Noar, S. M., & Zimmerman, R. S. (2005). Health Behavior Theory and cumulative knowledge regarding health behaviors: Are we moving in the right direction? *Health Education Research*, *20*(3), 275–290. https://doi.org/10.1093/her/cyg113

Pettee Gabriel, K. K., Morrow, J. R., & Woolsey, A. L. T. (2012). Framework for physical activity as a complex and multidimensional behavior. *Journal of Physical Activity & Health*, *9*(1), 11-8. https://doi.org/10.1123/jpah.9.s1.s11

Poh, M., Loddenkemper, T., Swenson, N. C., Goyal, S., Madsen, J. R., & Picard, R. W. W. (2010, August 31-September 4). *Continuous monitoring of electrodermal activity during epileptic seizures using a wearable sensor*. [Paper presentation]. 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina.

Poh, M., Swenson, N. C., & Picard, R. W. (2010). A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *Transactions on Bio-Medical Engineering*, *57*(5), 1243–1252. https://doi.org/10.1109/TBME.2009.2038487

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

Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology*, *51*(3), 390–395. https://doi.org/10.1037//0022-006x.51.3.390

Quinn, J. W., Mooney, S. J., Sheehan, D. M., Teitler, J. O., Neckerman, K. M., Kaufman, T. K., Lovasi, G. S., Bader, M. D., & Rundle, A. G. (2016). Neighborhood physical disorder in New York City. *Journal of Maps*, *12*(1), 53–60. https://doi.org/10.1080/17445647.2014.978910

Reichert, M., Braun, U., Lautenbach, S., Zipf, A., Ebner-Priemer, U. W., Tost, H., & Meyer-Lindenberg, A. (2020). Studying the impact of built environments on human mental health in everyday life: methodological developments, state-of-the-art and technological frontiers. *Current Opinion in Psychology*, *32*, 158–164. https://doi.org/10.1016/j.copsyc.2019.08.026

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*Article 918. https://doi.org/10.3389/fpsyg.2016.00918

Reis, H. T. (2012). Why researchers should think "real-world": A conceptual rationale. In Mehl, M. R., & Conner T. S. (Eds.), *Handbook of research methods for studying daily life e* (pp. 3–21). Guilford Press.

Revelle, W., & Wilt, J. (2019). Analyzing dynamic data: A tutorial. *Personality and Individual Differences*, *136*, 38–51. https://doi.org/10.1016/j.paid.2017.08.020

Rhemtulla, M., van Bork, R., & Borsboom, D. (2019). Worse than measurement error: Consequences of inappropriate latent variable measurement models. *Psychological Methods.* Advance online publication. https://doi.org/10.1037/met0000220

Rhodes, R. E., & Bruijn, G. J. de (2013). How big is the physical activity intention-behaviour gap? A meta-analysis using the action control framework. *British Journal of Health Psychology*, *18*(2), 296–309. https://doi.org/10.1111/bjhp.12032

Robinson, W. S. (1950). Ecological Correlations and the Behavior of Individuals. *American Sociological Review*, *15*(3), 351–357. https://doi.org/10.1093/ije/dyr081

Romanzini, C. L. P., Romanzini, M., Batista, M. B., Barbosa, C. C. L., Shigaki, G. B., Dunton, G., Mason, T., & Ronque, E. R. V. (2019). Methodology Used in Ecological Momentary Assessment Studies About Sedentary Behavior in Children, Adolescents, and Adults: Systematic Review Using the Checklist for Reporting Ecological Momentary Assessment Studies. *J Med Internet Res*, *21*(5), Article e11967. https://doi.org/10.2196/11967.

Rzotkiewicz, A., Pearson, A. L., Dougherty, B. V., Shortridge, A., & Wilson, N. (2018). Systematic review of the use of Google Street View in health research: Major themes, strengths, weaknesses and possibilities for future research. *Health & Place*, *52*, 240–246. https://doi.org/10.1016/j.healthplace.2018.07.001

Salles, G. F., Cardoso, C. R. L., & Muxfeldt, E. S. (2008). Prognostic influence of office and ambulatory blood pressures in resistant hypertension. *Archives of Internal Medicine*, *168*(21), 2340–2346. https://doi.org/10.1001/archinte.168.21.2340

Schlicht, W., Ebner-Priemer, U. W., & Kanning, M. (2013). Ecological momentary assessment and intervention in physical activity and well-being: Affective reactions, social-cognitive factors, and behaviors as determinants of physical activity and exercise. *Frontiers in Psychology*, *4*, Article 916. https://doi.org/10.3389/fpsyg.2013.00916

Schmiedek, F., & Neubauer, A. B. (2019). Experiments in the Wild: Introducing the Within-Person Encouragement Design. *Multivariate Behavioral Research*, 1–21. https://doi.org/10.1080/00273171.2019.1627660

Scott, J. J., Morgan, P. J., Plotnikoff, R. C., Trost, S. G., & Lubans, D. R. (2014). Adolescent pedometer protocols: Examining reactivity, tampering and participants' perceptions. *Journal of Sports Sciences*, *32*(2), 183–190. https://doi.org/10.1080/02640414.2013.815361

Sheeran, P. (2002). Intention—Behavior Relations: A Conceptual and Empirical Review. *European Review of Social Psychology*, *12*(1), 1–36. https://doi.org/10.1080/14792772143000003

Sheeran, P., & Abraham, C. (2003). Mediator of moderators: Temporal stability of intention and the intention-behavior relation. *Personality & Social Psychology Bulletin*, *29*(2), 205–215. https://doi.org/10.1177/0146167202239046

Sheeran, P., Webb, T. L., & Gollwitzer, P. M. (2005). The interplay between goal intentions and implementation intentions. *Personality & Social Psychology Bulletin*, *31*(1), 87–98. https://doi.org/10.1177/0146167204271308

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological Momentary Assessment. *Annual Review of Clinical Psychology*, *4*(1), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

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*, Article 515. https://doi.org/10.1186/s12889-016-3172-0

Society for Ambulatory Assessment. (n.d.). http://ambulatory-assessment.org/

Spenkelink, C. D., Hutten, M. M. R., Hermens, H. J., & Greitemann, B. O. L. (2002). Assessment of activities of daily living with an ambulatory monitoring system: A comparative study in patients with chronic low back pain and nonsymptomatic controls. *Clinical Rehabilitation*, *16*(1), 16–26. https://doi.org/10.1191/0269215502cr463oa

Stewart, O. T., Moudon, A. V., Fesinmeyer, M. D., Zhou, C., & Saelens, B. E. (2016). The association between park visitation and physical activity measured with accelerometer, GPS, and travel diary. *Health & Place*, *38*, 82–88. https://doi.org/10.1016/j.healthplace.2016.01.004

Stewart, O. T., Moudon, A. V., Littman, A. J., Seto, E., & Saelens, B. E. (2018). Why neighborhood park proximity is not associated with total physical activity. *Health & Place*, *52*, 163–169. https://doi.org/10.1016/j.healthplace.2018.05.011

Stone, A. A., & Shiffman, S. (1994). Ecological Momentary Assessment (Ema) in Behavioral Medicine. *Annals of Behavioral Medicine : A Publication of the Society of Behavioral Medicine*, *16*(3), 199–202. https://doi.org/10.1093/abm/16.3.199

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review. 8*(3), 220–247. https://doi.org/10.1207/s15327957pspr0803_1

Törnros, T., Dorn, H., Reichert, M., Ebner-Priemer, U. W., Salize, H. J., Tost, H., Meyer-Lindenberg, A., & Zipf, A. (2016). A comparison of temporal and location-based sampling strategies for global positioning system-triggered electronic diaries. *Geospatial Health*, *11*(3), 335–341. https://doi.org/10.4081/gh.2016.473

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.* https://doi.org/10.1038/s41593-019-0451-y

Troiano, R. P. (2005). A timely meeting: Objective measurement of physical activity. *Medicine and Science in Sports and Exercise*, *37*(11), 487-489. https://doi.org/10.1249/01.mss.0000185473.32846.c3

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

Trull, T. J., Ebner-Priemer, U. W., & Hall, M. (2019). Ambulatory Assessment in Psychopathology Research: A Review of Recommended Reporting Guidelines and Current Practices. *Journal of Abnormal Psychology*, (in press).

Voelkle, M. C., Brose, A., Schmiedek, F., & Lindenberger, U. (2014). Toward a Unified Framework for the Study of Between-Person and Within-Person Structures: Building a

Bridge Between Two Research Paradigms. *Multivariate Behavioral Research*, *49*(3), 193–213. https://doi.org/10.1080/00273171.2014.889593

Von Haaren-Mack, B., Bussmann, J. B.J., & Ebner-Priemer, U. W. (in press). Physical activity monitoring. In M. Robbins & K. Sweeny (Eds.), *Encyclopedia of Health Psychology*, New York: Wiley.

Wang, H. C., & Lee, A. R. (2015). Recent developments in blood glucose sensors. *Journal of Food and Drug Analysis*, *23*(2), 191–200. https://doi.org/10.1016/j.jfda.2014.12.001

Webb, T. L., & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin*, *132*(2), 249–268. https://doi.org/10.1037/0033-2909.132.2.249

Wen, C. K. F., Liao, Y., Maher, J. P., Huh, J., Belcher, B. R., Dzubur, E., & Dunton, G. F. (2018). Relationships among affective states, physical activity, and sedentary behavior in children: Moderation by perceived stress. *Health Psychology, 37*(10), 904–914. https://doi.org/10.1037/hea0000639

Wilhelm, F. H., Pfaltz, M. C., & Grossman, P. (2006). Continuous electronic data capture of physiology, behavior and experience in real life: towards ecological momentary assessment of emotion. *Interacting with Computers*, *18*(2), 171–186. https://doi.org/10.1016/j.intcom.2005.07.001

Wilhelm, F. H., & Roth, W. T. (1998). Taking the laboratory to the skies: Ambulatory assessment of self-report, autonomic, and respiratory responses in flying phobia. *Psychophysiology*, *35*(5), 596–606. https://doi.org/10.1017/S0048577298970196

Wilhelm, P., Perrez, M., & Pawlik, K. (2012). Conducting research in daily life: A historical review. In Mehl, M. R. & Conner, T. S. (Ed.), *Handbook of Research Methods for Studying Daily Life*, (pp. 62–86). Guilford Publications.

Wilhelm, P., & Schoebi, D. (2007). Assessing Mood in Daily Life: Structural Validity, Sensitvity to Change, and Reliability of a Short-Scale to Measure Three Basic

Dimensions of Mood. *European Journal of Psychological Assessment*, *23*(4), 258–267. https://doi.org/10.1027/1015-5759.23.4.258

Wood, W., Quinn, J. M., & Kashy, D. A. (2002). Habits in everyday life: Thought, emotion, and action. *Journal of Personality and Social Psychology*, *83*(6), 1281–1297. https://doi.org/10.1037/0022-3514.83.6.1281

Wunsch, K., Meier, M., Ueberholz, L., Strahler, J., & Kasten, N. (2019). Acute psychosocial stress and working memory performance: the potential of physical activity to modulate cognitive functions in children. *BMC Pediatri*cs *19*(271). https://doi.org/10.1186/s12887-019-1637-x

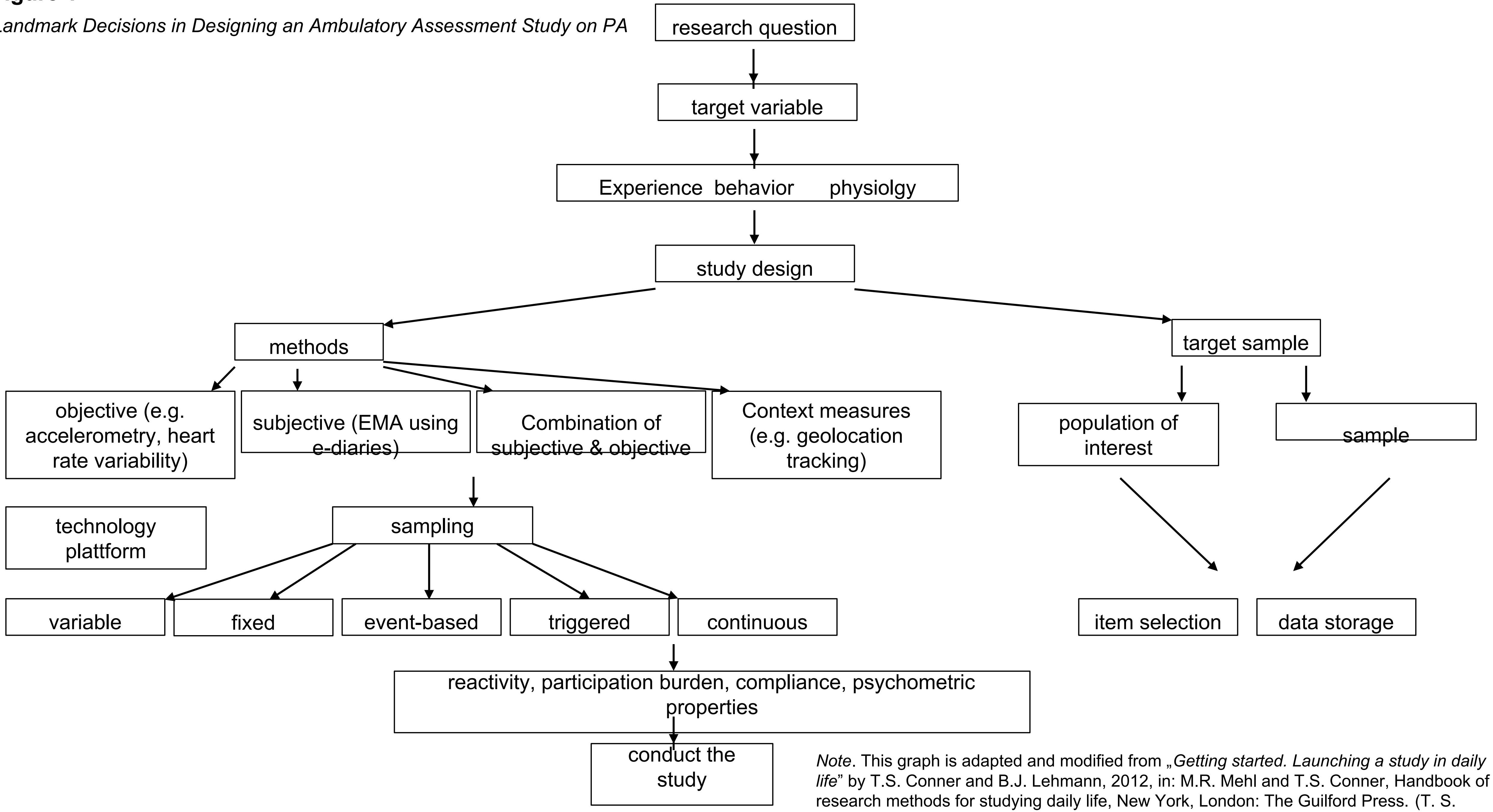
Xu, C., Li, S., Liu, G., Zhang, Y., Miluzzo, E., Chen, Y. F., Li, J., & Firner, B. (2013). Crowd++: Unsupervised Speaker Count with Smartphones. In Mattern, F., & Santini, S. (Eds.), *UbiComp'13: Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp 43-52). Association for Computing Machinery. https://doi.org/10.1145/2493432.2493435

Zawadzki, M. J., Smyth, J. M., Sliwinski, M. J., Ruiz, J. M., & Gerin, W. (2017). Revisiting the lack of association between affect and physiology: Contrasting between-person and within-person analyses. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, *36*(8), 811–818. https://doi.org/10.1037/hea0000466

Zipf, A., Mobasheri, A., Rousell, A., & Hahmann, S. (2016). Crowdsourcing for individual needs – the case of routing and navigation for mobility-impaired persons. In Capineri, C et al (Eds.), *European Handbook of Crowdsourced Geographic Information* (pp. 325-337). Ubiquity Press. https://doi.org/10.5334/bax.x

Figure 1

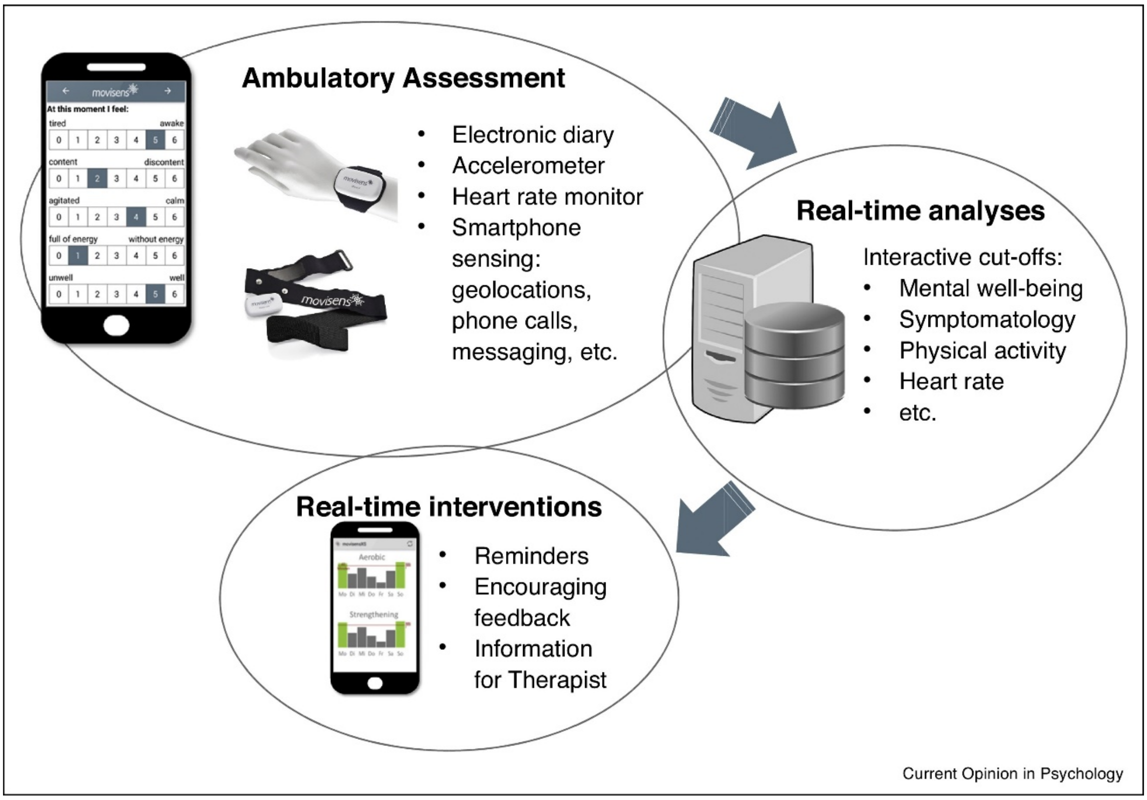
Landmark Decisions in Designing an Ambulatory Assessment Study on PA



Note. This graph is adapted and modified from „*Getting started. Launching a study in daily life*" by T.S. Conner and B.J. Lehmann, 2012, in: M.R. Mehl and T.S. Conner, Handbook of research methods for studying daily life, New York, London: The Guilford Press. (T. S. Conner & Lehman, 2012).

Figure 2

Ambulatory Assessment in combination with real-time analyses and interventions.



Note. Ambulatory Assessment can be combined with real-time analyses (MovisensXS platform, https://xs.movisens.com) to enable triggered real-time e-diary prompts and real-life interventions. The graph is adapted from the Journal Current Opinion in Psychology (Reichert et al., 2020) and reprinted with permission from Reichert and colleagues.

Conflicts of Interest: The authors declare no competing interests. UE-P works as a consultant for Boehringer Ingelheim.