Stress and Food Intake in Daily Life: Insights Based on a Novel Ecological Momentary Assessment Tool and an Advanced Data Analysis Approach

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

vorgelegt beim Fachbereich Psychologie und Sportwissenschaften

der Johann Wolfgang Goethe-Universität

in Frankfurt am Main

von

Alea Ruf

aus Frankfurt am Main

Frankfurt 2022

Vom Fachbereich Psychologie und Sportwissenschaften

der Johann Wolfgang Goethe-Universität

als Dissertation angenommen.

Dekanin:	Prof. Dr. Sonja Rohrmann
Gutachterinnen:	Prof. Dr. Monika Knopf
	Prof. Dr. Sonja Rohrmann
Datum der Disputation:	03.05.2023

Content

Abstract

Stress influences health not only directly, but also indirectly through changes in health-related behaviours, such as diet. Research has shown that stress influences individuals' eating behaviour in different ways: Some increase, some decrease food intake, while others show no change. Identifying individuals at risk for stress-induced eating is essential for the development of tailored strategies for the prevention and treatment of overweight and obesity. The individual-difference model of stress-induced eating suggests that individual differences in the dietary response to stress are determined by differences in learning history, attitudes, or biology. Even though many studies have tried to identify person-characteristics that explain individual differences in the dietary response to stress, evidence remains inconclusive. Considering that eating is a repeated-occurrence health behaviour which is performed multiple times a day, Ecological Momentary Assessment (EMA) seems particularly promising to study the complex relationship between stress and food intake when and where it naturally occurs. Despite its potential, the number of studies applying EMA to assess the stress and eating relationship is limited. Furthermore, previous EMA studies show two limitations: (1) Actual food intake is not assessed and (2) inappropriate data analysis approaches are applied to semicontinuous outcomes. Therefore, the first aim of the present dissertation was to address the lack of an EMA tool that allows the assessment of stress and actual food intake by developing and evaluating the APPetite-mobile-app. Feasibility and usability of the APPetite-mobile-app as well as validity of the incorporated food record were empirically examined (Paper 1). Given the lack of an appropriate data analysis procedure, the second aim of the present dissertation was the introduction of a sophisticated statistical approach for semicontinuous data (Paper 2): Multilevel two-part modelling allows studying the influence of stress on the occurrence (i.e., *whether* individuals eat) as well as the amount of food intake (i.e., *how much* individuals eat) while accounting for the potential dependency between the two. Lastly, the novel EMA tool and the advanced data analysis procedure were integrated in

order to gain novel insights into individual differences in the dietary response to stress and thereby identify individuals at risk for stress-induced eating in daily life (Paper 3). Results of Paper 1 showed good feasibility and acceptable usability of the APPetite-mobile-app as well as validity of the incorporated food record. Findings of Paper 2 highlight that multilevel two-part models offer novel and distinct insights in terms of the occurrence and the amount of food intake and are therefore not only methodologically but also conceptually promising. Paper 3 provides first evidence that the dietary response to stress might not be as stable as yet assumed. Time-varying factors might moderate the relationship between stress and actual food intake. Therefore, an expansion of the individual-difference model is proposed which accounts for time-varying factors. Further EMA studies are needed to verify the expanded model and identify time-varying factors which influence the dietary response to stress. Beyond that, improvements in the dietary assessment are required in order to allow prolonged EMA periods as well as larger samples. The present dissertation contributes to the research on the stress and eating relationship as it overcomes limitations of previous EMA studies and yields novel insights into the relationship between stress and actual food intake in daily life. Not only identifying individuals at risk for stress-induced eating, but also the identification of situations with an increased risk for stress-induced eating appears to be important for the development of targeted strategies for the prevention and treatment of overweight and obesity.

1. General Introduction

1.1. Diet and Health

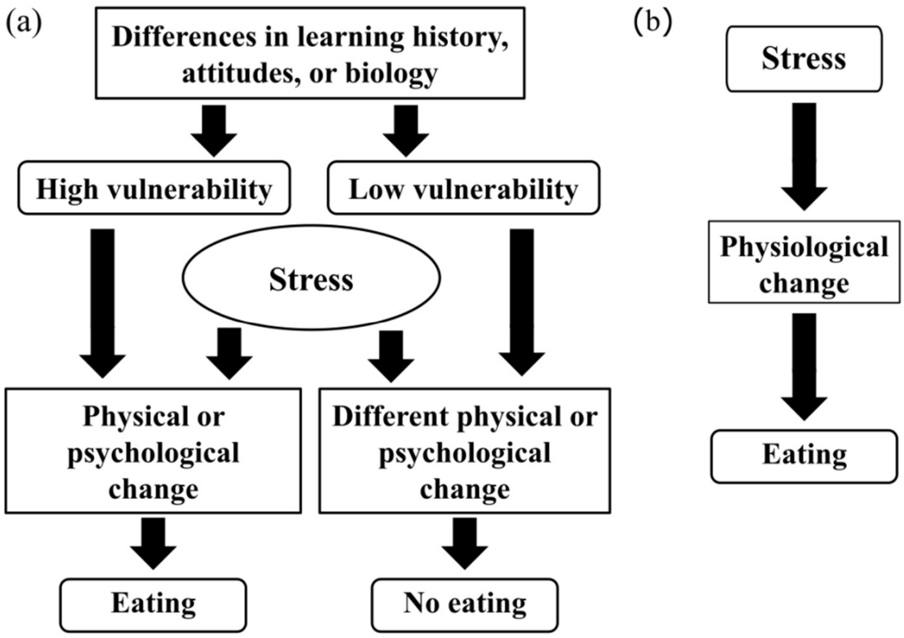
Diet is a key contributor to physical as well as mental health. While a healthy diet can protect human health, poor dietary habits can have adverse effects. For instance, Mediterranean dietary patterns were found to be a protective factor against coronary heart disease, while the consumption of trans-fatty acids was associated with a higher risk for coronary heart disease (Mente et al., 2009). In 2017, approximately 11 million deaths were associated with dietary risk factors (e.g., high intake of sodium) across 195 countries (Afshin et al., 2019). Though, not only the quality of food intake (i.e., types of food) but also the quantity of food intake (i.e., energy intake) influences health as it plays a central role in the regulation of body weight. Elevated levels of body mass index (BMI; i.e., overweight and obesity) are a major risk factor for non-communicable diseases, such as cardiovascular diseases and diabetes (World Health Organization, 2021). Despite this knowledge, the prevalence of obesity is rising globally (Swinburn et al., 2019). Even though the association between diet and mental health is not equally well understood, first evidence supports the presence of a direct link between diet and mental health (Adan et al., 2019). Beyond that, obesity is associated not only with an increased probability of somatic diseases but also of mental disorders, particularly depression (Kelly et al., 2011; Milaneschi et al., 2019; Rajan & Menon, 2017). These findings highlight the growing need to understand the “causes of the causes”. Only when factors and processes that underlie eating behaviour are understood, effective interventions targeting the prevention as well as the treatment of overweight and obesity can be developed.

1.2. Stress and Food Intake

It is well-known that not only physiological factors, such as hunger, influence eating behaviour. A large variety of factors shape eating, including environmental (e.g., price - Afshin et al., 2017), sociocultural (e.g., social norms - Higgs, 2015) and psychological (e.g., emotions - Evers et al., 2018) factors. One factor that stands out in this context due to its twofold impact on health is stress: Stress not only influences health directly, as such that elevated levels of stress are linked to various negative health outcomes, e.g., cardiovascular diseases (Kivimäki & Steptoe, 2018). Stress impacts health also indirectly through changes in health-related behaviours, such as diet (O'Connor et al., 2021). Substantial empirical evidence suggests that stress is associated with changes in food intake (for overviews, see Araiza & Lobel, 2018; Hill et al., 2021; Yau & Potenza, 2013). Furthermore, stress seems to play an important role in the development and maintenance of obesity through multiple pathways (Tomiyama, 2019). One major pathway through which stress contributes to obesity is stress-induced eating (Tomiyama, 2019). This pathway can manifest either as overeating (i.e., an increase in food intake when experiencing stress) or as eating more unhealthy (Tomiyama, 2019). Even though stress is widely thought to induce overeating, numerous studies have shown that stress affects individuals' eating behaviour in different ways: Some individuals increase, some decrease food intake, while others show no change (for overviews, see Araiza & Lobel, 2018; Hill et al., 2021; Torres & Nowson, 2007). As early as 1994, a review (Greeno & Wing, 1994) concludes that empirical evidence at the time supports an individual-difference model of stress-induced eating (see Figure 1a) more strongly, as opposed to a general effect model (see Figure 1b). The individual-difference model is based on the assumption that the effect of stress on eating is determined by individual differences in learning history, attitudes, or biology which lead to high or low vulnerability to stress-induced eating. Identifying individuals at risk for stress-induced eating is important in order to develop tailored measures for the prevention and treatment of overweight and obesity. Based on this, a large number of studies aimed at identifying person-characteristics that predict an individual's dietary response to stress and thereby explain individual differences in the dietary response to stress. However, results have been highly inconsistent. For instance, some studies report gender differences, e.g., decreased food intake under stress in men and increases in eating in some women (Grunberg & Straub, 1992). Yet, other studies (e.g., Conner et al., 1999) found no gender differences. Furthermore, some studies suggests that individuals higher in emotional eating and dietary restraint are more likely to increase food intake when feeling stressed (e.g., O'Connor et al., 2008; Wallis & Hetherington, 2004; Wardle et al., 2000). Contrary to these findings, no moderating effect of emotional eating (Conner et al., 1999) as well as restrained eating (Conner et al., 1999; Pollard et al., 1995) was observed in other studies.

Figure 1

(a) Individual-difference model of stress-induced eating, (b) general effect model of stress-induced eating (adopted from Greeno & Wing, 1994)



A recent meta-analysis found no evidence that gender, age, weight status, and eating style (dietary restraint) moderate the relationship between stress and overall food intake (Hill et al., 2021). The authors, however, point out that a large number of included studies were restricted to one aspect of eating behaviour only (e.g., between-meal snacking) or were based on food intake in artificial environments (i.e., the laboratory). Furthermore, Araiza and

Lobel (2018) emphasise that the stress and eating relationship could be studied with increased reliability and validity through novel and sophisticated methodological approaches, such as Ecological Momentary Assessment (EMA).

1.3. Stress and Food Intake in Daily Life

1.3.1. Ecological Momentary Assessment (EMA) of the Stress and Eating Relationship

Eating is a repeated-occurrence health behaviour which is performed multiple times per day (Dunton, 2018). Therefore, a real-life micro-timescale approach that captures the dynamics of food intake and associated factors ecologically and momentarily is needed to understand the processes underlying eating behaviour in everyday life (Dunton, 2018). EMA seems particularly promising to study the dietary response to stress where and when it naturally takes place. EMA comprises repeated assessments of behaviours (e.g., food intake), experiences (e.g., perceived stress), and/or physiological parameters throughout a day in real life which enables studying complex psychological, behavioural, and/or physiological processes (Smyth & Stone, 2003). EMA overcomes disadvantages of traditional approaches (e.g., questionnaire-based surveys and laboratory tasks), as it minimises recall bias, maximises ecological validity and captures within-person processes and variation across time and settings (Shiffman et al., 2008). EMA can make a valuable contribution to understanding individual differences in the dietary response to stress, as it provides knowledge on which individuals in the real world engage in stress-induced eating. In contrast, knowing which individuals engage in stress-induced eating in a laboratory setting might not generalise to everyday life (i.e., behaviour in the laboratory might not predict behaviour outside the laboratory).

1.3.2. Findings from EMA Studies

Despite its potential, the number of EMA studies assessing the stress and eating relationship in daily life is limited. Zenk et al. (2014) assessed the association between daily hassles and snack food intake in women through EMA. Results indicated that participants were more likely to consume snack foods on days with more daily hassles. However, on the momentary level (i.e., within-day level), no relationship between a stressful event and concurrent as well as subsequent snack food intake was found. A further EMA study found no relationship between stress and subsequent self-reported healthy eating (Schultchen et al., 2019). However, the authors point out that only the general effect of stress on healthy eating was studied, disregarding individual differences associated with eating styles and sociodemographic variables found in previous studies. In contrast, two studies by Reichenberger et al. (2018, 2021) included person-characteristics to account for individual differences in the dietary response to stress. The first study assessed the role of stress in the context of taste- and hunger-driven eating as well as gender, BMI, and eating styles (emotional, external, and restrained eating) as potential moderators (Reichenberger et al., 2018). The authors found that stress was associated with a decrease in taste-eating. Gender, BMI, and eating styles did not moderate this association. Furthermore, no relationship between hunger-eating and stress as well as no moderating effect of gender, BMI, and eating styles were identified. In the second study, the moderating role of trait stress-eating (i.e., an individual's self-reported tendency to eat more, less or the same in response to stress) in the relationship between stress and perceived food intake was assessed (Reichenberger et al., 2021). Findings revealed that trait stress-eating moderated the relationship between stress and food intake when data collected throughout the day were aggregated (i.e., to the day level). Accordingly, individuals high in trait stress-eating reported increased food intake on days they experienced higher stress. In individuals with low trait stress-eating, no effect of stress on food intake was found. In contrast to the day level, the stress and food intake relationship was not moderated by trait stress-eating on the within-day level.

1.3.3. Limitations of EMA Studies

The available EMA studies offer first insights into the relationship between stress and eating in daily life. However, two main limitations have to be taken into consideration:

(1) Lack of EMA Studies Assessing Actual Food Intake. Even though the importance of capturing actual food intake (e.g., energy intake) when studying the stress and eating relationship has been highlighted (Araiza & Lobel, 2018; Hill et al., 2021), there is a lack of EMA studies which assess the relationship between stress and actual food intake. Actual food intake refers to capturing all consumed foods and drinks as well as consumed amounts, which are then used to generate nutritional values (e.g., energy or macronutrient intake). Some EMA studies assessed only perceived food intake, that is food intake is assessed based on the participant's subjective evaluation, e.g., on a scale from 0 (*eaten too little*) to 100 (*eaten too much*; Reichenberger et al., 2021). Others focused on certain aspects of food intake only. For example, Zenk et al. (2014) assessed merely snack food intake by asking whether each of five snack food categories (cookies or sweetened baked goods, chocolate or candy, ice cream or frozen dessert, salty snacks, and French fries or other fried side dishes) had been consumed since the last signal (*yes* or *no*). Schultchen et al. (2019), who did not find an association between stress and healthy eating, acknowledge that important effects of stress on eating behaviour (e.g., increased/decreased food intake) might have been overlooked due to focusing on healthy eating alone. The association between stress and actual food intake (i.e., calorie intake) was assessed only in one EMA study with nine patients with type 2 diabetes (Inada et al., 2019). So far, no EMA study assessed the relationship between stress and actual food intake in a larger and/or healthy sample. Food intake is a highly complex phenomenon making its assessment challenging and potentially causing the present lack of EMA studies which assess stress and actual food intake. Nonetheless, EMA offers great potential to assess food intake accurately as it avoids typical reporting biases (e.g., memory bias) present in traditional dietary assessment methods (e.g., food frequency questionnaires; Maugeri & Barchitta, 2019).

(2) Lack of Adequate Statistical Approach for Semicontinuous Outcomes. EMA allows investigating whether dynamic factors (e.g., stress) assessed several times per day predict food intake (e.g., energy intake) within a predefined time interval (e.g., within the subsequent 1.5 hours). However, food intake typically does not occur within each of these predefined time intervals or only small amounts (e.g., a snack) are consumed. This results in an outcome that is zero-inflated (i.e., contains a large proportion of zeros) and often right-skewed (i.e., contains a large proportion of small positive values). For instance, in the study by Reichenberger et al. (2021), participants reported no food intake in almost half of the 2.5-hour-intervals (2,318 out of 4,656) and in one third of the 4-hour-intervals (1,574 out of 4,648). The shorter time intervals are in which food intake is studied, the more likely it is that zero-inflation is high (i.e., in more time intervals the behaviour of interest is not shown). This type of data is commonly referred to as semicontinuous.

Linear multilevel modelling (also known as linear mixed or linear hierarchical modelling) is commonly used to analyse EMA data, as it accounts for the nested data structure (repeated assessments nested within individuals; e.g., Viechtbauer, 2022). However, traditional linear multilevel modelling cannot be applied to semicontinuous outcomes as the assumption of normally distributed residuals is likely violated. Beyond that, Baldwin et al. (2016) demonstrate that incorrect conclusions can occur when traditional linear multilevel models are used to analyse semicontinuous data without accounting for the large proportion of zeros. To circumvent the difficulties related to zero-inflation, previous EMA studies used one of two approaches: (1) To allow the use of traditional multilevel models, time intervals in which no eating is reported and which are therefore equal to zero are excluded (e.g., Reichenberger et al., 2021: study 1 – 2,318 out of 4,656 2.5-hour-intervals excluded). This approach removes zero-inflation and merely studies *how much* an individual eats (i.e., the amount of food intake) during eating occasions. (2) The semicontinuous outcome is dichotomised (i.e., zero values are retained, while positive values are set to one) to permit the application of multilevel logistic regressions to study *whether* but not how much an individual eats (i.e., the occurrence of food intake). For instance, Zenk et al. (2014) captured snack food intake by asking participants whether each of five snack food categories had been consumed since the last signal (*yes* or *no*). The "yes" responses were summed and then dichotomised as "none" (= 0) or "one or more" (= 1) in order to allow multilevel binary logistic regressions. However, both approaches show considerable limitations. The first approach causes loss of important information (Tooze et al., 2002) and can cause bias in the parameter estimates (Liu et al., 2008; Su et al., 2009). The second approach disregards data with important implications given that the number or amounts of consumed foods and drinks were reported. Hence, an adequate statistical approach for semicontinuous outcomes which allows studying the influence of stress on the occurrence as well as the amount of food intake and thereby does not overlook relevant information is needed to study the relationship between stress and food intake comprehensively in daily life.

1.4. Summary and Implications for Empirical Studies

Stress affects individuals' food intake in different ways. Despite various efforts to identify person-characteristics that explain individual differences in the dietary response to stress, evidence remains inconclusive. EMA offers great potential to study the dietary response to stress, where and when it naturally occurs. However, only very few EMA studies assessed the relationship between stress and eating so far. Moreover, these studies show two substantial limitations, as they do not assess actual food intake and apply statistical approaches which involve the loss of important information. Based on these limitations, there is a need for (1) an EMA tool that allows the assessment of stress and actual food intake and (2) a statistical approach that accounts for the semicontinuous outcome appropriately. These advancements are needed in order to gain novel insights into the stress and eating relationship and to identify individuals at risk for stress-induced eating in daily life.

2. The Present Dissertation

2.1. Aims of the Present Dissertation

The relationship between stress and food intake is highly complex (Hill et al., 2021). Even though many studies have tried to shed light on it, central questions remain unanswered. "Our frequent reliance on global, retrospective reports seriously limits our ability to accurately characterize, understand, and change behavior in real-world settings and misses the dynamics of life as it is lived, day-to-day, hour by hour" (Shiffman et al., 2008, p. 3). In contrary, EMA opens up new perspectives and is a promising approach to provide novel scientific evidence and advance the research field. For these reasons, the present dissertation aims at deepening the understanding of the relationship between stress and food intake by applying EMA. Given the limitations of previous EMA studies, the present dissertation addresses the need for (1) an EMA tool that allows the assessment of stress and actual food intake as well as (2) a statistical approach that accounts for semicontinuous outcomes appropriately. The first aim of the dissertation was to develop and evaluate a mobile tool that allows capturing stress and actual food intake in daily life (Paper 1). The second aim of the dissertation was to introduce multilevel two-part modelling as a novel and sophisticated statistical approach for semicontinuous dietary outcomes (Paper 2). The newly developed and evaluated tool as well as the advanced statistical method built the foundation for the third aim of the dissertation. The EMA tool and the statistical approach were integrated in order to gain novel insights into individual differences in the dietary responses to stress and thereby identify individuals at risk for stress-induced eating in daily life (Paper 3).

2.2. Paper 1: Development and Evaluation of an EMA Tool

Ruf, A., Koch, E.D., Ebner-Priemer, U., Knopf, M., Reif, A., & Matura, S. (2021). Studying Microtemporal, Within-Person Processes of Diet, Physical Activity, and Related Factors Using the APPetite-Mobile-App: Feasibility, Usability, and Validation Study. *Journal of Medical Internet Research, 23*(7), e25850. https://doi.org/10.2196/25850

Paper 1 of the present dissertation addressed the need for an EMA tool that allows the assessment of the relationship between stress and actual food intake in daily life. As the suitability of novel EMA tools for the use in daily life should be examined, the APPetite-mobile-app was not only developed but also evaluated. Three central criteria need to be considered carefully when developing and evaluating EMA tools: (1) *Feasibility* refers to the possibility that something can be done or achieved. Since the data collection of EMA tools takes place in the participants' daily life, the repeated assessments should conflict as little as possible with other obligations. Only if participants are able to comply with the EMA protocol, sufficient data can be captured in order to draw meaningful conclusions.
(2) *Usability* describes the degree to which something is easy to use. Only if the EMA tool is easy to use, participants engage with it and sufficient data can be collected. (3) *Validity* refers to the extent to which a tool measures what it declares to measure. Given the complexity of food intake, this is of great importance in the context of the assessment of actual food intake. Only if the EMA tool enables a valid dietary assessment, the relationship between stress and actual food intake can be assessed reliably. Consequently, Paper 1 of the present dissertation examined the feasibility and usability of the APPetite-mobile-app as well as the validity of the incorporated food record empirically.[1]

2.2.1. Methods

The APPetite-mobile-app. The APPetite-mobile-app captures food intake event-contingent through a food record, i.e., participants are asked to record foods and drinks as soon as possible after consuming them. The obtained dietary data were transferred by trained staff to myfood24-Germany, a 24-hour dietary recall (Koch et al., 2020), in order to generate nutritional values (e.g., energy and macronutrient intake). Stress as well as context, affect, impulsivity, and food availability are assessed signal-contingent through eight semirandom prompts per day.

[1] Note that only parts of Paper 1 that are most relevant to the present dissertation are described in the following. For instance, the feasibility evaluation of the activity tracker is not covered.

Evaluation of Feasibility. Feasibility was separately assessed for the EMA prompts as well as the food record. Prompt feasibility was, amongst others, evaluated based on compliance rates (percentage of complete prompts within received prompts). Since the momentary nature of responses is a key feature of EMA, responding to prompts should be initiated without delays. Therefore, response latency (time from first prompt signal to prompt responding) was examined as a feasibility indicator. The feasibility of the food record could not be evaluated based on compliance rates, since it cannot be differentiated between someone not recording a food because of noncompliance or because of not actually consuming it. Therefore, other measures, such as the number of recorded eating/drinking events and the reporting latency (time between food intake and food recording), were assessed. Data of 157 participants who completed the APPetite-mobile-app for three days were used to examine feasibility.

Evaluation of Usability. The System Usability Scale (SUS; Brooke, 1996) was used to assess usability. A total SUS-score between 0 and 100 was calculated. Higher numbers indicate better usability. The questionnaire was completed by 84 participants.

Evaluation of Validity. Two approaches were used to assess the validity of the food record: (1) comparison with an established reference method (often referred to as the assessment of relative validity) and (2) comparison with total energy expenditure (TEE). A web-based, self-administered 24-hour dietary recall (Koch et al., 2020) was used as the reference method. Following a counterbalanced crossover design to control for unwanted order effects, participants were assigned to one of two groups: Group 1 completed three 24-hour recalls exactly a week before and Group 2 exactly a week after the completion of the APPetite-food record. Hence, the same weekdays, the week before or after, were assessed. Habitual energy, protein, fat, carbohydrates, sugar, and fibre intake (i.e., the mean of the three days) assessed through the APPetite-food record was compared with habitual dietary intake assessed through the 24-hour recall. Furthermore, energy intake in kilocalories (kcal) was compared to TEE assuming that energy intake is equivalent to TEE in weight-stable individuals (Trabulsi & Schoeller, 2001). TEE was estimated from accelerometry which was recorded by move 3 sensors (movisens GmbH, Karlsruhe, Germany) for seven days. Data from 44 participants (20 in group 1 and 24 in group 2) was used for the evaluation of validity. The two groups did not differ regarding gender distribution, age, and BMI.

Samples. Data of the present dissertation were collected within the APPetite study which is part of the Horizon 2020 project Eat2beNICE. The APPetite study recruited participants from existing study cohorts, such as the LORA (Longitudinal Resilience Assessment) study which enrolled individuals not affected by psychiatric conditions (Chmitorz et al., 2021), the PROUD (Prevention of Comorbid Depression and Obesity in Attention-Deficit/Hyperactivity Disorder [ADHD]) study which included patients with ADHD (Mayer et al., 2018), and the BipoLife-A1 study which follows up individuals with an increased risk for bipolar disorders, including patients affected by ADHD and/or depression (Pfennig et al., 2020; Ritter et al., 2016). Demographics of the samples of the feasibility, usability, and validation study are shown in Table 1.

Table 1

Demographics of the samples of the feasibility, usability, and validity evaluation

	Feasibility (n = 157)	Usability (n = 84)	Validity (n = 44)
Gender, *n* (%)			
Female	100 (63.7)	55	33
Male	57 (36.3)	29	11
Age, mean (*SD*)	28.04 (7.22)	29.26 (7.41)	28.64 (8.13)
BMI, mean (*SD*)	24.71 (4.81)	24.82 (5.26)	23.8 (3.62)
Cohort, *n*			
LORA	136	67	44
PROUD	7	6	-
BipoLife	14	11	-

2.2.2. Results and Discussion

Feasibility. Across the three days, 81.73% (*SD* = 21.65) of all received prompts were completed. Although there is no official criterion for good compliance, compliance rates above 80% are generally considered good (Stone & Shiffman, 2002). Therefore, the mean prompt compliance above 80% indicates good feasibility. Furthermore, no significant difference in compliance among the three days was found suggesting no decrease in engagement with the prompts. Other EMA studies which assessed more than three days found substantial declines in response rates (e.g., 40% decline from 63% on day 1 to 23% on day 7, even with only four prompts per day - Spook et al., 2013). Based on this, the duration of the 3-day EMA period in the present study seems feasible. Prompts were responded to on average after 189.32 seconds (*SD* = 388.65). Participants initiated responding to 70.54% (2,157/3,058) of all prompts within 60 seconds after the first prompt signal. Considering that responses up to 30 minutes after the first prompt signal were allowed, the response latency of just over three minutes can be rated as short and thereby highlights the momentary nature of the responses and the feasibility of the prompts.

A mean of 7.02 (*SD* = 3.33) eating and drinking events were entered by participants per day. A previous study found similar numbers of eating and drinking occasions, i.e., 20.7 eating and drinking occasions over a 3-day period (Ashman et al., 2017). Since the objective of the APPetite-food record was to record food intake in real time or near real time, it is important to examine the time between food intake and food recording. Foods and drinks were recorded on average 58.35 minutes (*SD* = 127.52) after intake which confirms that food recording was based on shorter retention intervals compared to traditional dietary assessment methods (e.g., 24-hour recalls). Hence, these findings highlight the feasibility of recording food intake in real time or near real time and the potential of EMA to avoid typical reporting biases (e.g., recall bias) present in traditional dietary assessment methods. However, food recording latency increased over the three days from 53.51 minutes on the first day to 90.81

minutes on the third day. This increase suggests that a decrease in motivation, potentially due to high burden, might have interfered with maintaining the promptly reporting of foods and drinks. Yet, the results of the feasibility evaluation indicate that the APPetite-mobile-app is overall a feasible EMA tool.

Usability. Usability of the APPetite-mobile-app was rated moderate with a SUS-score of 61.9 out of 100 (*SD* = 17.79; range 17.5–97.5). Compliance and rated usability showed no significant correlation. Interestingly, in a study which assessed the usability of the top seven diet-tracking apps, two apps (Lose It! = 59.2; MyDietCoach = 46.7) were rated lower on the SUS compared to the APPetite-mobile-app (Ferrara et al., 2019). Even though the comparison is difficult given that these apps focus on the dietary assessment only, while the usability of the APPetite-mobile-app was rated based on the dietary assessment as well as the EMA prompts, these findings are somewhat surprising. Typically lower usability is expected for scientific tools compared to commercial tools, as scientific tools are often developed without professional app developers due to considerable costs. Further usability issues have been reported for the commercial app "MyFitnessPal" as only 20% of participants would continue to use the app after study participation (Chen et al., 2019). These findings highlight that achieving good usability of apps designed to record food intake is generally challenging. Even though the usability of the APPetite-mobile-app was rated relatively low given the range of the SUS, it is nonetheless comparable with some commercial diet-tracking apps. Yet, more importantly, no negative effect of the usability of the APPetite-mobile-app on compliance was found. Therefore, improved usability is desirable but not essential for its scientific use.

Validity. The evaluation of relative validity identified considerable differences in habitual energy and macronutrients intake between the APPetite-food record and the 24-hour recall on the group and individual level. At first glance, one might think that these discrepancies indicate a lack of validity of the APPetite-food record. Yet, a closer look raises some doubts: Even though 24-hour recalls are often used as the reference method in the assessment of relative validity, inaccurate estimations of energy intake captured by 24-hour recalls have been previously reported (e.g., Lopes et al., 2016). The APPetite-food record captured higher intakes for energy as well as all macronutrients compared to the 24-hour recall which provides first evidence for lower levels of underreporting of the APPetite-food record compared to the 24-hour recall. This is supported by the comparison of energy intake and TEE. Results show that energy intake was assessed fairly accurate by the APPetite-food record on the group level on two of three days when compared to TEE of the exact same day, while the comparison with mean TEE (2417.8 kcal) indicated that the 24-hour recall (1909.2 kcal) underestimated habitual energy intake to a larger degree than the APPetite-food record (2146.4 kcal). These findings lead to the conclusion that the APPetite-food record might be a more accurate dietary assessment method compared to widely used 24-hour recalls.

2.3. Paper 2: Introduction of an Advanced Data Analysis Approach

Ruf, A., Neubauer, A.B., Ebner-Priemer, U., Reif, A., & Matura, S. (2021). Studying dietary intake in daily life through multilevel two-part modelling: a novel analytical approach and its practical application. *International Journal of Behavioral Nutrition and Physical Activity, 18*(130). https://doi.org/10.1186/s12966-021-01187-8

Paper 2 addressed the need for an appropriate data analysis approach for semicontinuous dietary outcomes. The aim of Paper 2 was not only to introduce multilevel two-part modelling as an advanced data analysis approach for semicontinuous dietary data, but also to provide practical guidance on the implementation of these models in freely available software. Multilevel two-part modelling is a highly informative, but less well-known statistical approach to analyse semicontinuous data. It allows studying the occurrence as well as the amount of food intake while accounting for their potential dependency. Paper 2 examined an exemplary research question and provides corresponding data and code to facilitate the model application to readers.

2.3.1. Methods

Multilevel Two-Part Model for Semicontinuous Dietary Data. Multilevel two-part models treat semicontinuous outcomes as a combination of two parts: (1) the zero part in which the outcome is either 0 or 1, indicating *whether* an individual eats in a given time interval, and (2) the continuous/positive part which contains the nonzero values of the outcome, indicating *how much* an individual eats if he/she eats in a given time interval. The two parts of the outcome follow different distributions, wherefore the present dissertation proposes a multilevel two-part model which combines a multilevel logistic regression for the zero part to study *whether* an individual eats (i.e., the occurrence of food intake) and a multilevel gamma regression for the right-skewed continuous part to study *how much* an individual eats when he/she eats (i.e., the amount of food intake). However, the two parts are likely not independent. To account for the potential relation between the occurrence and the amount of food intake, the cross-part correlation (i.e., the correlation between the random effects across the two parts) is modelled. The R-package brms (Bürkner, 2017, 2018) was chosen to implement the proposed multilevel two-part model as it provides a user-friendly and freely available application of the model.

Data and Material. To illustrate the model implementation and interpretation, it is assessed exemplarily whether momentary energetic arousal and gender predict the occurrence and/or the amount of energy intake. Data of 99 participants collected through the APPetite-mobile-app (developed and evaluated in Paper 1) were used for analysis. Each momentary assessment of energetic arousal ($N = 2{,}044$) was paired with energy intake in kcal within the subsequent up to two hours. Almost half (48.4%, $n = 989$) of the time intervals showed no energy intake and were therefore equal to 0.

2.3.2. Results and Discussion

Two findings of the exemplary analyses are of particular interest as they highlight the importance of applying multilevel two-part modelling to semicontinuous outcomes:

(1) Results indicated that gender was associated with the amount consumed during eating occasions, but not with the occurrence of eating (i.e., women consumed on average around 22% less energy in time intervals in which energy intake occurred compared to men. Yet, women and men did not differ regarding the occurrence of energy intake). This highlights that the differentiation between the two parts reveals part-specific associations which cannot be detected through traditional multilevel modelling. (2) Fairly strong to moderate positive cross-part correlations between the random intercepts (0.77, 0.71, and 0.56) were found. This indicates that the two model parts were related as such that participants who consume on average more energy during eating occasions eat on average less often. Not accounting for this relationship can cause bias in parameter estimates. This is of relevance particularly in the continuous part of the model as the cluster size of the continuous part (i.e., the number of observations with food intake within an individual) is determined by the zero part. For instance, given the moderate to strong positive cross-part correlations, individuals who eat less often have fewer observations in the continuous part. Yet, the few observations consist of larger amounts. On the contrary, individuals who eat more frequently have more observations in the continuous part which consist of smaller amounts. As a consequence, larger values of food intake are underrepresented and smaller values are overrepresented in the continuous part. Even when researchers are interested only in the continuous part of the semicontinuous outcome and therefore choose to fit a single model, the described bias will be present (Su et al., 2009).

The model proposed in Paper 2 overcomes several limitations of traditional linear multilevel modelling when it comes to semicontinuous data: (1) It accounts for the zero-inflation by incorporating two model parts, a zero and a continuous part, which prevents incorrect inferences (as shown by Baldwin et al., 2016); (2) It accommodates the skewness of the continuous part of the outcome by implementing a gamma regression which avoids controversial transformation of the outcome (e.g., logarithmizing) and maintains the original metric of the data; (3) It accounts for the dependency between the two model parts by integrating the cross-part correlation which prevents estimation bias present in separate models (as outlined above). Hence, multilevel two-part models not only solve a statistical issue, but also offer novel and distinct insights in terms of the occurrence and the amount of food intake while accounting for the potential dependency between them. Multilevel modelling is therefore a conceptually as well as methodologically highly promising approach for semicontinuous dietary outcomes.

2.4. Paper 3: Individual Differences in the Dietary Response to Stress in Daily Life

Ruf, A., Neubauer, A.B., Koch, E.D., Ebner-Priemer, U., Reif, A., & Matura, S. (2022). Individual differences in the dietary response to stress in ecological momentary assessment: Does the individual-difference model need expansion? *Applied Psychology: Health & Wellbeing,* 1-21. https://doi.org/10.1111/aphw.12400

Based on the inconclusive body of evidence in the context of individual differences in the dietary response to stress, the aim of Paper 3 of the present dissertation was to gain novel insights into the relationship between stress and actual food intake by integrating the accurate assessment of actual food intake (Paper 1) and the advanced data analysis approach for semicontinuous data (Paper 2). On basis of the recommendation by Hill et al. (2021) for (1) more detailed measures of the nature of the stressors, (2) more accurate assessments of food consumption, such as energy intake, (3) more studies that examine key moderating variables of the stress and eating relationship, (4) assessment of eating styles, and (5) accurate measures of weight, height, and diet status as well as the importance of accounting for dispositional stress-related eating (i.e., self-reported tendency to eat more, less or the same in response to stress), the aim of Paper 3 was to examine (1) whether individuals differ in the dietary response to stress in daily life, (2) whether individual differences in the dietary response to stress can be explained by gender, age, BMI, trait stress-eating, and eating styles, and (3) whether these findings support the individual-difference model of stress-eating (Greeno & Wing, 1994).

2.4.1. Methods

Procedure. Participants completed two in-person sessions as well as the EMA protocol of the APPetite-mobile-app for three days. Body weight, body height, eating styles, and trait stress-eating were assessed during the first in-person session. The German version of the Three-Factor-Eating-Questionnaire (Stunkard & Messick, 1985; German version: Pudel & Westenhöfer, 1989) was used to capture the following eating styles: cognitive restraint of eating, disinhibition, and hunger. Trait stress-eating was assessed using the Salzburg Stress Eating Scale (SSES; Meule et al., 2018). Furthermore, detailed training to familiarise with the APPetite-mobile-app was provided in the first session. During the EMA period, stress was assessed eight times per day through semirandom signal-contingent prompts. Three items were adapted from Reichenberger et al. (2018) to assess perceived stress since the last prompt or since waking up (in the first prompt per day). Food intake was captured event-contingent through the APPetite-food record. To obtain energy intake in kcal, the collected dietary data were transferred to myfood24-Germany (Koch et al., 2020).

Sample. In total, 185 healthy adults of the LORA cohort participated in the APPetite study. The inclusion criteria of the LORA study were the age between 18 and 50 years, normal or corrected eyesight, sufficient German language proficiency and the capacity to provide informed consent. The lifetime diagnosis of schizophrenia or bipolar disorder, organic mental disorders, substance dependence syndromes or other current severe axis I disorders, current severe medical conditions, learning disabilities, serious neurological disorders, and the participation in a drug trial in the previous 6 months were exclusion criteria. Beyond that, the International Neuropsychiatric Interview (M.I.N.I; Ackenheil et al., 1999; Lecrubier et al., 1997) was administered in the first in-person session of the LORA study to confirm the absence of current mental disorders. Four participants dropped out after the first in-person session of the APPetite study. Due to invalid data, one participant had to be excluded. Due to poor dietary records, 26 participants were excluded. Beyond that, 13 single days of the EMA period had to be excluded as dietary data was recorded poorly. The final sample consisted of 154 participants (see Table 2 for demographics of the total sample as well as for men and women separately).

Table 2

Demographics of the total sample as well as for men and women separately

	Total sample (N = 154)	Men (n = 42)	Women (n = 112)
Age in years, mean (*SD*)	28.91 (7.75)	30.48 (7.52)	28.32 (7.78)
BMI, mean (*SD*)	24.20 (4.09)	26.72 (4.15)	23.28 (3.67)
Education, *n* (%)			
GCSE (*Mittlere Reife*)	1 (0.7)	1 (2.4)	-
High-school diploma (*Abitur*)	55 (35.7)	12 (28.6)	43 (38.4)
Vocational training (*Berufsausbildung*)	10 (6.5)	2 (4.8)	8 (7.1)
Academic degree (*Hochschulabschluss*)	79 (51.3)	23 (54.8)	56 (50.0)
PhD (*Promotion*)	9 (5.8)	4 (9.5)	5 (4.5)

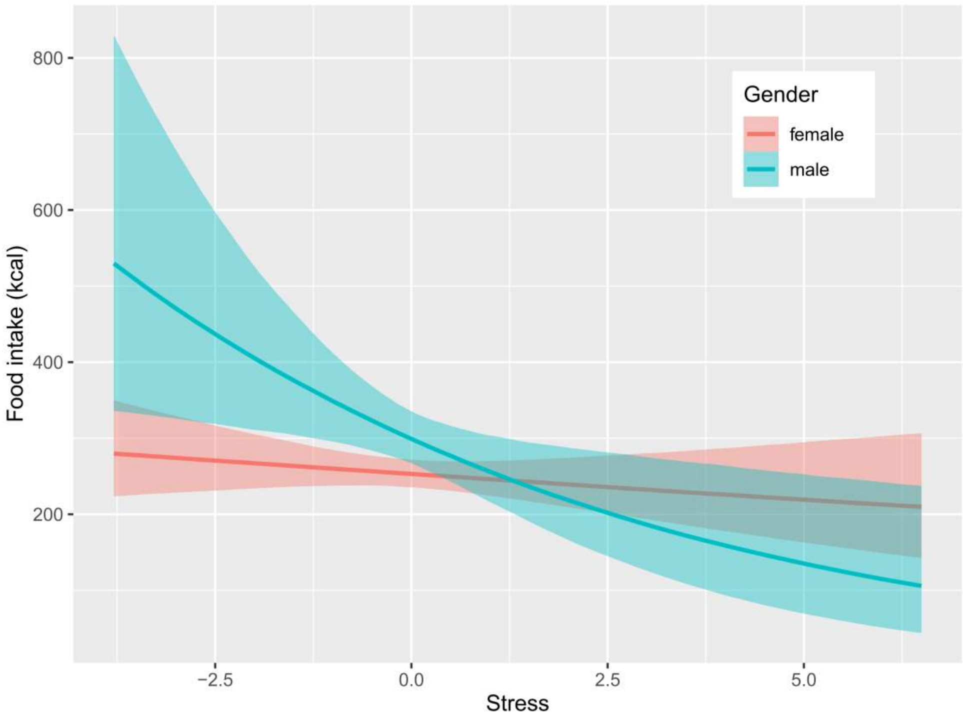
Data Preprocessing and Analysis. After data cleaning (e.g., exclusion of time intervals in which the stress items were not completed), the final dataset included 2,779 time intervals. Each time interval for which stress was assessed (i.e., time between current prompt and previous prompt/waking up) was matched to concurrent energy intake in kcal (i.e., sum of any intake of energy within the respective time interval). The multilevel two-part model described in Paper 2 was used for analysis due to the nested data structure (time intervals [Level 1] nested within individuals [Level 2]) and the zero-inflated, right-skewed (i.e., semicontinuous) outcome. To examine whether the effect of stress differs between individuals, a random slope/effect for stress was included. First, a model with the Level-1 predictor stress in both model parts (i.e., the zero part as well as the continuous part) was run to examine individual (i.e., between-person) differences in the within-person effect of stress on the occurrence and the amount of energy intake. Next, the association between the Level-1 predictor stress in interaction with the Level-2 predictor (1) gender, (2) age, (3) BMI, (4) trait stress-eating, (5) dietary restraint, (6) disinhibition, or (7) hunger (cross-level interaction) and energy intake in both model parts was tested.

2.4.2. Results and Discussion

Results indicate that stress was not related to whether individuals eat (i.e., the occurrence of energy intake). Gender, age, BMI, trait stress-eating, and eating styles did not moderate the relationship between stress and the occurrence of eating. Accordingly, stress does not seem to make individuals more or less likely to eat. Since this is the first study to differentiate between effects of stress on the occurrence and the amount of food intake, these findings require replication. BMI, age, trait stress-eating, and eating styles did not moderate the relationship between stress and the amount of food intake. Only gender moderated the relationship between stress and the amount of food intake, as such that stress had a significant effect on the amount of food intake in men, but not in women. Increased stress was associated with decreased amounts of food intake in men (see Figure 2). Hence, stress seems to affect men's eating behaviour more intensely compared with women. While this is in line with a study which found that men significantly decreased food intake in the stress condition (Grunberg & Straub, 1992), it stands in contrast to studies that identified no gender differences (e.g., Conner et al., 1999). Yet, gender differences in compliance could to some extent explain the gender differences found in the present study. Participants might be less likely to record foods when experiencing stress, which might appear as if food intake decreases as a response to stress. Systematic noncompliance in food recording due to stress might be particularly common in male participants as a recent meta-analysis suggests that men are generally less compliant in EMA studies compared to women (Wrzus & Neubauer, 2022). Yet, in order to reduce potential bias due to systematic noncompliance in food recording, participants and days with poor food records were thoroughly excluded.

Figure 2

Relationship between stress and the amount of food intake moderated by gender

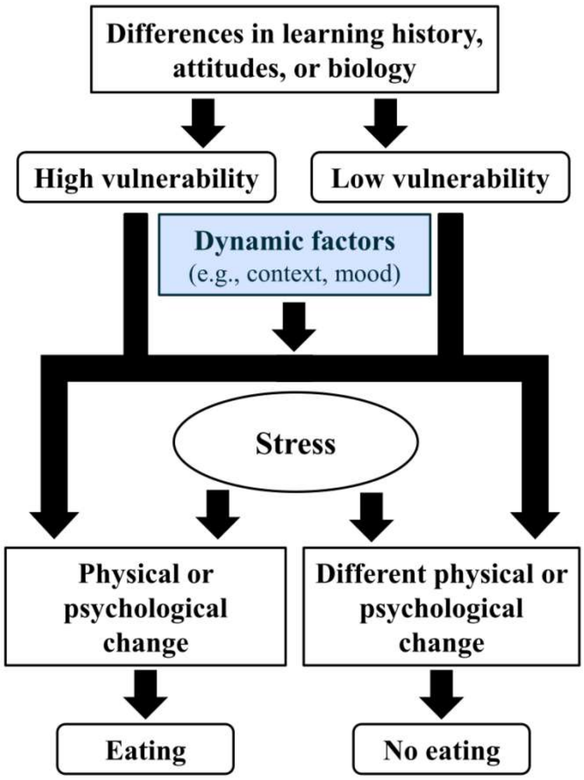


Contrary to previous EMA studies which indicate that stress is associated with eating behaviour only on the day level (Reichenberger et al., 2021; Zenk et al., 2014), the present study suggests that stress has short-term effect on eating, since an association between stress and food intake on the within-day level was found in men. Findings from laboratory studies underline the relevance of short-term effects, as effects of stress on food consumption were found during or shortly after stress-induction (e.g., Grunberg & Straub, 1992 – during a 14-minute stress-induction film; Epel et al., 2001 – within 30 minutes following stress-induction). Hence, more research is needed to identify the time frame in which stress influences food intake.

Surprisingly, stress had no significant random effect/slope in either of the two model parts indicating that individual differences in the stress and eating relationship were small. This finding suggests that individuals might not always show the same dietary response to stress, as intraindividual (i.e., within-person) variability might mask individual differences. This is in line with first evidence that indicates that time-varying factors, such as easy food availability, moderate the stress and eating relationship (Zenk et al., 2014). Hence, the dietary response to stress might not be as stable as yet assumed. Based on this, an extension of the individual-difference model which accounts for time-varying factors as potential moderators of the stress and eating relationship is proposed: the dynamic individual-difference model (illustrated in Figure 3). These findings suggest that not only identifying individuals at risk for stress-induced eating, but also situations with an increased risk for stress-induced eating is necessary to develop tailored strategies for the prevention and treatment of overweight and obesity. Research is needed to verify the expanded model.

Figure 3

Expansion of the individual-difference model of stress-eating by Greeno and Wing (1994)



2.5. Conclusions

The present dissertation aimed at integrating the assessment of actual food intake (Paper 1) and an appropriate data analysis approach for semicontinuous data (Paper 2) in order to gain novel insights into the relationship between stress and food intake in daily life (Paper 3). Therefore, a novel EMA tool, the APPetite-mobile-app, was developed and evaluated (Paper 1). Multilevel two-part modelling was introduced as a novel approach to study semicontinuous dietary data (Paper 2). Based on the inconclusive body of evidence in the context of individual differences in the dietary response to stress, individual differences in the dietary responses to stress in daily life were examined in order to identify individuals at risk for stress-induced eating (Paper 3).

The feasibility, usability, and validation study showed that the APPetite-mobile-app is overall a suitable EMA tool to capture micro-temporal, within-person processes underlying actual food intake in real time or near real time (Paper 1). These findings provide first evidence that the assessment of the relationship between stress and actual food intake in an EMA setting is attainable. Yet, two challenges became apparent: (1) Food recording in daily life can be burdensome, which might lead to poor dietary data (e.g., 26 out of 185 participants and 13 single days excluded due to poor dietary data in Paper 3) and might contribute to the commonly low usability ratings of diet-tracking apps. (2) Dietary data processing can be time-consuming. In the present study, quality of dietary data was checked manually and implausible food records were discussed with the participant to resolve any uncertainties. Even though this was a time-intensive procedure, it was needed to ensure high quality of the dietary data. Furthermore, the APPetite-mobile-app does not incorporate the automatised generation of nutritional values. Nutritional values were generated based on manual data transfer which is time-consuming and can be error-prone. These challenges highlight the need for advances in the assessment of actual food intake. Future tools should incorporate the automatisation of the nutritional value generation. Furthermore, photos taken of the foods and drinks could be used as time stamps of food intake and as memory aids. Another useful addition could be a feature which allows participants to register skipped meals. Ultimately, however, only technical advancements, such as the accurate passive detection of eating and the automatised photo-based dietary assessment, can significantly reduce participants' burden and the time required for manual data processing. Only if these advances are achieved, EMA study durations can be extended as well as sample sizes increased. Since a systematic review notes that only few EMA tools which capture food intake were validated against current dietary assessment methods or nutritional markers (Maugeri & Barchitta, 2019), future tools which are developed to capture the stress and eating relationship should be subject to an evaluation. The present dissertation hopes that providing first evidence that EMA is overall a feasible and valid approach to capture stress and actual food intake motivates other researchers to use EMA to capture the relationship between stress and actual food intake in daily life and initiates a discussion about feasible, usable, and valid methods to assess this relationship. Only when assessment strategies are developed, evaluated, shared, discussed, and improved, advances in the research field can be accomplished.

Results of Paper 2 showed that multilevel two-part modelling is a methodologically as well as conceptually promising approach to study food intake in daily life as it accounts for the semicontinuous data structure and provides novel and distinct insights in terms of the occurrence as well as the amount of food intake. Not only the exemplary analysis of Paper 2 but also the results of Paper 3 highlight that multilevel two-part modelling reveals part-specific associations which otherwise would be overlooked. Despite their potential, multilevel two-part models are still largely unknown which might be due to their complexity and their predominant availability in statistical software that is less common (e.g., WinBUGS) or not free to use (e.g., SAS Proc NLMIXED). To increase the visibility and accessibility of these models, Paper 2 offers an application-oriented introduction (including open data and code) to multilevel two-part modelling using brms, a package in the commonly used and freely available software environment R (R Core Team, 2020). Since the importance of EMA studies assessing factors influencing food intake in daily life is likely growing, Paper 2 contributes to establishing an appropriate data analysis procedure that accounts for the semicontinuous data structure. In fact, most behavioural outcomes show semicontinuous characteristics which can benefit from differentiating between factors either influencing *whether* the behaviour is shown or *how long/intensive/often* the behaviour is shown (e.g., social interaction: Has an individual interacted socially? If so, how many minutes?). The shorter time intervals are in which a behaviour is studied, the more likely it is that the outcome is zero-inflated, i.e., the behaviour is not shown in each assessment interval. Given the growing number of EMA studies, semicontinuous outcomes will become more prevalent and multilevel two-part models will likely gain greater recognition in the near future.

Following a novel approach (i.e., assessing actual food intake in an EMA setting combined with sophisticated multilevel two-part modelling), Paper 3 provides novel insights into the relationship between stress and actual food intake in daily life. Findings of Paper 3 suggest that individual differences in the dietary response to stress in daily life were small. This indicates that an individual's dietary response to stress might not be as stable as yet assumed. Time-varying factors might play a role in shaping food intake under stress. This is confirmed by first evidence showing that time-varying factors (i.e., easy food availability) moderate the stress and eating relationship (Zenk et al., 2014). Hence, the dietary response to stress might not only differ between individuals but also within individuals (i.e., between situations). So far, research is based predominantly on the assumption that the dietary response to stress is stable within an individual (i.e., a trait) as suggested by the individual-difference model of stress-induced eating (Greeno & Wing, 1994) and has therefore focused on person-characteristics as potential moderators of the stress and eating relationship. Even though research has tried for decades to identify individuals at risk for stress-induced eating, results are highly inconsistent and do not allow final conclusions. Therefore, the present dissertation proposes an expansion of the individual-difference model which accounts for time-varying factors as potential moderators of the stress and eating relationship. Based on the expanded model, research should not only aim at identifying individuals at risk for stress-induced eating, but also at identifying situations in which individuals more likely engage in stress-induced eating. This knowledge is needed for the development of tailored measures for the prevention and treatment of overweight and obesity.

Since the present dissertation is the first to assess the influence of stress on the occurrence as well as the amount of actual food intake in daily life through multilevel two-part modelling, future research is needed to replicate the findings and to verify the expansion of the individual-difference model. Furthermore, some limitations of the present work need to be acknowledged, from which recommendations for future studies can be derived: (1) The relationship between stress and actual food intake was assessed over three days. Three days are a rather short assessment period which might not allow capturing the whole spectrum of the stress and eating relationship. For instance, relatively low levels of stress were found. A prolonged assessment period might allow capturing a larger variance in stress intensity. Yet, this limitation is closely linked to the much needed advances in dietary assessments (as described above). Improvements are needed in order to reduce participants' burden and ultimately to allow assessing the relationship between stress and actual food intake on a larger scale (e.g., over longer assessment periods or in larger samples). (2) Stress was assessed based on self-reports. However, participants might be less likely to respond to prompts when experiencing stress, which might be one reason for the relatively low levels of stress found in the current study. While this bias is likely small in the current sample, given the high compliance rates with the EMA prompts, physiological stress responses (e.g., heart rate variability) could be a useful addition to avoid systematic non-compliance in future studies. (3) The present dissertation studied the association between stress and quantitative food intake (i.e., energy intake). However, some research suggests that stress might not be related to food intake per se, but to increased intake of unhealthy foods and decreased consumption of healthy foods (Araiza & Lobel, 2018). Although the aetiology of obesity is multifactorial (Aronne et al., 2009), on the individual level the fundamental cause of overweight and obesity is a positive energy balance between energy consumption (i.e., food intake) and energy expenditure (e.g., physical activity and diet-induced thermogenesis; World Health Organization, 2021). Consequently, total energy intake is a key target in the prevention and treatment of overweight and obesity and was therefore studied in the present dissertation. However, future studies should in addition examine the association between stress and qualitative food intake (e.g., healthy vs. unhealthy food intake). (4) Time pressure is an important construct to consider when assessing the influence of stress on eating (Reichenberger et al., 2018). For instance, stress accompanied by time pressure might be related to decreases in food intake, since there is simply no time to eat. Yet, the present dissertation did not differentiate between stress being related or unrelated to time pressure. Future studies should consider differentiating between the two.

Despite these limitations, the present dissertation extends the research on the stress and eating relationship as it overcomes limitations of previous EMA studies and yields novel insights into the relationship between stress and actual food intake in daily life. The present dissertation is the first to develop and evaluate an EMA tool which allows the assessment of the relationship between stress and actual food intake in daily life as well as to introduce and apply multilevel two-part modelling to study the influence of stress on the occurrence as well as the amount of food intake. By that, it provides food for thought for methodological as well as contentual considerations which may facilitate advances in the research field. These advances are needed in order to broaden the understanding of the relationship between stress and actual food intake and to identify specific situations and/or individuals at risk for stress-induced eating.

3. Zusammenfassung

3.1. Einleitung

Sowohl die Ernährung als auch das Stressempfinden haben einen direkten Einfluss auf die menschliche Gesundheit. Darüber hinaus beeinflusst Stress die Gesundheit indirekt durch Veränderungen von Gesundheitsverhalten, wie zum Beispiel der Ernährung. Studien konnten zeigen, dass sich Personen hinsichtlich ihrer Nahrungsaufnahme bei Stress unterscheiden: Manche essen mehr, manche weniger, während andere keine Veränderung zeigen. Das "individual-difference model" von Greeno und Wing (1994) besagt, dass Unterschiede in der Nahrungsaufnahme bei Stress auf individuelle Unterschiede in der Lerngeschichte, den Einstellungen oder der Biologie zurückzuführen sind. Die Identifikation von Personen mit erhöhtem Risiko für stressbedingte Nahrungsaufnahme ist demnach von entscheidender Bedeutung für die Entwicklung gezielter Maßnahmen zur Prävention und Behandlung von Übergewicht und Adipositas. Basierend darauf verfolgten zahlreiche Studien das Ziel, Personenmerkmale zu identifizieren, die individuelle Unterschiede in der Nahrungsaufnahme bei Stress erklären. Es zeigten sich jedoch widersprüchliche Befunde. Eine Metaanalyse von Hill et al. aus dem Jahr 2021 fand keine studienübergreifenden Hinweise darauf, dass Geschlecht, Alter, Gewichtsstatus und Essstil den Zusammenhang zwischen Stress und der Nahrungsaufnahme moderieren. Die Autoren weisen jedoch darauf hin, dass sich eine große Zahl der einbezogenen Studien entweder auf lediglich einen einzelnen Aspekt des Essverhaltens (z. B. Snacks zwischen den Mahlzeiten) beschränkte oder die Nahrungsaufnahme in einer künstlichen Umgebung (d. h. im Labor) untersuchte. Im Gegensatz dazu ermöglicht es der Ansatz des Ecological Momentary Assessments (EMA), die Beziehung zwischen Stress und der Nahrungsaufnahme im Alltag, unmittelbar wo und wann sie sich auf natürliche Weise zeigt, zu untersuchen. EMA beinhaltet die wiederholte Erfassung von Verhalten (z. B. Nahrungsaufnahme), Erfahrungen (z. B. Stress) und/oder physiologischen Parametern innerhalb eines Tages über mehrere Tage hinweg im Alltag.

Hierdurch wird die Untersuchung komplexer psychologischer, verhaltensbezogener und/oder physiologischer Prozesse im Alltag ermöglicht (Smyth & Stone, 2003). Darüber hinaus erlaubt EMA es, enge zeitliche Zusammenhänge zwischen Stress und der Nahrungsaufnahme über Minuten und Stunden hinweg zu untersuchen. Obwohl EMA einen vielversprechenden Ansatz für die Untersuchung des Zusammenhangs von Stress und der Nahrungsaufnahme im Alltag darstellt, wurde es bislang in nur wenigen Studien genutzt. Bisherige EMA-Studien weisen zudem zwei Limitationen auf: (1) Sie erfassen die Nahrungsaufnahme lediglich mittels subjektiver Selbsteinschätzung, z. B. von 0 – *zu wenig gegessen* bis 100 – *zu viel gegessen* (Reichenberger et al., 2021), oder erfassen ausschließlich einzelne Aspekte der Nahrungsaufnahme, z. B. den Konsum von Snacks (Zenk et al., 2014). So untersucht bislang keine Studie den Zusammenhang von Stress und der tatsächlichen Nahrungsaufnahme in einer gesunden Stichprobe mittels EMA. Die tatsächliche Nahrungsaufnahme bezieht sich auf die Erfassung aller verzehrten Lebensmittel und Getränke sowie der verzehrten Mengen, die nachfolgend zur Ermittlung von Nährwerten (z. B. Energie und Makronährstoffe) verwendet werden. (2) EMA ermöglicht es, zu untersuchen, ob das Stressempfinden, welches mehrmals am Tag erfasst wird, mit der Nahrungsaufnahme innerhalb eines festgelegten Zeitintervalls (z. B. innerhalb der nächsten Stunde) assoziiert ist. In der Regel erfolgt jedoch nicht in jedem dieser Zeitintervalle eine Nahrungsaufnahme oder es werden nur geringe Mengen (z. B. ein Snack) verzehrt. Infolgedessen ergibt sich eine Kriteriumsvariable, die zahlreiche Nullen sowie rechtsschiefe positive Werte enthält. Diese Art von Variablen wird häufig als semikontinuierlich bezeichnet und kann nicht mittels traditioneller linearer Mehrebenenmodelle analysiert werden, da die Annahme normalverteilter Residuen mit großer Wahrscheinlichkeit verletzt ist. Bisherige EMA-Studien nutzten für die Analyse der semikontinuierlichen Kriteriumsvariable ungeeignete statistische Verfahren, die sich entweder auf den Einfluss von Stress auf das Auftreten von Nahrungsaufnahme, d. h. *ob* gegessen wird (z. B. Zenk et al., 2014), oder auf die Menge der Nahrungsaufnahme, d. h. *wie viel* gegessen wird (z. B. Reichenberger et al., 2021), beschränken. Beide Analyseansätze sind jedoch mit Informationsverlust verbunden. Basierend auf den Limitationen bisheriger EMA-Studien bedarf es (1) eines EMA-Instruments, das die Erfassung von Stress und tatsächlicher Nahrungsaufnahme erlaubt, sowie (2) eines statistischen Verfahrens, das die semikontinuierliche Kriteriumsvariable angemessen berücksichtigt. Nur so kann das Verständnis der Beziehung zwischen Stress und der Nahrungsaufnahme im Alltag vertieft und Personen mit erhöhtem Risiko für stressbedingte Nahrungsaufnahme im Alltag identifiziert werden.

3.2. Ziele der vorliegenden Arbeit

Nur wenn eine valide Erfassung der tatsächlichen Nahrungsaufnahme sowie ein auf die semikontinuierlichen Daten angepasstes statistisches Verfahren integriert werden, ist eine reliable Untersuchung des Zusammenhangs von Stress und der tatsächlichen Nahrungsaufnahme im Alltag möglich. Aus diesem Grund verfolgte die vorliegende Dissertation drei Ziele: Das erste Ziel stellte die Entwicklung und Evaluation eines EMA-Instruments zur Erfassung von Stress und der tatsächlichen Nahrungsaufnahme dar (Publikation 1). Das zweite Ziel der Dissertation war es, ein fortschrittliches, bisher jedoch größtenteils unbekanntes statistisches Verfahren für semikontinuierliche Daten einzuführen und eine praktische Anleitung für die Modellanwendung bereitzustellen (Publikation 2). Das neu entwickelte und evaluierte EMA-Instrument sowie das für semikontinuierliche Variablen geeignete statistische Verfahren bildeten die Grundlage für das dritte Ziel der vorliegenden Dissertation. Das EMA-Instrument und das statistische Verfahren wurden integriert, um neue Erkenntnisse hinsichtlich individueller Unterschiede in der Nahrungsaufnahme bei Stress im Alltag zu gewinnen (Publikation 3).

3.3. Befunde der vorliegenden Arbeit

3.3.1. Publikation 1: Entwicklung und Evaluation eines EMA-Instruments

Das smartphonebasierte EMA-Instrument, die sogenannte APPetite-Mobile-App, wurde zur Erfassung von Stress und der tatsächlichen Nahrungsaufnahme entwickelt. Die App erfasst das Stressempfinden achtmal pro Tag sowie die tatsächliche Nahrungsaufnahme mittels eines integrierten Ernährungsprotokolls. Die Durchführbarkeit und Benutzerfreundlichkeit der App sowie die Validität des Ernährungsprotokolls wurden empirisch untersucht. Die Ergebnisse zeigen, dass die APPetite-Mobile-App insgesamt ein geeignetes Instrument zur Erfassung der tatsächlichen Nahrungsaufnahme sowie von Stress im Alltag darstellt. Jedoch zeigten sich ebenfalls Herausforderungen. So ist das Protokollieren aller Nahrungsmittel mit hohem Aufwand für die Teilnehmerinnen und Teilnehmer der Studie verbunden. Zudem geht die Weiterverarbeitung der Ernährungsdaten mit hohem zeitlichem Aufwand für die Forschenden einher. Dies verdeutlicht die Notwendigkeit der technischen Weiterentwicklung von Geräten und Software, die die passive Erkennung von Nahrungsaufnahme und die automatisierte Erfassung der tatsächlichen Nahrungsaufnahme mittels Fotos ermöglichen.

3.3.2. Publikation 2: Einführung eines geeigneten statistischen Verfahrens für semikontinuierliche Ernährungsdaten

Ein vielversprechendes statistisches Verfahren für semikontinuierliche Daten stellen zweiteilige Mehrebenenmodelle (engl. *multilevel two-part models)* dar. Sie untersuchen semikontinuierliche Variablen mittels zweier Modellkomponenten: (1) Der Modellteil der Nullen (engl. *zero part*) untersucht, ob Stress einen Einfluss auf das Auftreten von Nahrungsaufnahme hat (d. h. *ob* eine Person isst). (2) Der kontinuierliche Modellteil untersucht, ob Stress einen Einfluss auf die Menge der Nahrungsaufnahme hat (d. h. *wie viel* eine Person isst, wenn sie isst). Zweiteilige Mehrebenenmodelle erlauben es, nicht nur diese zwei Modellteile, sondern ebenfalls ihre Abhängigkeit anhand der Korrelation zwischen den Modellteilen (engl. *cross-part correlation*) zu untersuchen. Die modellteilspezifischen Befunde in Bezug auf das Auftreten und die Menge der Nahrungsaufnahme sowie die mittleren bis hohen Korrelationen zwischen den Modellteilen der Beispielanalyse zeigen, dass zweiteilige Mehrebenenmodelle nicht nur methodisch sondern auch konzeptuell vielversprechend für die Analyse semikontinuierlicher Ernährungsdaten sind.

3.3.3. Publikation 3: Individuelle Unterschiede in der Nahrungsaufnahme bei Stress

Zur Untersuchung, (1) ob sich individuelle Unterschiede in der Nahrungsaufnahme bei Stress auch im Alltag zeigen, (2) ob Geschlecht, Alter, BMI, Trait Stressessen und Essstile die Beziehung zwischen Stress und der tatsächlichen Nahrungsaufnahme moderieren und (3) ob diese Ergebnisse das „individual-difference model" stützen, wurden Daten, die mithilfe der APPetite-Mobile-App (Publikation 1) gesammelt wurden, mittels zweiteiliger Mehrebenenmodelle (Publikation 2) ausgewertet. Im Modellteil der Nullen moderierten Geschlecht, Alter, BMI, Trait Stressessen und Essstile die Beziehung zwischen Stress und dem Auftreten von Nahrungsaufnahme nicht. Im kontinuierlichen Modellteil moderierte ausschließlich Geschlecht die Beziehung zwischen Stress und der Menge der Nahrungsaufnahme. Demnach aßen Männer bei steigendem Stress signifikant weniger, während Frauen keine relevante Veränderung zeigten.

Überraschenderweise legen die zufälligen Effekte der Modelle geringe individuelle Unterschiede in der Nahrungsaufnahme bei Stress nahe. Demzufolge scheinen Personen nicht immer die gleiche Nahrungsaufnahme bei Stress zu zeigen. Intraindividuelle Variabilität könnte individuelle Unterschiede verdecken und situative Faktoren (z. B. Stimmung) könnten eine moderierende Rolle spielen. Diese Schlussfolgerung wird durch erste Befunde, die zeigen, dass situative Faktoren (z. B. einfache Verfügbarkeit von Essen) die Beziehung zwischen Stress und der Nahrungsaufnahme moderieren (Zenk et al., 2014), gestützt. Basierend darauf schlägt die vorliegende Dissertation eine Erweiterung des „individual-difference models" vor, die situative Faktoren, die die Beziehung zwischen Stress und der Nahrungsaufnahme moderieren könnten, berücksichtigt. Laut des erweiterten Modells sollte das Ziel nicht nur sein, Personen zu identifizieren, die ein höheres Risiko für stressbedingte

Nahrungsaufnahme aufweisen, sondern auch Situationen mit erhöhtem Risiko für stressbedingte Nahrungsaufnahme.

3.4. Fazit

Die vorliegende Dissertation eröffnet neue Einblicke in die Beziehung zwischen Stress und der Nahrungsaufnahme im Alltag, indem sie (1) ein EMA-Instrument, das die Untersuchung der Beziehung zwischen Stress und der tatsächlichen Nahrungsaufnahme im Alltag ermöglicht, entwickelt und evaluiert, und (2) zweiteilige Mehrebenenmodelle einführt und anwendet, um den Einfluss von Stress auf das Auftreten sowie die Menge der Nahrungsaufnahme zu untersuchen. Da die Beziehung zwischen Stress und der tatsächlichen Nahrungsaufnahme im Alltag mittels zweiteiliger Mehrebenenmodelle erstmalig im Rahmen der vorliegenden Arbeit untersucht wurde, sind weitere Studien notwendig, um die Befunde zu replizieren und die Gültigkeit des erweiterten Modells zu prüfen. Die vorliegende Arbeit weist einige Limitationen auf, die bei der Interpretation der Ergebnisse berücksichtigt werden müssen. So wurde beispielswiese der Zusammenhang von Stress und der tatsächlichen Nahrungsaufnahme über einen Zeitraum von lediglich drei Tagen erfasst. Hauptgrund dafür war der hohe zeitliche Aufwand, der mit der Ernährungserfassung für die Teilnehmerinnen und Teilnehmer der Studie und der Ernährungsauswertung für die Forschenden einherging. Verbesserungen der Ernährungserfassung, die längere EMA-Erfassungen sowie größere Stichproben erlauben, sind daher unerlässlich. Nichtsdestotrotz liefert die vorliegende Arbeit methodische sowie inhaltliche Ansatzpunkte, die dazu beitragen können, Fortschritte auf dem Forschungsgebiet anzuregen. Diese Fortschritte sind notwendig, um das Verständnis der Beziehung zwischen Stress und der tatsächlichen Nahrungsaufnahme im Alltag weiter zu vertiefen und spezifische Situationen und/oder Individuen mit besonderem Risiko für gesundheitsschädliche Veränderungen der Nahrungsaufnahme bei Stress zu identifizieren.

References

Ackenheil, M., Stotz, G., & Dietz-Bauer, R. (1999). *Mini International Neuropsychiatric Interview. German Version 5.0.0, DSM-IV*. Psychiatrische Universitätsklinik München.

Adan, R. A. H., van der Beek, E. M., Buitelaar, J. K., Cryan, J. F., Hebebrand, J., Higgs, S., Schellekens, H., & Dickson, S. L. (2019). Nutritional psychiatry: Towards improving mental health by what you eat. *European Neuropsychopharmacology*, *29*(12), 1321–1332. https://doi.org/10.1016/j.euroneuro.2019.10.011

Afshin, A., Peñalvo, J. L., Del Gobbo, L., Silva, J., Michaelson, M., O'Flaherty, M., Capewell, S., Spiegelman, D., Danaei, G., & Mozaffarian, D. (2017). The prospective impact of food pricing on improving dietary consumption: A systematic review and meta-analysis. *PLOS ONE*, *12*(3), e0172277. https://doi.org/10.1371/journal.pone.0172277

Afshin, A., Sur, P. J., Fay, K. A., Cornaby, L., Ferrara, G., Salama, J. S., Mullany, E. C., Abate, K. H., Abbafati, C., Abebe, Z., Afarideh, M., Aggarwal, A., Agrawal, S., Akinyemiju, T., Alahdab, F., Bacha, U., Bachman, V. F., Badali, H., Badawi, A., … Murray, C. J. L. (2019). Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, *393*(10184), 1958–1972. https://doi.org/10.1016/S0140-6736(19)30041-8

Araiza, A. M., & Lobel, M. (2018). Stress and eating: Definitions, findings, explanations, and implications. *Social and Personality Psychology Compass*, *12*(4), e12378. https://doi.org/10.1111/spc3.12378

Aronne, L. J., Nelinson, D. S., & Lillo, J. L. (2009). Obesity as a disease state: A new paradigm for diagnosis and treatment. *Clinical Cornerstone*, *9*(4), 9–29. https://doi.org/10.1016/S1098-3597(09)80002-1

Ashman, A. M., Collins, C. E., Brown, L. J., Rae, K. M., & Rollo, M. E. (2017). Validation of a smartphone image-based dietary assessment method for pregnant women. *Nutrients*, *9*(1), 1–17. https://doi.org/10.3390/nu9010073

Baldwin, S. A., Fellingham, G. W., & Baldwin, A. S. (2016). Statistical models for multilevel skewed physical activity data in health research and behavioral medicine. *Health Psychology*, *35*(6), 552–562. https://doi.org/10.1037/hea0000292

Brooke, J. (1996). SUS: a "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, A. L. McClelland, & B. A. Weerdmeester (Eds.), *Usability Evaluation in Industry* (pp. 189–194). Taylor and Francis.

Bürkner, P.-C. (2017). brms : An R Package for Bayesian Multilevel Models Using Stan.

Journal of Statistical Software, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, *10*(1), 395. https://doi.org/10.32614/RJ-2018-017

Chen, J., Berkman, W., Bardouh, M., Ng, C. Y. K., & Allman-Farinelli, M. (2019). The use of a food logging app in the naturalistic setting fails to provide accurate measurements of nutrients and poses usability challenges. *Nutrition*, *57*, 208–216. https://doi.org/10.1016/j.nut.2018.05.003

Chmitorz, A., Neumann, R. J., Kollmann, B., Ahrens, K. F., Öhlschläger, S., Goldbach, N., Weichert, D., Schick, A., Lutz, B., Plichta, M. M., Fiebach, C. J., Wessa, M., Kalisch, R., Tüscher, O., Lieb, K., & Reif, A. (2021). Longitudinal determination of resilience in humans to identify mechanisms of resilience to modern-life stressors: the longitudinal resilience assessment (LORA) study. *European Archives of Psychiatry and Clinical Neuroscience*, *271*, 1035–1051. https://doi.org/10.1007/s00406-020-01159-2

Conner, M., Fitter, M., & Fletcher, W. (1999). Stress and snacking: A diary study of daily hassles and between-meal snacking. *Psychology & Health*, *14*(1), 51–63. https://doi.org/10.1080/08870449908407313

Dunton, G. F. (2018). Sustaining Health-Protective Behaviors Such as Physical Activity and Healthy Eating. *JAMA*, *320*(7), 639–640. https://doi.org/10.1001/jama.2018.6621

Epel, E., Lapidus, R., McEwen, B., & Brownell, K. (2001). Stress may add bite to appetite in women: a laboratory study of stress-induced cortisol and eating behavior. *Psychoneuroendocrinology*, *26*(1), 37–49. https://doi.org/10.1016/S0306-4530(00)00035-4

Evers, C., Dingemans, A., Junghans, A. F., & Boevé, A. (2018). Feeling bad or feeling good, does emotion affect your consumption of food? A meta-analysis of the experimental evidence. *Neuroscience & Biobehavioral Reviews*, *92*, 195–208. https://doi.org/10.1016/j.neubiorev.2018.05.028

Ferrara, G., Kim, J., Lin, S., Hua, J., & Seto, E. (2019). A Focused Review of Smartphone Diet-Tracking Apps: Usability, Functionality, Coherence With Behavior Change Theory, and Comparative Validity of Nutrient Intake and Energy Estimates. *JMIR MHealth and UHealth*, *7*(5), e9232. https://doi.org/10.2196/mhealth.9232

Greeno, C. G., & Wing, R. R. (1994). Stress-induced eating. *Psychological Bulletin*, *115*(3), 444–464. https://doi.org/10.1037/0033-2909.115.3.444

Grunberg, N. E., & Straub, R. O. (1992). The role of gender and taste class in the effects of stress on eating. *Health Psychology*, *11*(2), 97–100. https://doi.org/10.1037/0278-

6133.11.2.97

Higgs, S. (2015). Social norms and their influence on eating behaviours. *Appetite*, *86*, 38–44. https://doi.org/10.1016/j.appet.2014.10.021

Hill, D., Conner, M., Clancy, F., Moss, R., Wilding, S., Bristow, M., & O'Connor, D. B. (2021). Stress and eating behaviours in healthy adults: a systematic review and meta-analysis. *Health Psychology Review*, 1–25. https://doi.org/10.1080/17437199.2021.1923406

Inada, S., Iizuka, Y., Ohashi, K., Kikuchi, H., Yamamoto, Y., Kadowaki, T., & Yoshiuchi, K. (2019). Preceding psychological factors and calorie intake in patients with type 2 diabetes: investigation by ecological momentary assessment. *BioPsychoSocial Medicine*, *13*(1), 20. https://doi.org/10.1186/s13030-019-0161-4

Kelly, S. J., Daniel, M., Dal Grande, E., & Taylor, A. (2011). Mental ill-health across the continuum of body mass index. *BMC Public Health*, *11*(1), 765. https://doi.org/10.1186/1471-2458-11-765

Kivimäki, M., & Steptoe, A. (2018). Effects of stress on the development and progression of cardiovascular disease. *Nature Reviews Cardiology*, *15*(4), 215–229. https://doi.org/10.1038/nrcardio.2017.189

Koch, S. A. J., Conrad, J., Hierath, L., Hancock, N., Beer, S., Cade, J. E., & Nöthlings, U. (2020). Adaptation and evaluation of myfood24-germany: A web-based self-administered 24-h dietary recall for the german adult population. *Nutrients*, *12*(1), 1–15. https://doi.org/10.3390/nu12010160

Lecrubier, Y., Sheehan, D. V, Weiller, E., Amorim, P., Bonora, I., Sheehan, K. H., Janavs, J., & Dunbar, G. C. (1997). The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *European Psychiatry*, *12*(5), 224–231. https://doi.org/10.1016/S0924-9338(97)83296-8

Liu, L., Ma, J. Z., & Johnson, B. A. (2008). A multi-level two-part random effects model, with application to an alcohol-dependence study. *Statistics in Medicine*, *27*(18), 3528–3539. https://doi.org/10.1002/sim.3205

Lopes, T. S., Luiz, R. R., Hoffman, D. J., Ferriolli, E., Pfrimer, K., Moura, A. S., Sichieri, R., & Pereira, R. A. (2016). Misreport of energy intake assessed with food records and 24-h recalls compared with total energy expenditure estimated with DLW. *European Journal of Clinical Nutrition*, *70*(11), 1259–1264. https://doi.org/10.1038/ejcn.2016.85

Maugeri, A., & Barchitta, M. (2019). A systematic review of ecological momentary assessment of diet: Implications and perspectives for nutritional epidemiology. *Nutrients*, *11*(11), 2696. https://doi.org/10.3390/nu11112696

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., de Ruiter, S., Ebner-Priemer, U. W., Kieser, M., Ortega, F. B., Thome, J., Buitelaar, J. K., … 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

Mente, A., de Koning, L., Shannon, H. S., & Anand, S. S. (2009). A Systematic Review of the Evidence Supporting a Causal Link Between Dietary Factors and Coronary Heart Disease. *Archives of Internal Medicine*, *169*(7), 659–669. https://doi.org/10.1001/archinternmed.2009.38

Meule, A., Reichenberger, J., & Blechert, J. (2018). Development and preliminary validation of the Salzburg Stress Eating Scale. *Appetite*, *120*, 442–448. https://doi.org/10.1016/j.appet.2017.10.003

Milaneschi, Y., Simmons, W. K., van Rossum, E. F. C., & Penninx, B. W. (2019). Depression and obesity: evidence of shared biological mechanisms. *Molecular Psychiatry*, *24*(1), 18–33. https://doi.org/10.1038/s41380-018-0017-5

O'Connor, D. B., Jones, F., Conner, M., McMillan, B., & Ferguson, E. (2008). Effects of daily hassles and eating style on eating behavior. *Health Psychology*, *27*(1, Suppl), S20–S31. https://doi.org/10.1037/0278-6133.27.1.S20

O'Connor, D. B., Thayer, J. F., & Vedhara, K. (2021). Stress and Health: A Review of Psychobiological Processes. *Annual Review of Psychology*, *72*(1), 663–688. https://doi.org/10.1146/annurev-psych-062520-122331

Pfennig, A., Leopold, K., Martini, J., Boehme, A., Lambert, M., Stamm, T., Bermpohl, F., Reif, A., Kittel-Schneider, S., Juckel, G., Fallgatter, A. J., Kircher, T., Jansen, A., Pfeiffer, S., Berndt, C., Rottmann-Wolf, M., Sauer, C., Ritter, P., Correll, C. U., … Bauer, M. (2020). Improving early recognition and intervention in people at increased risk for the development of bipolar disorder: study protocol of a prospective-longitudinal, naturalistic cohort study (Early-BipoLife). *International Journal of Bipolar Disorders*, *8*(1). https://doi.org/10.1186/s40345-020-00183-4

Pollard, T. M., Steptoe, A., Canaan, L., Davies, G. J., & Wardle, J. (1995). Effects of academic examination stress on eating behavior and blood lipid levels. *International Journal of Behavioral Medicine*, *2*(4), 299.

https://doi.org/10.1207/s15327558ijbm0204_2

Pudel, V., & Westenhöfer, J. (1989). *Fragebogen zum Eßverhalten*. Hogrefe.

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

Rajan, T., & Menon, V. (2017). Psychiatric disorders and obesity: A review of association studies. *Journal of Postgraduate Medicine*, *63*(3), 182–190. https://doi.org/10.4103/jpgm.JPGM_712_16

Reichenberger, J., Kuppens, P., Liedlgruber, M., Wilhelm, F. H., Tiefengrabner, M., Ginzinger, S., & Blechert, J. (2018). No haste, more taste: An EMA study of the effects of stress, negative and positive emotions on eating behavior. *Biological Psychology*, *131*, 54–62. https://doi.org/10.1016/J.BIOPSYCHO.2016.09.002

Reichenberger, J., Pannicke, B., Arend, A.-K., Petrowski, K., & Blechert, J. (2021). Does stress eat away at you or make you eat? EMA measures of stress predict day to day food craving and perceived food intake as a function of trait stress-eating. *Psychology & Health*, *36*(2), 129–147. https://doi.org/10.1080/08870446.2020.1781122

Ritter, P. S., Bermpohl, F., Gruber, O., Hautzinger, M., Jansen, A., Juckel, G., Kircher, T., Lambert, M., Mulert, C., Pfennig, A., Reif, A., Rienhoff, O., Schulze, T. G., Severus, E., Stamm, T., & Bauer, M. (2016). Aims and structure of the German Research Consortium BipoLife for the study of bipolar disorder. *International Journal of Bipolar Disorders*, *4*(1). https://doi.org/10.1186/s40345-016-0066-0

Schultchen, D., Reichenberger, J., Mittl, T., Weh, T. R. M., Smyth, J. M., Blechert, J., & Pollatos, O. (2019). Bidirectional relationship of stress and affect with physical activity and healthy eating. *British Journal of Health Psychology*, *24*(2), 315–333. https://doi.org/10.1111/bjhp.12355

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

Smyth, J. M., & Stone, A. A. (2003). Ecological Momentary Assessment Research in Behavioral medicine. *Journal of Happiness Studies*, *4*(1), 35–52. https://doi.org/10.1023/A:1023657221954

Spook, J. E., Paulussen, T., Kok, G., & Van Empelen, P. (2013). Monitoring dietary intake and physical activity electronically: Feasibility, usability, and ecological validity of a mobile-based ecological momentary assessment tool. *Journal of Medical Internet Research*, *15*(9), 1–13. https://doi.org/10.2196/jmir.2617

Stone, A. A., & Shiffman, S. (2002). Capturing momentary, self-report data: A proposal for reporting guidelines. *Annals of Behavioral Medicine*, *24*(3), 236–243. https://doi.org/10.1207/S15324796ABM2403_09

Stunkard, A. J., & Messick, S. (1985). The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. *Journal of Psychosomatic Research*, *29*(1), 71–83. https://doi.org/10.1016/0022-3999(85)90010-8

Su, L., Tom, B. D. M., & Farewell, V. T. (2009). Bias in 2-part mixed models for longitudinal semicontinuous data. *Biostatistics*, *10*(2), 374–389. https://doi.org/10.1093/biostatistics/kxn044

Swinburn, B. A., Kraak, V. I., Allender, S., Atkins, V. J., Baker, P. I., Bogard, J. R., Brinsden, H., Calvillo, A., De Schutter, O., Devarajan, R., Ezzati, M., Friel, S., Goenka, S., Hammond, R. A., Hastings, G., Hawkes, C., Herrero, M., Hovmand, P. S., Howden, M., … Dietz, W. H. (2019). The Global Syndemic of Obesity, Undernutrition, and Climate Change: The Lancet Commission report. *The Lancet*, *393*(10173), 791–846. https://doi.org/10.1016/S0140-6736(18)32822-8

Tomiyama, A. J. (2019). Stress and Obesity. *Annual Review of Psychology*, *70*(1), 703–718. https://doi.org/10.1146/annurev-psych-010418-102936

Tooze, J. A., Grunwald, G. K., & Jones, R. H. (2002). Analysis of repeated measures data with clumping at zero. *Statistical Methods in Medical Research*, *11*(4), 341–355. https://doi.org/10.1191/0962280202sm291ra

Torres, S. J., & Nowson, C. A. (2007). Relationship between stress, eating behavior, and obesity. *Nutrition*, *23*(11), 887–894. https://doi.org/10.1016/j.nut.2007.08.008

Trabulsi, J., & Schoeller, D. A. (2001). Evaluation of dietary assessment instruments against doubly labeled water, a biomarker of habitual energy intake. *American Journal of Physiology - Endocrinology and Metabolism*, *281*(5), E891–E899. https://doi.org/10.1152/ajpendo.2001.281.5.e891

Viechtbauer, W. (2022). Statistical methods for ESM data. In I. Myin-Germeys & P. Kuppens (Eds.), *The open handbook of Experience Sampling Methodology: A step-by-step guide to designing, conducting, and analyzing ESM studies* (2nd ed., pp. 153–183). Center for Research on Experience Sampling and Ambulatory Methods Leuven.

Wallis, D. J., & Hetherington, M. M. (2004). Stress and eating: the effects of ego-threat and cognitive demand on food intake in restrained and emotional eaters. *Appetite*, *43*(1), 39–46. https://doi.org/10.1016/j.appet.2004.02.001

Wardle, J., Steptoe, A., Oliver, G., & Lipsey, Z. (2000). Stress, dietary restraint and food intake. *Journal of Psychosomatic Research*, *48*(2), 195–202. https://doi.org/10.1016/S0022-3999(00)00076-3

World Health Organization. (2021). *Obesity and overweight*. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

Wrzus, C., & Neubauer, A. B. (2022). Ecological Momentary Assessment: A Meta-Analysis on Designs, Samples, and Compliance Across Research Fields. *Assessment*. https://doi.org/10.1177/10731911211067538

Yau, Y. H. C., & Potenza, M. N. (2013). Stress and eating behaviors. *Minerva Endocrinologica*, *38*(3), 255–267. https://pubmed.ncbi.nlm.nih.gov/24126546

Zenk, S. N., Horoi, I., McDonald, A., Corte, C., Riley, B., & Odoms-Young, A. M. (2014). Ecological momentary assessment of environmental and personal factors and snack food intake in African American women. *Appetite*, *83*, 333–341. https://doi.org/10.1016/j.appet.2014.09.008

Appendix

…………………………………………………………………………………………………...

Paper 1

Ruf, A., Koch, E.D., Ebner-Priemer, U., Knopf, M., Reif, A., & Matura, S. (2021). Studying Microtemporal, Within-Person Processes of Diet, Physical Activity, and Related Factors Using the APPetite-Mobile-App: Feasibility, Usability, and Validation Study. *Journal of Medical Internet Research, 23*(7), e25850. https://doi.org/10.2196/25850

…………………………………………………………………………………………………...

Original Paper

Studying Microtemporal, Within-Person Processes of Diet, Physical Activity, and Related Factors Using the APPetite-Mobile-App: Feasibility, Usability, and Validation Study

Alea Ruf[1], BSc, MSc; Elena Doris Koch[2], BA, MA; Ulrich Ebner-Priemer[2,3], BA, MA, PhD, Prof Dr; Monika Knopf[4], BA, MA, PhD, Prof Dr; Andreas Reif[1], Dr med, Prof Dr; Silke Matura[1], BA, MA, Dr rer med

[1]Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital, Goethe University, Frankfurt, Germany

[2]Mental mHealth Lab, Institute of Sports and Sports Science, Karlsruhe Institute of Technology, Karlsruhe, Germany

[3]Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

[4]Department of Developmental Psychology, Goethe University, Frankfurt, Germany

Corresponding Author:
Alea Ruf, BSc, MSc
Department of Psychiatry, Psychosomatic Medicine and Psychotherapy
University Hospital
Goethe University
Heinrich-Hoffmann-Straße 10
Frankfurt, 60528
Germany
Phone: 49 69 6301 83348
Email: alea.ruf@kgu.de

Abstract

Background: Diet and physical activity (PA) have a major impact on physical and mental health. However, there is a lack of effective strategies for sustaining these health-protective behaviors. A shift to a microtemporal, within-person approach is needed to capture dynamic processes underlying eating behavior and PA, as they change rapidly across minutes or hours and differ among individuals. However, a tool that captures these microtemporal, within-person processes in daily life is currently not present.

Objective: The APPetite-mobile-app is developed for the ecological momentary assessment of microtemporal, within-person processes of complex dietary intake, objectively recorded PA, and related factors. This study aims to evaluate the feasibility and usability of the APPetite-mobile-app and the validity of the incorporated APPetite-food record.

Methods: The APPetite-mobile-app captures dietary intake event-contingently through a food record, captures PA continuously through accelerometers, and captures related factors (eg, stress) signal-contingently through 8 prompts per day. Empirical data on feasibility (n=157), usability (n=84), and validity (n=44) were collected within the Eat2beNICE-APPetite-study. Feasibility and usability were examined in healthy participants and psychiatric patients. The relative validity of the APPetite-food record was assessed with a subgroup of healthy participants by using a counterbalanced crossover design. The reference method was a 24-hour recall. In addition, the energy intake was compared with the total energy expenditure estimated from accelerometry.

Results: Good feasibility, with compliance rates above 80% for prompts and the accelerometer, as well as reasonable average response and recording durations (prompt: 2.04 min; food record per day: 17.66 min) and latencies (prompts: 3.16 min; food record: 58.35 min) were found. Usability was rated as moderate, with a score of 61.9 of 100 on the System Usability Scale. The evaluation of validity identified large differences in energy and macronutrient intake between the two methods at the group and individual levels. The APPetite-food record captured higher dietary intakes, indicating a lower level of underreporting, compared with the 24-hour recall. Energy intake was assessed fairly accurately by the APPetite-food record at the group level on 2 of 3 days when compared with total energy expenditure. The comparison with mean total energy expenditure (2417.8 kcal, SD 410) showed that the 24-hour recall (1909.2 kcal, SD 478.8) underestimated habitual energy intake to a larger degree than the APPetite-food record (2146.4 kcal, SD 574.5).

Conclusions: The APPetite-mobile-app is a promising tool for capturing microtemporal, within-person processes of diet, PA, and related factors in real time or near real time and is, to the best of our knowledge, the first of its kind. First evidence supports the good feasibility and moderate usability of the APPetite-mobile-app and the validity of the APPetite-food record. Future findings in this context will build the foundation for the development of personalized lifestyle modification interventions, such as just-in-time adaptive interventions.

(J Med Internet Res 2021;23(7):e25850) doi: 10.2196/25850

KEYWORDS

diet; physical activity; microtemporal processes; within-person factors; ecological momentary assessment; smartphone-app; mobile phone; mHealth; dietary assessment; feasibility; usability; validity

Introduction

Background

Diet is a key contributor to both physical and mental health. Elevated BMI is a major risk factor for noncommunicable diseases, such as cardiovascular diseases [1]. Since 1975, the prevalence of obesity has nearly tripled globally [1]. Accordingly, in 2016, 13% of adults were obese and 39% were overweight [1]. Approximately 11 million deaths were associated with dietary risk factors (eg, low intake of whole grains) across 195 countries in 2017 [2]. Although the link between diet and mental health is not equally well understood, first evidence supports the presence of a direct association among diet, mental health, and mental functioning [3]. Obesity not only increases the probability of somatic diseases but also of mental illness, particularly depression [4-6]. These numbers and findings highlight the growing need to understand the “causes of the causes.”

Although factors and processes underlying eating behavior have been studied for many years [7,8], interventions remain ineffective in sustaining health-protective behaviors for the long term [9]. One reason for this could be the main focus on between-person characteristics (eg, age) and macrotemporal processes (across weeks, months, or years) [10]. Diet is a highly complex health behavior that is performed multiple times per day and is influenced by a variety of fluctuating factors and their interactions [11]. A real-life microtimescale approach is needed to capture the dynamics of diet and associated factors ecologically and momentarily and, ultimately, to understand the processes underlying eating behavior in everyday life [10]. In contrast to some between-person characteristics (eg, age), within-person factors are modifiable and therefore a promising target for interventions. For this reason, the identification of within-person factors that influence eating behavior in daily life is needed for the development of novel, more effective, and personalized interventional approaches.

It is not only diet that has a large impact on both physical and mental health. Physical activity (PA) represents another impactful, repeated-occurrence health behavior [12,13]. To untangle the complex association between diet and health [14], it is important to consider possible interactions. For instance, diet does not independently regulate body weight. Body weight is regulated through the interplay of energy intake (ie, diet) and energy expenditure (eg, PA) [15]. Therefore, the assessment of microtemporal, within-person processes of diet and PA should be combined, and possible interactions should be taken into consideration.

Ecological Momentary Assessment of Diet, PA, and Related Factors

The repeated or continuous assessment of experiences, behaviors, or physiological processes in real life through smartphones or wearable devices is a highly promising approach for studying microtemporal, within-person processes [16]. This approach is referred to as ecological momentary assessment (EMA), ambulatory assessment, experience sampling, and real-time data capture [17]. Although different terms have been used, they have in common the assessment of various parameters, multiple times per day in daily life [17].

Even though EMA studies do not allow causal conclusions, they offer insight into three important aspects of microtemporal, within-person processes: (1) temporal specificity (eg, Does diet influence mood to a greater extent than mood influences diet?), (2) situational specificity (eg, Is unhealthy eating more likely when being alone or with others?), and (3) person specificity (eg, Is stress more predictive for engaging in eating for some individuals compared with others?) [10].

Diet is a highly complex phenomenon that makes its assessment difficult. However, to avoid typical reporting biases that are present in traditional dietary assessment methods (eg, food frequency questionnaires), the number of studies using EMA to capture self-reported dietary intake or aspects of it in real time or near real time instead of retrospectively has rapidly grown in the last decade [18-20]. There are two categories of EMA approaches present so far: on the one hand, there are mobile-based dietary assessment tools that focus on the assessment of complex dietary intake and the generation of nutritional values. Complex dietary intake refers to assessing all consumed foods and drinks and consumed amounts, which are then used to generate nutritional values. Even though a small number of tools that assess complex dietary intake also allow assessing contextual correlates during eating occasions [21], no tool allows capturing a wider repertoire of factors preceding or succeeding eating occasions [22]. On the other hand, there are a number of studies that use EMA to study a variety of factors related to diet (eg, affect [23]). However, to the best of our knowledge, none of these studies assessed diet in its full complexity. Most of them focus on specific aspects of diet only, for example, snacks or sweetened beverages [24-29], a limited number of food and drink categories [23,30-33], portion sizes [34], or the type of eating events (main meals vs snacks) and the type of drinking occasions (alcoholic vs nonalcoholic) [35]. Hence, complex dietary intake was not assessed, and the generation of nutritional values was not possible. Although some of these studies reported a more comprehensive approach which captured all consumed foods and, in some studies, drinks through a free input field [23,33], the foods and drinks were only assigned to a limited number of food and drink categories and were not used to generate nutritional values. There is a need to study the processes underlying complex dietary intake instead of processes underlying only aspects of diet.

Despite the importance of taking possible interactions into account, most EMA studies focus on either the assessment of diet or PA. One study identified the need for an EMA tool to capture *complex lifestyle behavior*, that is, dietary intake and PA simultaneously [36]. However, the tool developed for this purpose failed to assess diet and PA in their complex nature. It only assessed specific food categories and used self-reports for the assessment of PA, which is unsatisfactory, given that 2 systematic reviews showed that indirect measures of PA (ie, self-reports) differ substantially from direct, objective measures (eg, accelerometers) [37,38].

In conclusion, there is a strong need for an EMA tool that allows capturing complex dietary intake, objectively measured PA, and a broad range of associated factors simultaneously in daily life to study microtemporal, within-person processes underlying these health-protective behaviors.

Objectives

As no EMA tool allows the study of microtemporal, within-person processes of complex dietary intake, objectively measured PA, and related factors, we developed an EMA tool for the simultaneous assessment of these complex health behaviors and related factors in daily life: the APPetite-mobile-app (this term also covers the assessment of PA, although it is not performed by the APPetite-mobile-app itself but by an accelerometer).

The suitability of novel EMA tools for use in daily life should be evaluated. Therefore, feasibility, usability, and validity were examined empirically in this study. The following questions will be addressed: Is the APPetite-mobile-app a feasible and usable tool for the combined assessment of complex dietary intake, PA, and associated factors in daily life and a valid tool for the assessment of complex dietary intake in real time or near real time?

Methods

The APPetite-Mobile-App

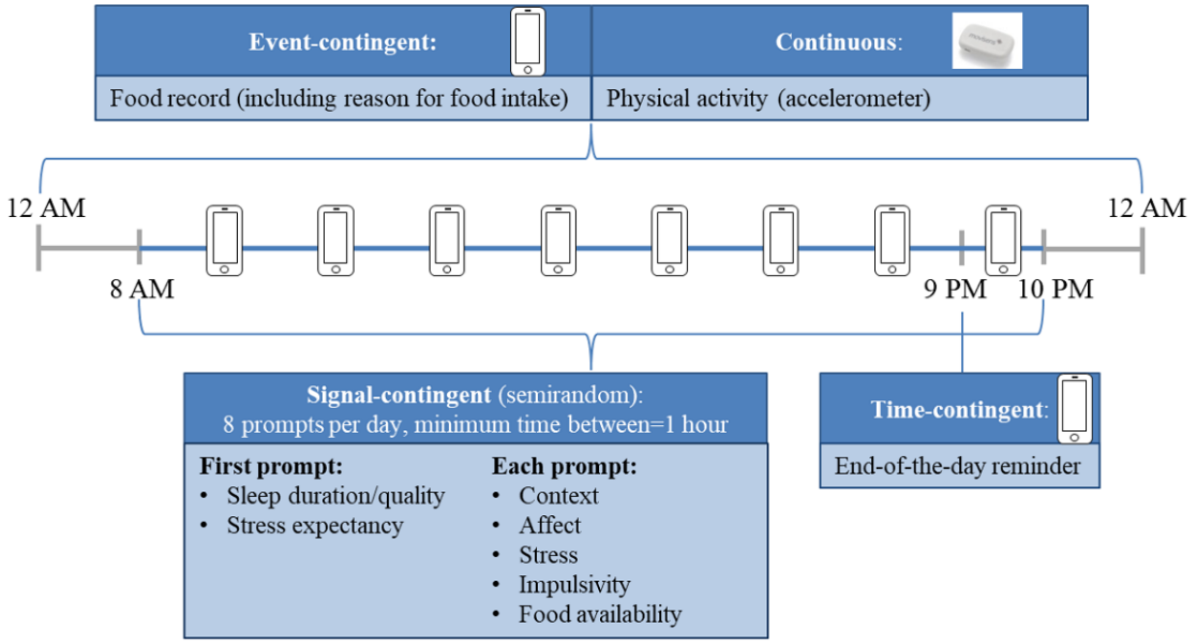
Software and Hardware

The APPetite-mobile-app was developed and run through movisensXS (version 1.4.7, movisens GmbH), a web-based platform for the development of EMA tools. It supports a broad range of sampling schemes, item formats, and multimedia records, allowing flexible and tailored study configurations. The APPetite-mobile-app is run through the movisensXS app (available for Android devices). If the mobile device has access to mobile data during the EMA assessment, participants' entries will be uploaded instantly to the platform. In this way, compliance can be monitored throughout the EMA assessment, and a chat function allows direct messaging with participants. All participants received a study smartphone (Motorola Moto G 3rd generation), with access to mobile data. The movisensXS app was previously tested on this particular mobile device, ensuring its smooth functioning and increasing the standardization of the mobile-based assessment.

Sampling Strategy

The APPetite-mobile-app uses event-, signal-, and time-contingent as well as continuous sampling (Figure 1).

Figure 1. Sampling strategy of the APPetite-mobile-app.



Food intake was recorded event-contingently through a food record. Participants were asked to enter foods and drinks as soon as possible after consuming them. Accordingly, participants are able to initiate the APPetite-food record at any time and capture their food intake in real time. This was chosen to minimize memory effects and record the exact time of food intake. In addition, this allows capturing food intake even during the night, when signal-contingent prompts are inappropriate. At 9 PM, a time-contingent prompt asks if all consumed foods and drinks of the day have been recorded, ensuring that no foods and drinks consumed on this day are missed.

The prompts are initiated signal-contingent at eight semirandom times per day between 8 AM and 10 PM. The minimum time between 2 prompts is 1 hour. Therefore, participants cannot predict the exact time of the next prompt, and the assessed situation is a better reflection of the participant's real life. Participants were instructed to respond immediately to the prompt. However, if participants are unable to reply instantly, it is possible to postpone the prompt for 5, 10, 15, 20, or 25 minutes to avoid missing data and reduce the participants' burden. If no reaction is registered, the prompt is deactivated and cannot be reactivated.

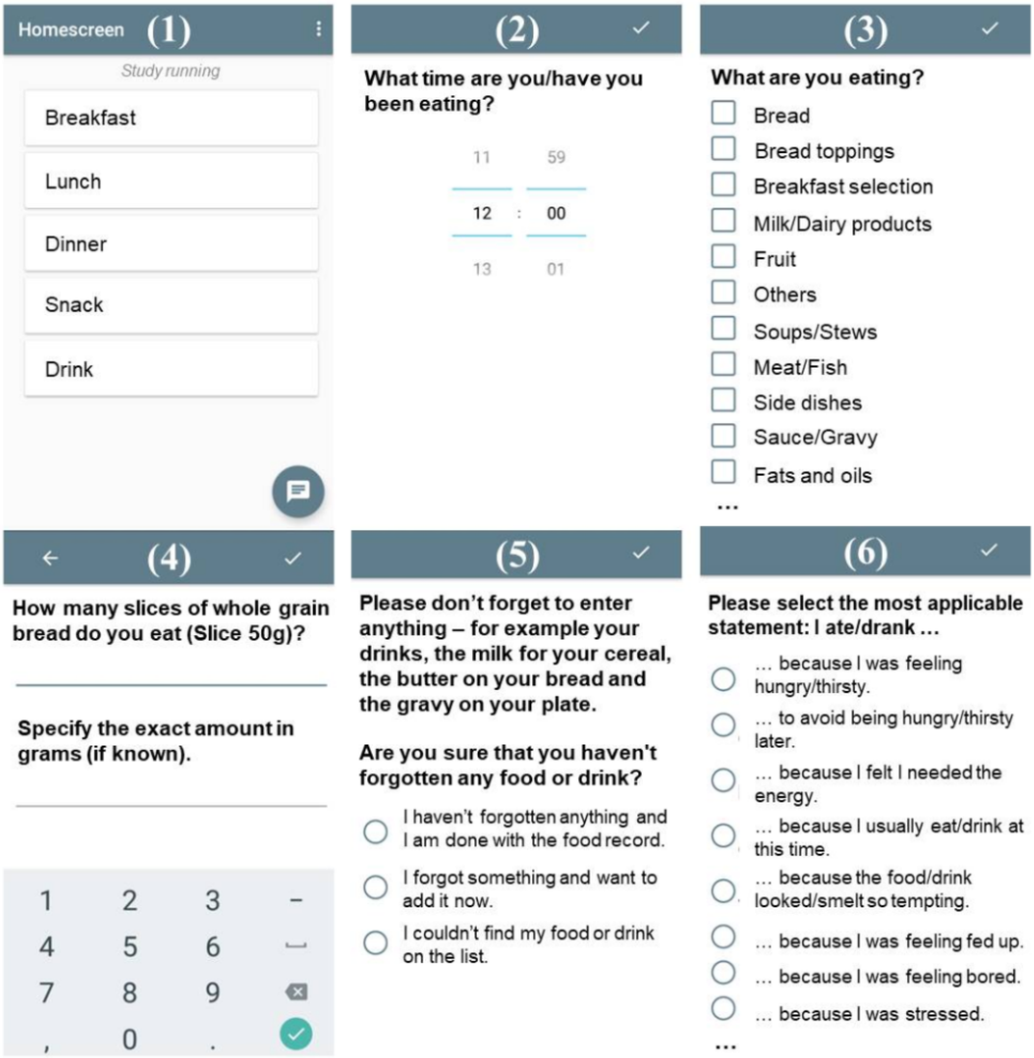
Continuous sampling through an accelerometer is used for the assessment of PA.

EMA Measures

APPetite-Food Record

The APPetite-food record comprises a 6-step process: (1) selection of meal type, (2) entry of time of intake, (3) selection of consumed foods and drinks, (4) specification of consumed amounts, (5) presentation of reminder for commonly forgotten foods, and (6) indication of predominant reason for eating or drinking (Figure 2 presents screenshots of the 6-step process). To generate nutritional values, the obtained dietary data were transferred by trained staff to myfood24-Germany, a 24-hour dietary recall [39]. A detailed description of the APPetite-food record and nutritional value generation is provided in Multimedia Appendix 1 [39-44]. All reasons for eating and drinking are presented in Multimedia Appendix 2 [40-44].

Figure 2. Screenshots of the 6-step process of the APPetite-food record.



Prompts

Each prompt assesses the context, affect, stress, impulsivity, and food availability either since the last prompt or immediately before the prompt. In addition, the first prompt of a day captures sleep quality and quantity as well as stress expectancy. All prompt measures and items are described in Multimedia Appendix 3 [45-49].

Physical Activity

Move 3 sensors from movisens were used to objectively record PA. The accelerometer was worn on the nondominant wrist.

Participants were asked to wear it at any time (also when sleeping) and only take it off when showering or performing water activities. The Move 3 sensor captures raw data on 3D acceleration, barometric air pressure, and temperature. Secondary parameters such as activity class, body position, steps, metabolic equivalents, and PA metrics can be extracted using the DataAnalyzer (movisens GmbH).

Evaluation of Feasibility

Measures

The feasibility of the APPetite-mobile-app was separately assessed for the EMA prompts, the APPetite-food record, and the accelerometer. The feasibility of the prompts is determined by prompt delivery, total number of answered prompts across all subjects, number of answered prompts per participant, compliance (percentage of complete prompts within received prompts), response latency (time from first prompt signal to answering), and the time needed to complete a single prompt. The food record's feasibility was evaluated based on the number of recorded eating and drinking events per day, reporting latency (time between the meal and meal recording), and the time needed to record food intake per day. The amount of time wearing the accelerometer and compliance (percentage wearing the accelerometer within the 7-day assessment period) are measures of feasibility of the accelerometer.

Sample

The data were collected within the ongoing APPetite study. The APPetite study is part of the European Union Horizon2020 project Eat2beNICE and recruits participants from three existing studies: LORA (Longitudinal Resilience Assessment) study [50], PROUD (Prevention of Comorbid Depression and Obesity in Attention-Deficit/Hyperactivity Disorder) study [51], and the BipoLife-A1 study (improving early recognition and intervention in people at risk of developing bipolar disorder [52,53]). The LORA study included individuals who were not affected by psychiatric conditions. The PROUD sample consisted of patients affected by attention-deficit/hyperactivity disorder. The BipoLife-A1 study follows up on patients with an increased risk for the development of bipolar disorder, including patients affected by attention-deficit/hyperactivity disorder or depression.

From November 2018 to March 2020, 161 participants were included in the APPetite study (140 LORA, 7 PROUD, and 14 BipoLife-A1). After the first in-person session, 3 LORA participants dropped out. Of these, 2 realized that they were unable to respond to prompts. The third person was mistakenly given a smartphone that was not coupled with the EMA protocol. Another person dropped out after the additional in-person session of the validation study for private reasons. Hence, EMA data of 157 participants are available for the evaluation of feasibility (see demographics in Table 1).

Table 1. Demographics of the total sample and the 3 cohorts (only individuals who completed the ecological momentary assessment were included; N=157).

Variables	Total (N=157)	LORA[a] (n=136)	PROUD[b] (n=7)	BipoLife-A1 (n=14)
Gender, n (%)				
Female	100 (63.7)	94 (69.1)	1 (14.3)	5 (35.7)
Male	57 (36.3)	42 (30.9)	6 (85.7)	9 (64.3)
Age (years), mean (SD)	28.04 (7.22)	28.08 (7.55)	26.43 (2.51)	28.5 (5.39)
BMI, mean (SD)	24.71 (4.81)	24.26 (3.87)	24.98 (6.11)	28.9 (9.13)

[a]LORA: Longitudinal Resilience Assessment.

[b]PROUD: Prevention of Comorbid Depression and Obesity in Attention-Deficit/Hyperactivity Disorder.

Procedure

The APPetite study consists of 2 in-person sessions, the EMA assessment, and a follow-up session from home. In the first in-person session, participants received detailed training on how to use the APPetite-mobile-app and the accelerometer. Participants received a smartphone with the APPetite-mobile-app and an accelerometer, including a wristband. Participants used the APPetite-mobile-app for 3 consecutive days (2 weekdays and 1 weekend day, not including the day of the first in-person session) and wore the accelerometer for 7 consecutive days (overlapping the 3 days of the APPetite-mobile-app assessment, not including the day of the first in-person session). During the 3 days of the app-based assessment, prompt compliance was tracked. If compliance fell below the threshold of 80%, a motivational message was sent to the participant. Participants who completed at least 80% (19/24) of the prompts were included in a raffle to win a €100 (US $121.74) voucher and a cooking class. Before the second in-person session, EMA data were checked, and questions regarding implausible prompt entries (eg, 8 AM as bedtime) and food records (eg, missing meals) were collected. These questions were reviewed in the second in-person session to resolve any uncertainties. Usability of the APPetite-mobile-app, reactivity, and representativity of the EMA assessment were assessed via questionnaires in the second in-person session.

Participants received €40 (US $48.7) after the second in-person session and €10 (US $12.17) after completing the follow-up. In addition, individual feedback on diet and PA was provided after the follow-up, which consisted of a web-based 24-hour recall from home.

Statistical Methods

Descriptive statistics were used to assess feasibility measures. We investigated whether compliance differed among the 3 cohorts, the 3 days, and between male and female participants. As compliance is not normally distributed, this is done using the following nonparametric tests: Kruskal-Wallis rank sum test, Friedman test, and Wilcoxon rank sum test. In addition, the Spearman rank correlation coefficient was calculated to investigate the association between compliance and age. The α level was set to .05. The analyses were performed using R 3.6.1 (R Core Team) with RStudio (RStudio, Public-benefit corporation).

Evaluation of Usability

Measures

Usability is assessed using the System Usability Scale (SUS; [54]), a commonly used questionnaire for the evaluation of websites or mobile apps. The questionnaire consists of 10 items. Each item represents a statement (eg, I thought the system was easy to use). Participants' agreement with the statement was rated on a 5-point scale. A total score between 0 and 100 was calculated. Higher numbers indicated better usability.

Sample

Data were collected within the APPetite study. However, SUS was subsequently added to this study (August 2019). Therefore, it is available only for a subsample of 84 participants (55 women and 29 men; 67 from LORA, 6 from PROUD, and 11 from BipoLife-A1). The mean age of the sample was 29.26 (SD 7.41) years, and the mean BMI was 24.82 (SD 5.26) kg/m^2.

Procedure

The SUS was completed during the second in-person session.

Statistical Methods

Total usability scores were calculated according to the study by Brooke [54] and presented through descriptive statistics (mean, SD, and range). We investigated whether usability was rated differently by the 3 cohorts using a one-way analysis of variance, as data are normally distributed and homogeneity of variance is given. An unpaired *t* test (two-tailed) was used to study gender differences, as assumptions of normal distribution and homogeneity of variance were met. The associations between usability and age (not normally distributed) as well as usability and compliance (not normally distributed) were investigated using Spearman rank correlations. The data were analyzed using R 3.6.1 with RStudio. The α level was set to .05.

Evaluation of Validity

Measures

The relative validity was assessed using a counterbalanced crossover design. Myfood24 Germany (Measure Your Food on One Day), a 24-hour recall, was chosen as the reference method. Myfood24 is a web-based, self-administered 24-hour dietary recall tool (refer to Koch et al [39] for details). It is based on two German nutritional databases (the German Food Code and Nutrient Data Base, Bundeslebensmittelschlüssel version 3.02, and the database LEBTAB of the Dortmund Nutritional and Anthropometric Longitudinally Designed study) and includes 11,501 food items. A comparison between habitual energy and macronutrient intake assessed through the APPetite-food record and 24-hour recall was drawn. Habitual intake was operationalized as the mean dietary intake of 3 days.

Furthermore, energy intake is compared with total energy expenditure (TEE) based on the assumption that energy intake equals TEE in weight-stable individuals [55]. TEE is estimated from nondominant wrist accelerometry according to White et al [56], which has been shown to be a precise approach to estimate TEE on population levels in free-living conditions when compared with TEE by doubly labeled water. The Euclidean norm minus one was extracted from the raw acceleration data using the DataAnalyzer from movisens (version 1.13.5; June 18, 2019) and inserted into the quadratic Euclidean norm minus one equation from White et al [56].

Sample

A total of 50 healthy participants from the LORA study (group 1: n=26; group 2: n=24) volunteered for the validation study. However, 6 participants from group 1 had to be excluded, as they did not complete all relevant parts within the predefined time schedule. Therefore, the evaluation of validity was based on data from 44 participants (33 women and 11 men)—20 from group 1 and 24 from group 2. This sample had a mean age of 28.64 (SD 8.13) years and a mean BMI of 23.8 (SD 3.62) kg/m^2. The groups did not significantly differ in terms of sex (group 1: 15 women and 5 men; group 2: 18 women and 6 men; X^2_1=0; *P*=.99), age (group 1: mean 30.15, SD 8.65 years; group 2: mean 27.38, SD 7.63 years; Mann-Whitney *U*=314; *P*=.08), and BMI (group 1: mean 23.99, SD 3.69 kg/m^2; group 2: mean 23.66, SD 3.63 kg/m^2; $t_{40.34}$=0.3 [unpaired; two-tailed]; *P*=.77).

Procedure

Participants from the LORA cohort who agreed to participate in the APPetite study were asked whether they wanted to also participate in the validation study: recording their food intake through a 24-hour recall on 3 additional days. Of the participants who agreed, 26 were assigned to group 1 and 24 to group 2, following a counterbalanced crossover design. Hence, participants in group 1 completed three 24-hour recalls exactly a week before the APPetite-food record was used. In group 2, participants completed three 24-hour recalls exactly a week after the APPetite-food record was used. The same weekdays, the week before or after, were assessed. Both groups received the same training to familiarize themselves with the 24-hour recall and the APPetite-food record. Participants received €30 (US $36.52) to participate in the validation study.

Statistical Methods

Habitual energy and macronutrient intake assessed through the APPetite-food record was compared with habitual dietary intake assessed through the 24-hour recall.

Habitual energy and macronutrient intake from the two methods were compared at the group level using two-tailed paired *t* tests (for normally distributed data including energy and carbohydrates) and Wilcoxon signed-rank tests (for skewed data including protein, fat, sugar, and fiber). Agreement between the two methods at the individual level was assessed using

Bland-Altman analysis of the mean differences [57]. For this, the difference between the two methods (y-axis) is plotted against the mean of the two methods (x-axis) for each participant. For reference, the mean difference between the two methods across all participants and the limits of agreement (LoA) estimated by the mean difference above and below 1.96 SD of the differences are shown in the plot. Thus, a systematic bias throughout the range of measurements can be identified. Acceptable LoA must be predefined. We predefined acceptable LoA for energy and macronutrient intake as 10% of the group mean across the two methods. Daily energy intake from the APPetite-food record (not normally distributed on days 2 and 3 of the EMA assessment) was compared with TEE (normally distributed) through a two-tailed paired *t* test for the first day and Wilcoxon signed-rank tests for the second and third days. Paired *t* tests (two-tailed) were calculated to compare mean TEE and habitual energy intake from the APPetite-food record and the 24-hour recall. The α level was set to .05. The analyses were performed using R 3.6.1 with RStudio.

Results

Feasibility

A total of 98.28% (3703/3768) of all scheduled prompts were delivered. The failure of prompt delivery was either due to technical problems or because the smartphone was switched off. Overall, 80.31% (3026/3768) of the prompts were answered completely. In total, 0.9% (34/3768) of prompts were registered as incomplete as a result of technical problems or extensive breaks during prompt completion. A total of 1.81% (68/3768) of prompts were dismissed, and 15.26% (575/3768) of prompts were ignored. The relatively large proportion of ignored prompts was, to some extent, a result of participants unintentionally leaving their smartphone at home or in another room. Furthermore, a number of participants reported that they had missed the first prompt or prompts of the day, as they were still sleeping.

Overall mean compliance (percentage of complete prompts within received prompts) was 81.73% (SD 21.65%). The compliance rate of 67.5% (106/157) of participants was above 80% (LORA: 94/136, 69.1%; PROUD: 4/7, 57%; and BipoLife-A1: 8/14, 57%). The mean compliance rate was 81.56% (SD 25.98%) on the first day, 83.28% (SD 23.55%) on the second day, and 79.97% (SD 25.8%) on the third day. The Friedman test showed no significant difference in compliance among the 3 days (X^2_2=3.6; *P*=.17), indicating no decline in motivation.

Compliance was highest in the LORA cohort and lowest in the PROUD cohort (see cohort means and SDs in Table 2). However, the Kruskal-Wallis rank sum test showed that compliance of the 3 cohorts did not differ significantly (X^2_2=0.7; *P*=.72).

Female participants had, on average, a compliance of 83.95% (SD 19.02%). The mean compliance for male participants was 77.83% (SD 25.34%). The Wilcoxon rank sum test found no gender difference in compliance (*P*=.16). No significant correlation was found between age and compliance (ρ=0.13; *P*=.12).

Participants responded to prompts after a mean of 189.32 seconds (SD 388.65). Responding to 70.54% (2157/3058) of all prompts was started within the first 60 seconds after the first prompt signal. The mean time needed to complete a single prompt was 122.63 seconds (SD 70.01). The prompt response latency and response duration for each of the 3 cohorts are shown in Table 3.

Table 2. Mean number and percentage of complete, incomplete, dismissed, and ignored prompts within received prompts for the total sample and each cohort.

Samples	Complete	Incomplete	Dismissed	Ignored
Total (N=157), mean (SD)				
Values	19.27 (5.32)	0.22 (0.44)	0.43 (1.07)	3.66 (4.79)
Percentage	81.73 (21.65)	0.93 (1.89)	1.82 (4.48)	15.52 (20.08)
LORA[a] (n=136), mean (SD)				
Values	19.43 (5.04)	0.23 (0.46)	0.4 (1.05)	3.5 (4.46)
Percentage	82.48 (20.41)	0.98 (1.95)	1.71 (4.4)	14.83 (18.7)
PROUD[b] (n=7), mean (SD)				
Values	17.14 (7.73)	0.43 (0.54)	0.86 (1.46)	5.57 (6.95)
Percentage	71.43 (32.22)	1.79 (2.23)	3.57 (6.09)	23.21 (28.95)
BipoLife-A1 (n=14), mean (SD)				
Values	18.86 (6.76)	0 (0)	0.5 (1.09)	4.29 (6.63)
Percentage	79.59 (27.46)	0 (0)	2.08 (4.55)	18.33 (27.8)

[a]LORA: Longitudinal Resilience Assessment.

[b]PROUD: Prevention of Comorbid Depression and Obesity in Attention-Deficit/Hyperactivity Disorder.

Table 3. Response latency and duration for a single prompt, reporting latency of the food record, and recording duration for the food record per day for the total sample and each cohort.

Variables	Total	LORA[a]	PROUD[b]	BipoLife-A1
Prompts (s), mean (SD)				
Latency	189.32 (388.65)	179.38 (375.82)	242.24 (447.85)	265.24 (469.8)
Duration	122.63 (70.01)	119.13 (64.21)	175.12 (144.31)	134.01 (64.07)
Food record (min), mean (SD)				
Latency	58.35 (127.52)	50.82 (115.8)	147.02 (197.85)	116.99 (190.16)
Duration	17.66 (8.66)	17.8 (8.57)	22.17 (10.86)	14.26 (7.56)

[a]LORA: Longitudinal Resilience Assessment.

[b]PROUD: Prevention of Comorbid Depression and Obesity in Attention-Deficit/Hyperactivity Disorder.

Dietary data of 8.8% (12/136) LORA, 14% (1/7) PROUD, and 14% (2/14) BipoLife-A1 participants had to be excluded, as the number of recorded meals or entered foods was evidently too low or entries were incomplete or implausible. In addition, 3 LORA participants had no food entry on 1 day. However, the remaining 2 days were recorded sufficiently well to be included. Included participants (n=142) recorded a total of 2969 eating and drinking events. In total, 3.03% (90/2969) entries were registered as incomplete, mainly due to technical problems. Participants entered on average 7.02 (SD 3.33) eating and drinking events per day (first day: mean 7.49, SD 3.14; range 2-17; no data available for 1 participant; second day: mean 7.08, SD 3.43; range 2-22; third day: mean 6.49, SD 3.37; range 2-17; no data available for 2 participants).

The mean latency from food intake to food recording was 58.35 (SD 127.52) minutes. Latency increased over the course of the 3 days (first day: mean 53.51, SD 72.01 min; second day: mean 69.5, SD 88.1 min; third day: 90.81, SD 116.12 min). The mean time to complete the food record of one day was 17.66 (SD 8.66) minutes. On the first day, participants took 21.01 (SD 9.68) minutes, on the second day they took 17.22 (SD 7.76) minutes, and on the third day they took 14.67 (SD 7.16) minutes. The cohort-specific food record latencies and recording durations are presented in Table 3.

The accelerometer records of 2 participants stopped during the second day. It is unknown if this was due to technical problems or because participants connected the sensor to a computer that instantly stopped the recording. In total, 11 participants did not wear the sensor on at least one day or stopped wearing it before the end of the 7-day assessment period. On average, participants (N=157 including the abovementioned) wore the sensor for 6 days 3 hours and 57 minutes (mean 8876.96, SD 1815.36 min; range 771-10,403). Hence, the mean compliance was 88.07% (SD 18.01%).

Usability

The SUS total score was 61.9 (SD 17.79; range 17.5-97.5) out of 100. The SUS score of the LORA cohort (n=67) was 61.23 (SD 16.8; range 17.5-95). The lowest usability with an SUS score of 60 (SD 24.08; range 32.5-92.5) was rated by the PROUD cohort (n=6). The highest SUS score was found for the Bipolife-A1 cohort (n=11), with a score of 67.05 (SD 20.97; range 22.5-97.5). However, the 3 cohorts did not differ in the ratings of usability according to a one-way analysis of variance ($F_{2,81}$=0.54; *P*=.59).

Female participants (mean 62.82, SD 17.36) scored usability on average marginally higher than male participants (mean 60.17, SD 18.77; t_{82}=0.65 [two-tailed]; *P*=.52). Age and usability were not significantly negatively correlated (ρ=−0.18; *P*=.10). Compliance and rated usability showed no significant correlation (ρ=0.13; *P*=.26).

Validity

Habitual intake of energy, protein, fat, carbohydrates, sugar, and fiber assessed through the APPetite-food record and the 24-hour recall are shown in Table 4. All nutritional intake was higher for the APPetite-food record. The difference between the two methods is significant for energy, protein, fat, and fiber intake (Table 4).

With regard to possible order effects, both groups (APPetite-food record first and 24-hour recall first) showed higher energy intake assessed through the APPetite-food record (group 1: 8494 kJ/2029 kcal; group 2: 9327 kJ/2228 kcal) compared with the 24-hour recall (group 1: 7881.23 kJ/1882 kcal; group 2: 8086.01 kJ/1931 kcal).

Agreement between the two methods at the individual level was investigated through Bland-Altman plots for energy and macronutrient intake. Mean energy difference between the APPetite-food record and the 24-hour recall was 994.18 kJ (95% CI 370.8-1617.6). A normal distribution of the difference was observed. Figure 3 shows the Bland-Altman plot of the habitual energy intake. The LoA are −3024.841 (95% CI −4104.6 to −1945.1) to 5013.2 (95% CI 3933.4-6093) and therefore larger than the predefined acceptable LoA of 849 kJ.

Bland-Altman analyses for protein, fat, carbohydrate, sugar, and fiber intake can be found in Multimedia Appendix 4 [57]. All LoA exceeded the predefined acceptable LoA.

Table 4. Mean habitual intake of energy and macronutrients from the APPetite-food record and the 24-hour recall; mean difference between the two methods; paired *t* tests (two-tailed) for normally distributed data including energy and carbohydrates; and Wilcoxon signed-rank tests for skewed data, including protein, fat, sugar, and fiber.

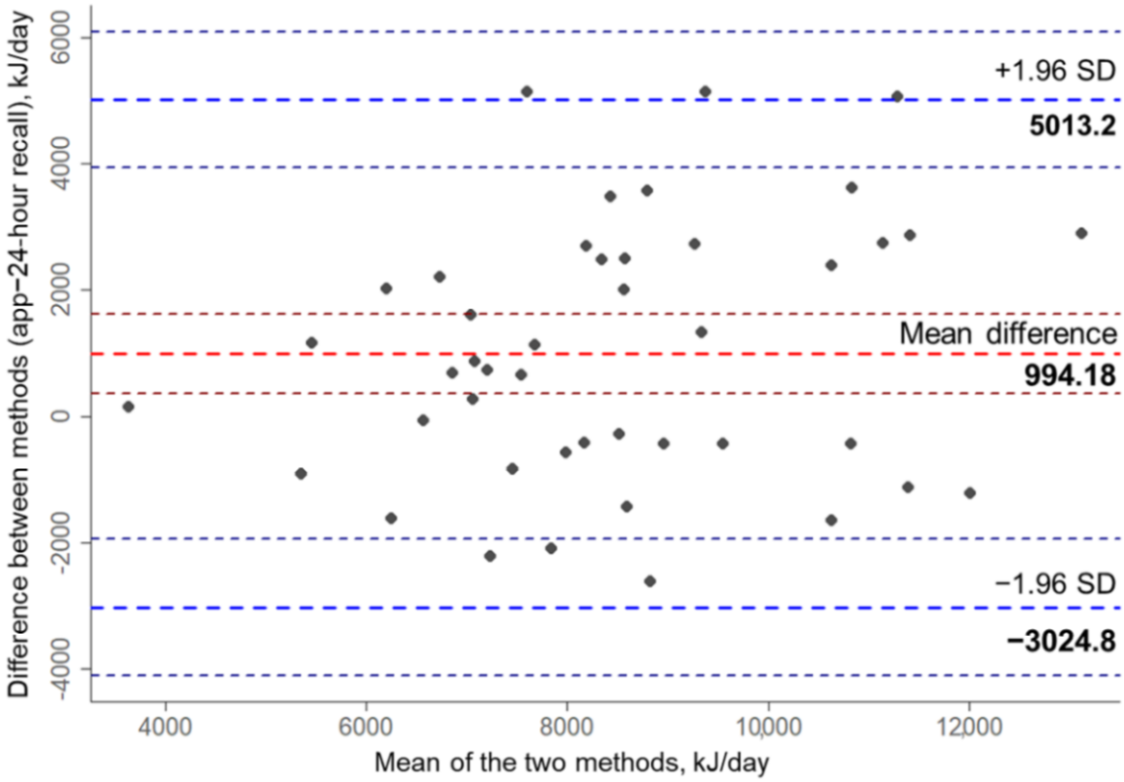
Dietary intake per day	APPetite-food record, mean (SD)	24-hour recall, mean (SD)	Mean difference (SD)	*t* test (*df*)	*P* value
Energy (kJ/day)	8987.11 (2404.65)	7992.93 (2002.61)	994.18 (2050.52)	3.21 (43)	.003[a]
Energy (kcal/day)	2146.42 (574.5)	1909.16 (478.8)	237.26 (489.94)	3.21 (43)	.003
Protein (g/day)	80.77 (27.6)	69.68 (24.72)	11.09 (22.63)	N/A[b]	.004[a]
Total fat (g/day)	92.25 (32.54)	76.95 (26.54)	15.3 (29.41)	N/A	.002[a]
Carbohydrate (g/day)	228.81 (63.12)	212.99 (52.65)	15.82 (64.03)	1.64 (43)	.11
Sugars (g/day)	81.21 (32.6)	76.97 (29.76)	4.24 (31.42)	N/A	.40
Fiber (g/day)	25.86 (9.22)	23.3 (8.5)	2.56 (7.43)	N/A	.04[c]

[a]*P*<.01.

[b]N/A: not applicable to Wilcoxon signed-rank test.

[c]*P*<.05.

Figure 3. Bland-Altman plot assessing agreement between habitual energy intake in kJ per day captured by the APPetite-food record and the 24-hour recall (red line: mean difference=app–24-hour recall; dark red lines: 95% CI of mean difference; blue lines: lower and upper limits of agreement; dark blue lines: 95% CI of lower and upper limits of agreement).



Energy intake from the APPetite-food record was significantly lower than the TEE estimated from accelerometry on the first day (t_{43}=5.33; *P*<.001; TEE mean 2425.4, SD 468.2; app mean 1897.53, SD 616.32), but did not significantly differ on days 2 and 3: day 2 (*P*=.051; TEE mean 2442.04, SD 447.5; app mean 2242.94, SD 769.78) and day 3 (*P*=.23; TEE mean 2435.6, SD 482.9; app mean 2317.77, SD 780.6). Mean TEE estimated from 7 days of accelerometry was 2417.8 kcal (SD 410) compared with the habitual energy intake of 2146.42 kcal (SD 574.5) from the APPetite-food record and 1909.16 kcal (SD 478.8) from the 24-hour recall. Paired *t* tests (two-tailed) showed that habitual energy intake was underestimated by both methods when compared with TEE: APPetite-food record (t_{43}=3.40; *P*=.002) and 24-hour recall (t_{43}=6.33; *P*<.001).

Discussion

Principal Findings

The APPetite-mobile-app was developed to capture complex dietary intake, objectively recorded PA, and related factors for studying microtemporal, within-person processes underlying eating behavior and PA in daily life. This study evaluated the feasibility and usability of the EMA tool as well as the validity of the APPetite-food record. The APPetite-mobile-app demonstrated good feasibility. Compliance with responding to prompts and wearing the accelerometer was above 80%, and reasonable response times and latencies were found for the prompts as well as the food record. Usability was rated moderate, with a mean SUS score of 61.9. Large differences in energy and macronutrient intake assessed with the APPetite-food record versus the 24-hour recall were found at the group and individual levels, whereby the APPetite-food record captured higher dietary intakes. Energy intake was assessed fairly accurately by the APPetite-food record on the group level on 2 of 3 days when compared with TEE. The comparison of habitual energy intake to mean TEE showed that the 24-hour recall underestimated energy intake to a larger degree than the APPetite-food record. These results indicate that the discrepancies between the two dietary assessment methods do not imply a lack of validity of the APPetite-food record; rather, they indicate a more accurate dietary assessment compared with the 24-hour recall and therefore provide the first evidence that the APPetite-food record is a valid tool for capturing complex dietary intake.

Comparison With Previous Work

The good feasibility of EMA tools is crucial to ensure unbiased data collection and prevent systematic missing data. Compliance rates are an important indicator of feasibility. Although there is no official criterion indicating good compliance, Stone and Shiffman [58] proposed compliance rates above 80% to be acceptable. However, they emphasized the arbitrariness of this criterion and the need to define acceptable compliance ranges for each study individually, especially when noncompliance may be systematic and not random. The mean prompt compliance in our study was above 80% and can therefore generally be rated as good. These good compliance rates may be partly due to the notifications participants received when falling below the 80% threshold and the incentive to be included in raffles when reaching a compliance rate of 80% or above.

Furthermore, the results demonstrate that prompt compliance did not decrease over the course of 3 days. In other EMA studies that assessed more than 3 days, response rates declined substantially (eg, 40% from 63% on day 1 to 23% on day 7), even with only 4 prompts per day [36]. As studying microtemporal processes requires an illustration of a day in high resolution, it is more important to focus on a larger number of completed prompts per day compared with a large number of EMA assessment days. On the basis of our constant compliance rate over 3 days, the length of the EMA assessment seems feasible, and no decline in motivation was evident.

We found marginally lower prompt compliance rates in the clinical cohorts than in the healthy cohort. In a study by Porras-Segovia et al [59] comparing EMA compliance rates from suicidal patients and student controls, lower compliance was found for the clinical sample. These findings were consistent with our results. However, Porras-Segovia et al [59] found a significant difference between patients and healthy controls, which was not the case in our study. These results suggest that the prompt schedule of the APPetite-mobile-app is equally well suited for healthy individuals and patients with a mental disorder.

A mean of 2.04 (SD 1.17) minutes was needed to complete 1 prompt. In accordance with the high prompt compliance rate, the response duration of the prompts can be considered feasible. Responding to a prompt was initiated on average 3.16 (SD 6.48) minutes after the first prompt signal. Short prompt latencies are essential to guarantee the momentary nature of the response and should therefore be taken into account thoroughly. However, most studies have not reported prompt latencies [60,61]. Some studies have predefined response windows. This ensures the momentary nature of the response but can cause lower compliance rates, for example, 69% in a study with an 8-minute response window [62]. We chose to allow a longer response period and prompt postponement of up to 25 minutes to reduce participants' burden and maintain high compliance. Nevertheless, participants were instructed to respond to EMA prompts instantly, if possible. Considering that we allowed responses up to 30 minutes after the first prompt signal, the mean latency of just over 3 minutes is short and underlines the feasibility of the prompts.

Compliance with the food record cannot be directly determined, as it is not possible to differentiate between someone not recording a food item because of noncompliance or because of not actually consuming it. However, other quality measures could also be used. The time spent reporting daily dietary intake or the number of recorded eating and drinking occasions per day can be used for quality checks. On average, participants needed 17.6 minutes to complete the food record of 1 day. Other technology-based tools for assessing dietary intake show similar times to complete, ranging between 13 and 45 minutes [20]. Participants entered on average 7 eating and drinking events per day. This number is in line with a previous study that found a mean of 20.7 eating and drinking occasions per individual over a 3-day period [29].

The APPetite-food record was developed to record food intake in real time or near real time. Therefore, it is important to consider the amount of time between food intake and recording. Foods and drinks were recorded on average 58.35 minutes after intake. This shows that participants did not wait until the evening to record all eating and drinking occasions for 1 day. Hence, food intake was recorded in real time or near real time.

The food recording behavior of our participants suggests that the APPetite-food record is feasible. However, we noticed that the participants' motivation was crucial for successfully capturing sufficient and accurate dietary data. Training is needed to ensure that participants understand the importance of food recording in real time or near real time. Furthermore, participants reported that receiving detailed dietary feedback at the end of the study increased their motivation to enter food intake accurately.

As expected from previous studies [63], a high compliance rate (88.07%) for the accelerometer worn on the wrist were found. All measures of feasibility regarding prompts, the APPetite-food record, and the accelerometer indicate that the APPetite-mobile-app is a feasible EMA tool.

In addition to good feasibility, usability is an important criterion that should be considered when developing new EMA tools. The usability of the APPetite-mobile-app was rated as moderate, with an SUS score of 61.9 out of 100. In a previous study, the usability of the top 7 iPhone operating system and Android diet-tracking apps was assessed [22]. The usability of 2 apps was rated even lower than the APPetite-mobile-app (Lose It!=59.2; MyDietCoach=46.7). However, a comparison is difficult as these tools focus purely on dietary assessment. The SUS score of the APPetite-mobile-app was rated on the basis of both dietary assessment and EMA prompts. The relatively low usability of our tool can be explained by the fact that it is a scientific device and was therefore developed independently without professional app developers. High costs are involved in the professional development of an app. For this reason, we chose the platform movisensXS to independently develop the app. Although movisensXS has many configuration options, it still has its limitations. For example, a search function within the food record cannot be implemented. The app was developed for scientific purposes only and not for consumer use. However, usability challenges have been reported even for commercial tools; for instance, only 20% of participants would continue to use MyFitnessPal after study participation [64]. Even though the usability of the APPetite-mobile-app was rated relatively low, no negative effect on feasibility, including compliance, was evident. Therefore, an improvement in usability is desirable but not essential for its use in scientific research.

A food record was incorporated into the APPetite-mobile-app to capture complex dietary intake in real time or near real time. An evaluation of validity was needed to test whether the APPetite-food record accurately assessed dietary intake. Hence, the APPetite-food record was compared with a 24-hour recall and TEE estimated from nondominant wrist accelerometry.

With regard to relative validity, low agreement between habitual dietary intake measured by the APPetite-food record and the 24-hour recall was found at both the group and individual levels. At the group level, energy, protein, fat, and fiber intake from the APPetite-food record was significantly higher than the 24-hour recall. Wide LoA, which exceeded the predefined acceptable LoA, were found for energy, protein, fat, carbohydrate, sugar, and fiber intake at the individual level. One could argue that these discrepancies indicate a lack of validity in the APPetite-food record. However, even though 24-hour recalls are frequently used as the established reference method when assessing relative validity, the true intake remains unknown [65]. Therefore, possible reasons for this discrepancy must be taken into account for both methods. Most validation studies that compared an EMA dietary assessment tool with a 24-hour recall found lower values for energy intake as well as intake of some macronutrients assessed through the EMA tool on the descriptive or even statistical level (eg, [64] for energy [statistical], proteins [statistical], fat [statistical], carbohydrates [statistical], and sugar [statistical]; [65] for energy, protein, fat, and carbohydrates; [66] for energy and fat [statistical], not for proteins and carbohydrates; [67] for energy, fat, and carbohydrates, not for proteins; [68] for energy, protein, sugar, and fat, not for carbohydrates and fiber, no statistical hypothesis test reported; [69] for energy, fat [statistical], carbohydrates, and fiber, not for proteins). This was not the case in our study. Habitual energy, protein, fat, carbohydrate, sugar, and fiber intake was higher when assessed with the APPetite-food record on the descriptive or even statistical level, indicating a lower degree of underreporting. This leads to the conclusion that the APPetite-food record could be a more precise dietary assessment method than the 24-hour recall.

This interpretation is underlined by the comparison of energy intake and TEE estimated from accelerometry. Energy intake from the APPetite-food record was not significantly different from TEE on 2 of 3 days, indicating that the APPetite-food record assesses energy intake fairly accurately at the group level. However, the comparison with the mean TEE showed that both methods underestimated habitual energy intake. In this context, it must be mentioned that over one-third of the participants in the validation study (17/44, 39%) indicated that they are currently trying to lose weight. Therefore, the discrepancy between the TEE and the reported energy intake could, to some extent, be due to diet and weight loss. However, the 24-hour recall underestimated habitual energy intake to a greater extent. Inaccurate estimates of energy intake captured by 24-hour recalls have been reported in previous studies [70]. A reason for the improved reporting accuracy in dietary assessments in real time or near real time compared with retrospective assessments could be the minimized retention interval [71,72]. Memory effects can cause bias in retrospective dietary assessments, as the demand for memory increases simultaneously with the retention interval. Memory lapses can cause two types of errors in the context of 24-hour dietary recalls: the failure to recall foods actually consumed (errors of omission) and the reporting of foods that were actually not consumed during the recalled day (errors of intrusion) [73]. Furthermore, incorrect estimations of portion sizes have been reported to constitute the largest measurement error in 24-hour recalls [73]. This error is closely related to memory bias, as consumed amounts must not only be accurately estimated but also be correctly remembered [74]. Food records in real time or near real time can minimize memory errors [74]. In a recent study, 65% of participants reported that remembering meal items and portion sizes was easier in a progressive assessment than in a traditional retrospective 24-hour recall [75]. Nonetheless, food records in real time or near real time are also affected by potential bias, which was also shown in our study as underestimation of food intake became evident. In particular, the change in dietary intake as a result of recording it has to be taken into account. Participants may choose not to eat complex meals or eat less to avoid extensive and time-consuming records [74]. Furthermore, keeping a record of food intake in daily life can be burdensome. Participants may not be able to record everything eaten due to other commitments. However, our results suggest that the impact of reactivity and high burden on the APPetite-food record might be smaller than the effect of memory loss on 24-hour recalls.

The results of the evaluation of validity indicate that the APPetite-food record might assess dietary intake more accurately than the 24-hour recall and capture daily energy intake fairly accurately at the group level. Nevertheless, both dietary assessment methods seem to underestimate habitual energy intake.

Strengths and Limitations

A common validation approach is the assessment of relative validity, which compares a novel tool to an established dietary assessment method. However, most of the available validation studies show methodological issues as they assess relative validity on overlapping days [64-68] and do not use a counterbalanced crossover design [69]. Assessing overlapping days can lead to an overestimation of agreement between two self-report methods, as recording dietary intake actively throughout the day may improve memory for completing the 24-hour dietary recall of the same day. Empirical evidence for the overestimated agreement has been found, detecting improved accuracy of 24-hour recalls of days when diet was tracked throughout [65]. A further problem becomes apparent in studies that do not assess overlapping days. When two methods are used one after another, order effects can bias the assessment. However, most studies did not control for possible order effects [69]. We were able to counteract these methodological issues by choosing a counterbalanced crossover study design that assessed no overlapping days. A counterbalanced crossover design is crucial for controlling learning, boredom, and other unwanted order effects. We understand this to be the most significant strength of our validation study.

One limitation of our validation study is due to the fact that dietary intake varies from day to day. Bland and Altman call this case "method where true value varies" [76]. When the true value varies, measurements of two methods have to be taken at the same time point to obtain an accurate estimate of agreement [77]. In the context of dietary assessment methods that would translate to assessing food intake using two methods on the same day. However, because an inflated agreement when assessing overlapping days is likely to occur [26], as mentioned earlier, this does not represent a suitable approach. Therefore, we were not able to compare dietary intake on a day level (eg, Thursday compared with Thursday the week before or after) and chose to compare habitual dietary intake instead. However, when comparing habitual dietary intake, two aspects must be considered: (1) the target of interest of the APPetite-mobile-app is not regular or habitual food intake but rather microtemporal dynamics of food intake in daily life. Using habitual intake as the measure of comparison sets aside this fact and might therefore not be the most appropriate measure for the evaluation of validity. (2) Day-to-day variability in dietary intake represents a problem when assessing habitual intake. It could be argued that capturing 3 days to operationalize habitual intake is not sufficient to obtain an accurate estimate.

Many studies that use Bland-Altman agreement analyses to evaluate the validity of food records in real time or near real time have inaccuracies. To the best of our knowledge, our study is the only one that has a predefined acceptable LoA. These pre-established limits are necessary to avoid misleading interpretations. A consensus on the acceptable LoA for dietary intake is desirable. This will improve the comparability of the results from studies assessing relative validity. Furthermore, the use of established but biased dietary assessment methods, such as 24-hour recalls, to study relative validity should be questioned critically. New approaches to evaluate the validity of food records in real time or near real time are needed.

Our findings are limited because of the lack of control for possible weight changes during study participation. The comparison of TEE and energy intake is based on the assumption that energy expenditure is equal to energy intake. However, this assumption is valid only for weight-stable individuals.

Two further limitations concern the APPetite-food record itself: (1) nutritional values are generated manually, which is time-consuming and can be error-prone. Automated generation is preferable. (2) The APPetite-food record relies on self-reports of dietary intake. Self-reports are subjective and therefore more likely to be biased. To add a more objective component to the dietary assessment, photos of the foods and drinks consumed could be taken in addition to self-reports.

The strength of our assessment of feasibility and usability is that the sample of healthy participants was enriched with data from patients suffering from a mental disorder. Therefore, it was possible to show that the APPetite-mobile-app is equally feasible and usable in this population. This finding is particularly important as diet and PA play an important role in mental health. This opens up the possibility of studying microtemporal, within-person processes of diet, PA, and related factors in psychiatric patients, which is crucial for the understanding of the link among diet, PA, and mental health. However, the unequal sample sizes of the 3 cohorts limit the results. This is of concern in the context of cohort comparisons, as well as the interpretation of the means of the total sample. Furthermore, a selection bias could be present, as the participants were exclusively recruited from 3 existing study cohorts.

Recommendations for Future Studies

The development of novel EMA tools for assessing microtemporal processes of diet, PA, and related factors is required. Studies comparing these new EMA tools are needed to establish empirical evidence on which assessment approaches are most effective in the study of microtemporal processes. Future EMA studies should consider that participants' motivation is the key to success, especially when complex dietary intake is assessed. Therefore, participants' burden needs to be kept minimal, and incentives for prompt responding and food recording, such as dietary feedback and raffle inclusions, are essential.

New technologies and wearable sensors are a promising advancement in the area of dietary assessment in naturalistic settings, as they can passively detect eating behavior [78]. They can be used for longer assessment periods because they require minimal user interaction. These sensors will improve the validity of self-reported dietary assessments to a great extent. We believe they will soon be of tremendous relevance, especially for the assessment of microtemporal processes of diet in daily life.

Conclusions

The APPetite-mobile-app is a promising tool for studying microtemporal, within-person processes of diet, PA, and related factors in real time or near real time and is, to the best of our knowledge, the first of its kind. First evidence supports that the APPetite-mobile-app is feasible and the APPetite-food record is a valid tool for capturing complex dietary intake. We hope this motivates other researchers to use EMA to capture complex dietary intake, PA, and associated factors in daily life, and it initiates a discussion about feasible, usable, and valid methods to assess these dynamics. Assessment strategies need to be developed, shared, and discussed to advance the research field. A solid empirical foundation regarding within-person, microtemporal associations of diet, PA, and associated factors is needed for the development of personalized lifestyle modification interventions, such as intensively adaptive interventions or just-in-time adaptive interventions [10].

Acknowledgments

This work was supported by the European Union's Horizon 2020 Research and Innovation Program under grant agreement number 728018. The funding source had no involvement in the study design, data collection, interpretation of the findings, or writing of this manuscript.

Conflicts of Interest

UEP reports consultancy for Boehringer-Ingelheim. All other authors declare no conflicts of interest.

Multimedia Appendix 1

APPetite-food record and nutritional value generation.
[DOCX File , 111 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Reasons for food intake.
[DOCX File , 33 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Prompt measures and items.
[DOCX File , 39 KB-Multimedia Appendix 3]

Multimedia Appendix 4

Bland-Altman analysis of protein, fat, carbohydrate, sugar, and fiber intake.
[DOCX File , 415 KB-Multimedia Appendix 4]

References

1. World Health Organization. Obesity and Overweight.: World Health Organization; 2020. URL: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight [accessed 2020-09-25]
2. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2019 May 11;393(10184):1958-1972 [FREE Full text] [doi: 10.1016/S0140-6736(19)30041-8] [Medline: 30954305]
3. Adan RA, van der Beek EM, Buitelaar JK, Cryan JF, Hebebrand J, Higgs S, et al. Nutritional psychiatry: towards improving mental health by what you eat. Eur Neuropsychopharmacol 2019 Dec;29(12):1321-1332 [FREE Full text] [doi: 10.1016/j.euroneuro.2019.10.011] [Medline: 31735529]
4. Kelly SJ, Daniel M, Grande ED, Taylor A. Mental ill-health across the continuum of body mass index. BMC Public Health 2011 Oct 05;11:765 [FREE Full text] [doi: 10.1186/1471-2458-11-765] [Medline: 21975214]
5. Rajan TM, Menon V. Psychiatric disorders and obesity: a review of association studies. J Postgrad Med 2017;63(3):182-190 [FREE Full text] [doi: 10.4103/jpgm.JPGM_712_16] [Medline: 28695871]
6. Milaneschi Y, Simmons WK, van Rossum EF, Penninx BW. Depression and obesity: evidence of shared biological mechanisms. Mol Psychiatry 2019 Jan;24(1):18-33. [doi: 10.1038/s41380-018-0017-5] [Medline: 29453413]
7. Sleddens EF, Kroeze W, Kohl LF, Bolten LM, Velema E, Kaspers P, et al. Correlates of dietary behavior in adults: an umbrella review. Nutr Rev 2015 Aug;73(8):477-499 [FREE Full text] [doi: 10.1093/nutrit/nuv007] [Medline: 26106126]
8. Sleddens EF, Kroeze W, Kohl LF, Bolten LM, Velema E, Kaspers PJ, et al. Determinants of dietary behavior among youth: an umbrella review. Int J Behav Nutr Phys Act 2015 Feb 01;12:7 [FREE Full text] [doi: 10.1186/s12966-015-0164-x] [Medline: 25638322]
9. Wood W, Neal DT. Healthy through habit: Interventions for initiating & maintaining health behavior change. Behav Sci Policy 2016;2(1):71-83 [FREE Full text] [doi: 10.1353/bsp.2016.0008]
10. Dunton GF. Sustaining health-protective behaviors such as physical activity and healthy eating. J Am Med Assoc 2018 Aug 21;320(7):639-640. [doi: 10.1001/jama.2018.6621] [Medline: 29852046]

11. Hummel E, Hoffmann I. Complexity of nutritional behavior: capturing and depicting its interrelated factors in a cause-effect model. Ecol Food Nutr 2016;55(3):241-257. [doi: 10.1080/03670244.2015.1129325] [Medline: 26828451]
12. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. Can Med Assoc J 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
13. Rhodes RE, Janssen I, Bredin SS, Warburton DE, Bauman A. Physical activity: health impact, prevalence, correlates and interventions. Psychol Health 2017 Aug;32(8):942-975. [doi: 10.1080/08870446.2017.1325486] [Medline: 28554222]
14. Bleich SN, Jones-Smith J, Wolfson JA, Zhu X, Story M. The complex relationship between diet and health. Health Aff (Millwood) 2015 Nov;34(11):1813-1820. [doi: 10.1377/hlthaff.2015.0606] [Medline: 26526238]
15. Moore MS. Interactions between physical activity and diet in the regulation of body weight. Proc Nutr Soc 2000 May;59(2):193-198. [doi: 10.1017/s0029665100000215] [Medline: 10946787]
16. Smyth JM, Stone AA. Ecological momentary assessment research in behavioral medicine. J Happiness Stud 2003;4(1):35-52. [doi: 10.1023/A:1023657221954]
17. Ebner-Priemer UW, Trull TJ. Ecological momentary assessment of mood disorders and mood dysregulation. Psychol Assess 2009 Dec;21(4):463-475. [doi: 10.1037/a0017075] [Medline: 19947781]
18. Maugeri A, Barchitta M. A systematic review of ecological momentary assessment of diet: implications and perspectives for nutritional epidemiology. Nutrients 2019 Nov 07;11(11):2696 [FREE Full text] [doi: 10.3390/nu11112696] [Medline: 31703374]
19. Mason TB, Do B, Wang S, Dunton GF. Ecological momentary assessment of eating and dietary intake behaviors in children and adolescents: a systematic review of the literature. Appetite 2020 Jan 01;144:104465. [doi: 10.1016/j.appet.2019.104465] [Medline: 31541670]
20. Eldridge AL, Piernas C, Illner A, Gibney MJ, Gurinović MA, de Vries JH, et al. Evaluation of new technology-based tools for dietary intake assessment-an ILSI Europe dietary intake and exposure task force evaluation. Nutrients 2018 Dec 28;11(1):55 [FREE Full text] [doi: 10.3390/nu11010055] [Medline: 30597864]
21. Pendergast FJ, Leech RM, McNaughton SA. Novel online or mobile methods to assess eating patterns. Curr Nutr Rep 2017;6(3):212-227 [FREE Full text] [doi: 10.1007/s13668-017-0211-0] [Medline: 28944099]
22. Ferrara G, Kim J, Lin S, Hua J, Seto E. A focused review of smartphone diet-tracking apps: usability, functionality, coherence with behavior change theory, and comparative validity of nutrient intake and energy estimates. JMIR Mhealth Uhealth 2019 May 17;7(5):e9232 [FREE Full text] [doi: 10.2196/mhealth.9232] [Medline: 31102369]
23. Jeffers AJ, Mason TB, Benotsch EG. Psychological eating factors, affect, and ecological momentary assessed diet quality. Eat Weight Disord 2020 Oct;25(5):1151-1159. [doi: 10.1007/s40519-019-00743-3] [Medline: 31388844]
24. Roy PG, Jones KK, Martyn-Nemeth P, Zenk SN. Contextual correlates of energy-dense snack food and sweetened beverage intake across the day in African American women: an application of ecological momentary assessment. Appetite 2019 Jan 01;132:73-81. [doi: 10.1016/j.appet.2018.09.018] [Medline: 30261234]
25. Powell DJ, McMinn D, Allan JL. Does real time variability in inhibitory control drive snacking behavior? An intensive longitudinal study. Health Psychol 2017 Dec;36(4):356-364. [doi: 10.1037/hea0000471] [Medline: 28192005]
26. Richard A, Meule A, Reichenberger J, Blechert J. Food cravings in everyday life: an EMA study on snack-related thoughts, cravings, and consumption. Appetite 2017 Jun 01;113:215-223. [doi: 10.1016/j.appet.2017.02.037] [Medline: 28249745]
27. Zenk SN, Horoi I, McDonald A, Corte C, Riley B, Odoms-Young AM. Ecological momentary assessment of environmental and personal factors and snack food intake in African American women. Appetite 2014 Dec;83:333-341. [doi: 10.1016/j.appet.2014.09.008] [Medline: 25239402]
28. Allan J, McMinn D, Powell D. Tracking snacking in real time: time to look at individualised patterns of behaviour. Nutr Health 2019 Sep;25(3):179-184. [doi: 10.1177/0260106019866099] [Medline: 31347450]
29. Franja S, Wahl DR, Elliston KG, Ferguson SG. Comfort eating: an observational study of affect in the hours immediately before, and after, snacking. Br J Health Psychol 2021 Jan 04:12505. [doi: 10.1111/bjhp.12505] [Medline: 33398921]
30. Ashurst J, van Woerden I, Dunton G, Todd M, Ohri-Vachaspati P, Swan P, et al. The association among emotions and food choices in first-year college students using mobile-ecological momentary assessments. BMC Public Health 2018 May 02;18(1):573 [FREE Full text] [doi: 10.1186/s12889-018-5447-0] [Medline: 29716572]
31. Grenard JL, Stacy AW, Shiffman S, Baraldi AN, MacKinnon DP, Lockhart G, et al. Sweetened drink and snacking cues in adolescents: a study using ecological momentary assessment. Appetite 2013 Aug;67:61-73 [FREE Full text] [doi: 10.1016/j.appet.2013.03.016] [Medline: 23583312]
32. Bruening M, Ohri-Vachaspati P, Brewis A, Laska M, Todd M, Hruschka D, et al. Longitudinal social networks impacts on weight and weight-related behaviors assessed using mobile-based ecological momentary assessments: Study Protocols for the SPARC study. BMC Public Health 2016 Dec 30;16:901 [FREE Full text] [doi: 10.1186/s12889-016-3536-5] [Medline: 27576358]
33. Wahl DR, Villinger K, König LM, Ziesemer K, Schupp HT, Renner B. Healthy food choices are happy food choices: Evidence from a real life sample using smartphone based assessments. Sci Rep 2017 Dec 06;7(1):17069. [doi: 10.1038/s41598-017-17262-9] [Medline: 29213109]

34. Seto E, Hua J, Wu L, Shia V, Eom S, Wang M, et al. Models of individual dietary behavior based on smartphone data: the influence of routine, physical activity, emotion, and food environment. PLoS One 2016;11(4):e0153085 [FREE Full text] [doi: 10.1371/journal.pone.0153085] [Medline: 27049852]
35. Schüz B, Bower J, Ferguson SG. Stimulus control and affect in dietary behaviours. An intensive longitudinal study. Appetite 2015 Apr;87:310-317. [doi: 10.1016/j.appet.2015.01.002] [Medline: 25579222]
36. Spook JE, Paulussen T, Kok G, Van Empelen P. Monitoring dietary intake and physical activity electronically: feasibility, usability, and ecological validity of a mobile-based Ecological Momentary Assessment tool. J Med Internet Res 2013;15(9):e214 [FREE Full text] [doi: 10.2196/jmir.2617] [Medline: 24067298]
37. Prince SA, Adamo KB, Hamel ME, Hardt J, Connor GS, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. Int J Behav Nutr Phys Act 2008;5:56 [FREE Full text] [doi: 10.1186/1479-5868-5-56] [Medline: 18990237]
38. Adamo KB, Prince SA, Tricco AC, Connor-Gorber S, Tremblay M. A comparison of indirect versus direct measures for assessing physical activity in the pediatric population: a systematic review. Int J Pediatr Obes 2009;4(1):2-27. [doi: 10.1080/17477160802315010] [Medline: 18720173]
39. Koch SA, Conrad J, Hierath L, Hancock N, Beer S, Cade JE, et al. Adaptation and evaluation of Myfood24-Germany: a web-based self-administered 24-h dietary recall for the German adult population. Nutrients 2020 Jan 06;12(1):160 [FREE Full text] [doi: 10.3390/nu12010160] [Medline: 31935885]
40. Cleobury L, Tapper K. Reasons for eating 'unhealthy' snacks in overweight and obese males and females. J Hum Nutr Diet 2014 Aug;27(4):333-341. [doi: 10.1111/jhn.12169] [Medline: 24134077]
41. Verhoeven AA, Adriaanse MA, de Vet E, Fennis BM, de Ridder DT. It's my party and I eat if I want to. Reasons for unhealthy snacking. Appetite 2015 Jan;84:20-27. [doi: 10.1016/j.appet.2014.09.013] [Medline: 25261101]
42. Schüz B, Papadakis T, Ferguson SG. Situation-specific social norms as mediators of social influence on snacking. Health Psychol 2018 Feb;37(2):153-159. [doi: 10.1037/hea0000568] [Medline: 29154607]
43. Renner B, Sproesser G, Strohbach S, Schupp HT. Why we eat what we eat. The Eating Motivation Survey (TEMS). Appetite 2012 Aug;59(1):117-128. [doi: 10.1016/j.appet.2012.04.004] [Medline: 22521515]
44. Evers C, Adriaanse M, de Ridder DT, de Witt Huberts JC. Good mood food. Positive emotion as a neglected trigger for food intake. Appetite 2013 Sep;68:1-7. [doi: 10.1016/j.appet.2013.04.007] [Medline: 23602962]
45. Dunton GF, Liao Y, Intille S, Huh J, Leventhal A. Momentary assessment of contextual influences on affective response during physical activity. Health Psychol 2015 Dec;34(12):1145-1153. [doi: 10.1037/hea0000223] [Medline: 26053885]
46. Wilhelm P, Schoebi D. Assessing mood in daily life: structural validity, sensitivity to change, and reliability of a short-scale to measure three basic dimensions of mood. Eur J Psychol Assess 2007 Jan;23(4):258-267. [doi: 10.1027/1015-5759.23.4.258]
47. Reichenberger J, Kuppens P, Liedlgruber M, Wilhelm FH, Tiefengrabner M, Ginzinger S, et al. No haste, more taste: an EMA study of the effects of stress, negative and positive emotions on eating behavior. Biol Psychol 2018 Jan;131:54-62. [doi: 10.1016/j.biopsycho.2016.09.002] [Medline: 27654506]
48. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
49. Tomko RL, Solhan MB, Carpenter RW, Brown WC, Jahng S, Wood PK, et al. Measuring impulsivity in daily life: the momentary impulsivity scale. Psychol Assess 2014 Jun;26(2):339-349 [FREE Full text] [doi: 10.1037/a0035083] [Medline: 24274047]
50. Chmitorz A, Neumann RJ, Kollmann B, Ahrens KF, Öhlschläger S, Goldbach N, et al. Longitudinal determination of resilience in humans to identify mechanisms of resilience to modern-life stressors: the longitudinal resilience assessment (LORA) study. Eur Arch Psychiatry Clin Neurosci 2020 Jul 18:A. [doi: 10.1007/s00406-020-01159-2] [Medline: 32683526]
51. Mayer JS, Hees K, Medda J, Grimm O, Asherson P, Bellina M, et al. 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 2018 Feb 26;19(1):140. [doi: 10.1186/s13063-017-2426-1] [Medline: 29482662]
52. Pfennig A, Leopold K, Martini J, Boehme A, Lambert M, Stamm T, et al. Improving early recognition and intervention in people at increased risk for the development of bipolar disorder: study protocol of a prospective-longitudinal, naturalistic cohort study (Early-BipoLife). Int J Bipolar Disord 2020 Jul 01;8(1):22 [FREE Full text] [doi: 10.1186/s40345-020-00183-4] [Medline: 32607662]
53. Ritter PS, Bermpohl F, Gruber O, Hautzinger M, Jansen A, Juckel G, et al. Aims and structure of the German Research Consortium BipoLife for the study of bipolar disorder. Int J Bipolar Disord 2016 Dec;4(1):26 [FREE Full text] [doi: 10.1186/s40345-016-0066-0] [Medline: 27873290]
54. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland AL, Weerdmeester BA, editors. Usability Evaluation in Industry. London: Taylor and Francis; 1996:1-6.
55. Trabulsi J, Schoeller DA. Evaluation of dietary assessment instruments against doubly labeled water, a biomarker of habitual energy intake. Am J Physiol Endocrinol Metab 2001 Nov;281(5):891-899. [doi: 10.1152/ajpendo.2001.281.5.E891] [Medline: 11595643]

56. White T, Westgate K, Hollidge S, Venables M, Olivier P, Wareham N, et al. Estimating energy expenditure from wrist and thigh accelerometry in free-living adults: a doubly labelled water study. Int J Obes (Lond) 2019 Nov;43(11):2333-2342 [FREE Full text] [doi: 10.1038/s41366-019-0352-x] [Medline: 30940917]
57. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 8;1(8476):307-310. [Medline: 2868172]
58. Stone AA, Shiffman S. Capturing momentary, self-report data: a proposal for reporting guidelines. Ann Behav Med 2002;24(3):236-243. [doi: 10.1207/S15324796ABM2403_09] [Medline: 12173681]
59. Porras-Segovia A, Molina-Madueño RM, Berrouiguet S, López-Castroman J, Barrigón ML, Pérez-Rodríguez MS, et al. Smartphone-based ecological momentary assessment (EMA) in psychiatric patients and student controls: a real-world feasibility study. J Affect Disord 2020 Sep 01;274:733-741. [doi: 10.1016/j.jad.2020.05.067] [Medline: 32664009]
60. Liao Y, Skelton K, Dunton G, Bruening M. 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). J Med Internet Res 2016 Jun 21;18(6):e151 [FREE Full text] [doi: 10.2196/jmir.4954] [Medline: 27328833]
61. Heron KE, Everhart RS, McHale SM, Smyth JM. Using mobile-technology-based Ecological Momentary Assessment (EMA) methods with youth: a systematic review and recommendations. J Pediatr Psychol 2017 Nov 01;42(10):1087-1107. [doi: 10.1093/jpepsy/jsx078] [Medline: 28475765]
62. Rusby JC, Westling E, Crowley R, Light JM. Psychosocial correlates of physical and sedentary activities of early adolescent youth. Health Educ Behav 2014 Feb;41(1):42-51 [FREE Full text] [doi: 10.1177/1090198113485753] [Medline: 23640122]
63. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. Br J Sports Med 2014 Jul;48(13):1019-1023 [FREE Full text] [doi: 10.1136/bjsports-2014-093546] [Medline: 24782483]
64. Chen J, Berkman W, Bardouh M, Ng CY, Allman-Farinelli M. The use of a food logging app in the naturalistic setting fails to provide accurate measurements of nutrients and poses usability challenges. Nutrition 2019 Jan;57:208-216. [doi: 10.1016/j.nut.2018.05.003] [Medline: 30184514]
65. Carter MC, Burley VJ, Nykjaer C, Cade JE. 'My Meal Mate' (MMM): validation of the diet measures captured on a smartphone application to facilitate weight loss. Br J Nutr 2013 Feb 14;109(3):539-546. [doi: 10.1017/S0007114512001353] [Medline: 22717334]
66. Bucher DT, Carrard I, Farina E, Danuser B, Kruseman M. Development and evaluation of e-CA, an electronic mobile-based food record. Nutrients 2017 Jan 18;9(1):76 [FREE Full text] [doi: 10.3390/nu9010076] [Medline: 28106767]
67. Ambrosini GL, Hurworth M, Giglia R, Trapp G, Strauss P. Feasibility of a commercial smartphone application for dietary assessment in epidemiological research and comparison with 24-h dietary recalls. Nutr J 2018 Jan 09;17(1):5 [FREE Full text] [doi: 10.1186/s12937-018-0315-4] [Medline: 29316930]
68. Rangan AM, O'Connor S, Giannelli V, Yap ML, Tang LM, Roy R, et al. Electronic Dietary Intake Assessment (e-DIA): comparison of a mobile phone digital entry app for dietary data collection with 24-hour dietary recalls. JMIR Mhealth Uhealth 2015 Oct 27;3(4):e98 [FREE Full text] [doi: 10.2196/mhealth.4613] [Medline: 26508282]
69. Ashman AM, Collins CE, Brown LJ, Rae KM, Rollo ME. Validation of a smartphone image-based dietary assessment method for pregnant women. Nutrients 2017 Jan 18;9(1):73 [FREE Full text] [doi: 10.3390/nu9010073] [Medline: 28106758]
70. Lopes TS, Luiz RR, Hoffman DJ, Ferriolli E, Pfrimer K, Moura AS, et al. Misreport of energy intake assessed with food records and 24-h recalls compared with total energy expenditure estimated with DLW. Eur J Clin Nutr 2016 Nov;70(11):1259-1264. [doi: 10.1038/ejcn.2016.85] [Medline: 27273069]
71. Baxter SD, Thompson WO, Davis HC, Johnson MH. Impact of gender, ethnicity, meal component, and time interval between eating and reporting on accuracy of fourth-graders' self-reports of school lunch. J Am Diet Assoc 1997 Nov;97(11):1293-1298. [doi: 10.1016/S0002-8223(97)00309-X] [Medline: 9366868]
72. Baxter SD, Hardin JW, Guinn CH, Royer JA, Mackelprang AJ, Smith AF. Fourth-grade children's dietary recall accuracy is influenced by retention interval (target period and interview time). J Am Diet Assoc 2009 May;109(5):846-856 [FREE Full text] [doi: 10.1016/j.jada.2009.02.015] [Medline: 19394471]
73. Gibson RS, Charrondiere UR, Bell W. Measurement errors in dietary assessment using self-reported 24-hour recalls in low-income countries and strategies for their prevention. Adv Nutr 2017 Nov;8(6):980-991. [doi: 10.3945/an.117.016980] [Medline: 29141979]
74. Thompson FE, Subar AF, Loria CM, Reedy JL, Baranowski T. Need for technological innovation in dietary assessment. J Am Diet Assoc 2010 Jan;110(1):48-51 [FREE Full text] [doi: 10.1016/j.jada.2009.10.008] [Medline: 20102826]
75. Osadchiy T, Poliakov I, Olivier P, Rowland M, Foster E. Progressive 24-hour recall: usability study of short retention intervals in web-based dietary assessment surveys. J Med Internet Res 2020 Feb 03;22(2):e13266 [FREE Full text] [doi: 10.2196/13266] [Medline: 32012055]
76. Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. J Biopharm Stat 2007;17(4):571-582. [doi: 10.1080/10543400701329422] [Medline: 17613642]
77. Olofsen E, Dahan A, Borsboom G, Drummond G. Improvements in the application and reporting of advanced Bland-Altman methods of comparison. J Clin Monit Comput 2015 Feb;29(1):127-139. [doi: 10.1007/s10877-014-9577-3] [Medline: 24806333]

78. Bell BM, Alam R, Alshurafa N, Thomaz E, Mondol AS, de la Haye K, et al. Automatic, wearable-based, in-field eating detection approaches for public health research: a scoping review. NPJ Digit Med 2020;3:38 [FREE Full text] [doi: 10.1038/s41746-020-0246-2] [Medline: 32195373]

Abbreviations

EMA: ecological momentary assessment
LoA: limits of agreement
LORA: Longitudinal Resilience Assessment
PA: physical activity
PROUD: Prevention of Comorbid Depression and Obesity in Attention-Deficit/Hyperactivity Disorder
SUS: System Usability Scale
TEE: total energy expenditure

Edited by R Kukafka; submitted 22.11.20; peer-reviewed by J Reichenberger, I Gabashvili; comments to author 01.01.21; revised version received 12.01.21; accepted 04.05.21; published 05.07.21

Please cite as:
Ruf A, Koch ED, Ebner-Priemer U, Knopf M, Reif A, Matura S
Studying Microtemporal, Within-Person Processes of Diet, Physical Activity, and Related Factors Using the APPetite-Mobile-App: Feasibility, Usability, and Validation Study
J Med Internet Res 2021;23(7):e25850
URL: https://www.jmir.org/2021/7/e25850
doi: 10.2196/25850
PMID:

Multimedia Appendix 1. APPetite-food record and nutritional value generation.

Participants are asked to enter foods and drinks as soon as possible after consuming them. The home-screen of the APPetite-mobile-app displays 5 buttons “breakfast”, “lunch”, “dinner”, “snack”, and “drink”. Pressing the appropriate meal button triggers the initiation of the APPetite-food record. After confirming the intent to start entering a meal as all subsequent steps are mandatory, participants set the time of the meal on a digital 24-hour clock (hh:mm). Participants are then redirected to a list of 16 categories of which they select the applicable ones. The subcategories as well as the food and drink items of each selected category are displayed subsequently. If the button “drink” is pressed, participants can only choose from the drink category. After selecting all relevant food and drink items, each item will be presented again for the assessment of the amount consumed. Participants can specify the amount by entering the number of consumed standard portions (eg, 200 ml glass for water) or an exact amount in milliliters or grams. To ensure participants do not forget to enter certain foods or drinks, a reminder to record everything (eg, milk when having cereals) is presented. After the reminder participants indicate whether they (1) have entered everything, (2) want to add something else from the list, or (3) want to use the Not-on-the-list-option. Within the Not-on-the-list-option participants are asked to describe the food and amount consumed as accurately as possible. In the last step, participants select the predominant reason for the current food or drink intake out of 19 presented reasons. These reasons were adopted from previous studies [1–5], translated to German and adapted to incorporate reasons for drinking (see Multimedia Appendix 2).

At 9 PM, the APPetite-mobile-app initiates an end-of-the-day prompt asking if all consumed foods and drink of the day have been recorded. If participants deny this question, they are requested to add missing meals following the described 6-step process.

The APPetite-food record includes 14 food-related categories, 1 drink category, and a Not-on-the-list-category. The food-related categories are divided into 31 subcategories and into around 500 food items overall. The drink category differentiates a total of 40 drinks in 2 subcategories: nonalcoholic and alcoholic drinks. The food and drink items are predominantly generic. To facilitate the search of certain foods and drinks, some items are present in more than one category (eg, milk in dairy products and nonalcoholic drinks).

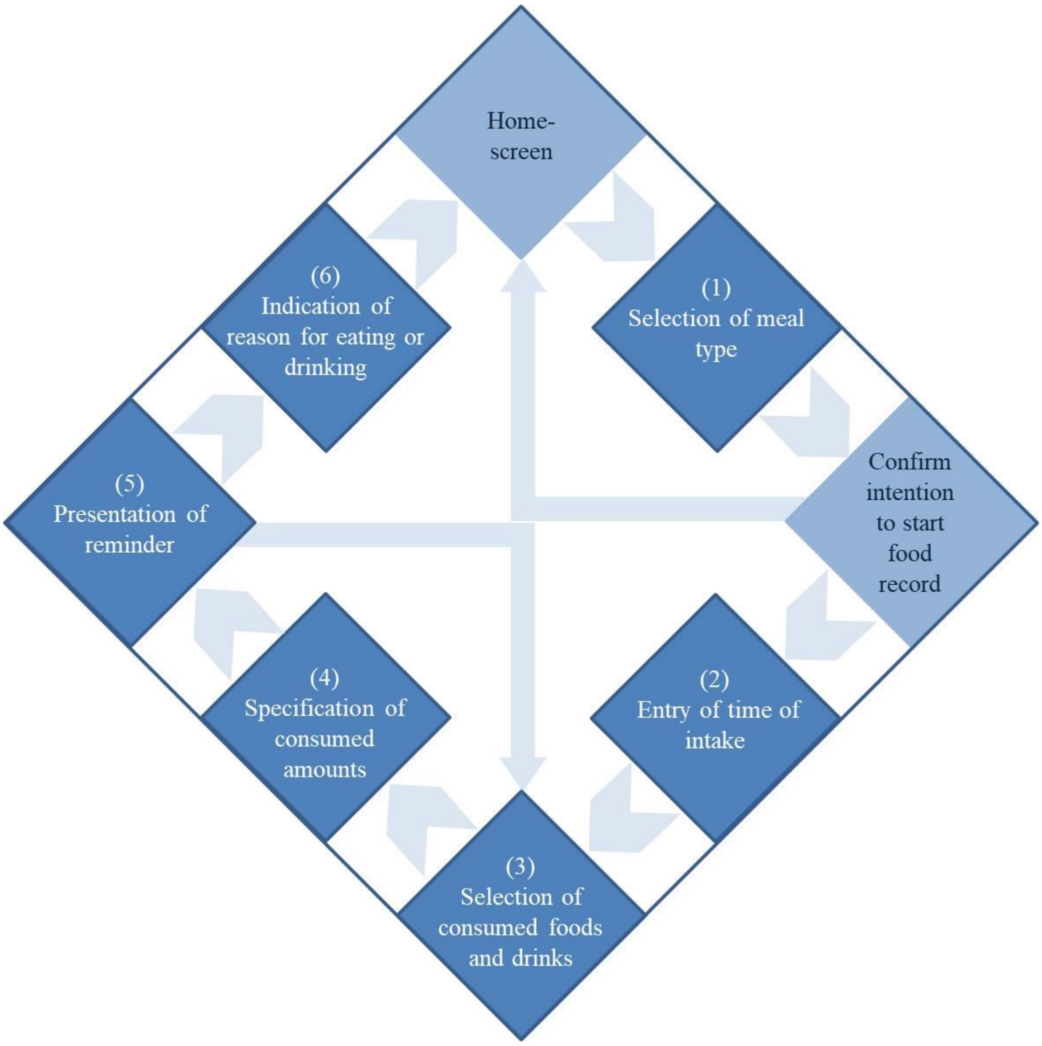


Figure 1. Illustration of the six-step process of the APPetite-food record.

A few additional features were implemented to improve the accuracy of the APPetite-food record. These features address further details about entered foods or the consumption of commonly added foods: (1) Preparation method of fruits and vegetables (eg, raw, cooked, or fried), (2) Caffeine content of coffee and tea beverages, (3) Type of fruit of juices, nectars or spritzers, (4) Added sugar, honey, sweetener pills or sweetener liquid to hot drinks, and (5) Added seeds (eg, sunflower seeds) to salads.

Participants cannot modify food records retrospectively. As compliance of the food record cannot be assessed directly, we provide dietary feedback to our participants to increase motivation to enter all foods and drinks completely and truly.

Nutritional value generation

The APPetite-food record captures complex dietary information. However, it does not allow the automated generation of nutritional values. Therefore, we created a workflow starting with the data download and data preprocessing, followed by the data plausibility check and the data transfer, resulting in the generation of nutritional values. Data preprocessing is done using RStudio extracting all consumed food and drink items and their amounts from the movisensXS output file. The extracted dietary data are then checked for plausibility. Questionable entries are identified (eg, “100 apples”) and reviewed with the participant in the second in-person session. Furthermore, additional information (eg, product brand) on recorded generic foods is acquired if possible. The checked and corrected dietary data are then transferred to myfood24-Germany, a 24-hour dietary recall [6], by trained staff. Data plausibility check, data correction, and data transfer are done according to the 4-eyes principle to minimize data loss and errors.

1. Cleobury L, Tapper K. Reasons for eating “unhealthy” snacks in overweight and obese males and females. J Hum Nutr Diet. 2014;27(4):333–341. PMID: 24134077
2. Verhoeven AAC, Adriaanse MA, de Vet E, Fennis BM, de Ridder DTD. It’s my party and I eat if I want to. Reasons for unhealthy snacking. Appetite. 2015;84:20–27. PMID: 25261101
3. Schüz B, Papadakis T, Ferguson SG. Situation-specific social norms as mediators of social influence on snacking. Heal Psychol. 2018;37(2):153–159. PMID: 29154607
4. Renner B, Sproesser G, Strohbach S, Schupp HT. Why we eat what we eat. The Eating Motivation Survey (TEMS). Appetite. 2012;59(1):117–128. PMID: 22521515
5. Evers C, Adriaanse M, de Ridder DTD, de Witt Huberts JC. Good mood food. Positive emotion as a neglected trigger for food intake. Appetite. 2013;68:1–7. PMID: 23602962
6. Koch SAJ, Conrad J, Hierath L, Hancock N, Beer S, Cade JE, et al. Adaptation and evaluation of myfood24-germany: A web-based self-administered 24-h dietary recall for the german adult population. Nutrients. 2020;12(1):1–15. PMID: 31935885

Multimedia Appendix 2. Reasons for food intake.

Bitte wählen Sie die zutreffendste Aussage aus:

Ich habe gegessen/getrunken, …

… weil ich hungrig/durstig war. [1]

… um später nicht hungrig/durstig zu sein. [1]

… weil ich Energie gebraucht habe. [1]

… weil ich normalerweise zu dieser Zeit esse/trinke. [1]

… weil das Essen/Getränk so verlockend aussah/gerochen hat. [1]

… weil ich traurig war. [1]

… weil mir langweilig war. [1]

… weil ich gestresst war. [1]

… weil ich mich einsam gefühlt habe. [2]

… weil ich glücklich war. [3]

… um jemand anderem/anderen Gesellschaft zu leisten. [1]

… weil ich mich verpflichtet gefühlt habe, zu essen/trinken. [1]

… weil ich nicht aufhören konnte, an Essen/Trinken zu denken. [1]

… weil ich verhindern wollte, dass Essen/Getränke verschwendet/weggeschmissen wird/werden. [1]

… um einen besonderen Anlass zu feiern (Geburtstag, Traditionelles Fest, etc.) [4]

… weil ich Fernsehen/einen Film geschaut habe. [4]

… um mich selbst zu belohnen. [4]

… weil andere gegessen/getrunken haben. [5]

Ich kann mich nicht erinnern, mich entschieden zu haben, zu essen/trinken. – Ich habe einfach gegessen/getrunken. [1]

Please select the most applicable statement:

I ate/drank …

… because I was feeling hungry/thirsty. [1]

… to avoid being hungry/thirsty later. [1]

… because I felt I needed the energy. [1]

… because I usually eat/drink at this time. [1]

… because the food/drink looked/smelt so tempting. [1]

… because I was feeling fed up. [1]

… because I was feeling bored. [1]

… because I was stressed. [1]

… because I felt lonely. [2]

… because I was happy. [3]

… to keep somebody else/other people company. [1]

… because I felt obliged to. [1]

… because I couldn’t stop thinking about food. [1]

… because I wanted to avoid food/drinks going to waste. [1]

… to celebrate a special occasion (birthday, traditional celebration, etc.). [4]

… because I was watching television/a movie. [4]

… to reward myself. [4]

… because others ate/drank. [5]

I don’t recall deciding to eat/drink – I just found myself eating/drinking. [1]

1. Cleobury L, Tapper K. Reasons for eating “unhealthy” snacks in overweight and obese males and females. J Hum Nutr Diet. 2014;27(4):333–341. PMID: 24134077
2. Renner B, Sproesser G, Strohbach S, Schupp HT. Why we eat what we eat. The Eating Motivation Survey (TEMS). Appetite. 2012;59(1):117–128. PMID: 22521515
3. Evers C, Adriaanse M, de Ridder DTD, de Witt Huberts JC. Good mood food. Positive emotion as a neglected trigger for food intake. Appetite. 2013;68:1–7. PMID: 23602962
4. Verhoeven AAC, Adriaanse MA, de Vet E, Fennis BM, de Ridder DTD. It’s my party and I eat if I want to. Reasons for unhealthy snacking. Appetite. 2015;84:20–27. PMID: 25261101
5. Schüz B, Papadakis T, Ferguson SG. Situation-specific social norms as mediators of social influence on snacking. Heal Psychol. 2018;37(2):153–159. PMID: 29154607

Multimedia Appendix 3. Prompt measures and items.

The German items of the prompts which are used within the APPetite-mobile-app are shown in black. The corresponding English items are shown in grey. All changes that have been made to the original items of the cited publications are described and explained in italic. Psychometric properties (McDonald's Omega for between- and within-subject reliability) are reported for all constructs/scales assessed with at least 3 items. The item order is not randomized within the prompts.

Sleep

The first prompt of each day assesses quantity and quality of the previous night's sleep. The time of falling asleep the night before and the time waking up this morning is captured on a digital 24-hour clock (hh:mm). Additionally, participants rate the sleep quality on a 5-point scale from bad to good.

Wann sind Sie gestern eingeschlafen?

hh:mm

What time did you fall asleep yesterday night?

hh:mm

Wie gut haben Sie geschlafen?

Schlecht	Eher schlecht	Teils-teils	Eher gut	Gut

How was your sleep?

Bad	Quite bad	So-so	Quite good	Good

Wann sind Sie heute Morgen aufgewacht?

hh:mm

What time have you woken up today?

hh:mm

Stress expectancy

The intensity of expected stress of the present day is rated in the first prompt of each day on a visual analogue scale from 0 (not stressful at all) to 100 (very stressful).

Wie stressig wird der heutige Tag insgesamt Ihrer Erwartung nach werden?

0 - gar nicht stressig 100 - sehr stressig

How stressful do you expect today to be?

0 – not stressful at all 100 – very stressful

Context

Each prompt captures the context (main activity, social and physical context) a person was in right before the prompt. The assessment was adapted from Dunton, Liao, Intille, Huh, and Leventhal [1].

„In a conversation" was added to the list of main activities as well as "laying" to the last question of the main activity assessment.

Was haben Sie gemacht, bevor das Signal zur Abfrage ertönte? Wählen Sie Ihre Haupttätigkeit.

Lesen / Computer

Fernsehen / Film schauen

Essen / Trinken

Körperliche Aktivität / Sport

Gespräch

Andere Tätigkeit

What were you DOING right before the beep went off [Choose your main activity]?

Reading/Computer

Watching TV/movies

Eating/drinking

Physical activity/exercising

In a conversation

Other

If "Physical activity/exercising" was selected:

Welche Art von Körperlicher Aktivität/Sport haben Sie gemacht? Bitte tippen Sie Ihre Antwort ein.

____

What type of PHYSCIAL ACTIVITY/EXERCISE? Please type in your answer.

If “Other” was selected:

Was war Ihre Tätigkeit?

Telefonieren

Kochen / Hausarbeiten

Auto fahren

Kinderbetreuung / Kindern helfen

Andere Tätigkeit

English:

What was this OTHER activity?

Talking on the phone

Cooking/chores

Riding in a car

Childcare/helping children

Something else

If “Something else” was selected:

Haben / Sind Sie …?

gelegen

gesessen

gestanden

gegangen

gejoggt / gerannt

Were you …?

laying

sitting

standing

walking

jogging/running

Waren Sie alleine?

ja

nein

English:

Were you alone?

Yes

No

If "No" was selected:

Mit wem waren Sie zusammen?

Partner / Partnerin

Kind(er)

Andere Familienmitglieder

Freund(e)/in(nen)

Arbeitskollege(n)/Arbeitskollegin(nen)

Andere Bekannte

Person(en), die ich nicht kenne

Who were you with?

Spouse

Child(ren)

Other family members

Friend(s)

Coworkers

Other type of acquaintances

People I did not know

Wo waren Sie, bevor das Signal zur Abfrage ertönte?

Zuhause (drinnen)

Zuhause (draußen)

Auf der Arbeit (drinnen)

Draußen (nicht Zuhause)

Auto / Transporter / LKW

Anderer Ort

English:

WHERE were you just before the beep went off?

Home (indoors)

Home (outdoors)

Work (indoors)

Outdoors (not at home)

Car/van/truck

Other

If “outdoors (not at home)” was selected:

Wo waren Sie draußen? Bitte tippen Sie Ihre Antwort ein.

WHERE were you OUTDOORS just before the beep went off? Please type in your answer.

If “Other” was selected:

Wo waren Sie? Bitte tippen Sie Ihre Antwort ein.

Where were you? Please type in your answer.

Mood (MDBF)

Affect is assessed using an EMA version of the Multidimensional Mood State Questionnaire (Mehrdimensionaler Befindlichkeitsfragebogen [2]). The author of this version provided us with an improved revision of the German questionnaire. Participants are instructed to rate their mood in the moment “right before the prompt”. McDonald’s Omegas for the scale valence are 0.829 (within) and 0.983 (between). The scale calmness shows McDonald’s Omegas of 0.772 (within) and 0.970 (between).

The revised version measures the three scales “valence”, “calmness” and “energetic arousal” with 8 instead of 6 bipolar items. One item was added to the valence scale (gut – schlecht; English: good – bad) and one to the calmness scale (aufgeregt – gelassen; English: aroused – composed) to increase reliability of these scales when measuring situational change. The answer format was changed from a 7-point scale to an 8-point scale, since the “neutral” response option was excluded and the scale endpoints “extreme” were added. This was done as the previous endpoints “very” were chosen extensively resulting in a negative skew of the scales valence and calmness. The adjective “energiegeladen” (English: full of energy) was exchanged for the more commonly used term “energievoll”.

Wie fühlen Sie sich jetzt im Moment? (Moment vor Beginn der aktuellen Befragung, nicht der Moment der Befragung selbst.)

How do you feel right now? (Moment right before the prompt, not moment of prompt itself.)

Unwohl Wohl

Extrem	Sehr	Ziemlich	Eher	Eher	Ziemlich	Sehr	Extrem

Unwell Well

Extremely	Very	Quite	Rather	Rather	Quite	Very	Extreme

Entspannt Angespannt

Extrem	Sehr	Ziemlich	Eher	Eher	Ziemlich	Sehr	Extrem

Relaxed Tense

Extremely	Very	Quite	Rather	Rather	Quite	Very	Extreme

Müde Wach

Extrem	Sehr	Ziemlich	Eher	Eher	Ziemlich	Sehr	Extrem

Tired Awake

Extremely	Very	Quite	Rather	Rather	Quite	Very	Extreme

Zufrieden Unzufrieden

Extrem	Sehr	Ziemlich	Eher	Eher	Ziemlich	Sehr	Extrem

Content Discontent

Extremely	Very	Quite	Rather	Rather	Quite	Very	Extreme

Unruhig Ruhig

Extrem	Sehr	Ziemlich	Eher	Eher	Ziemlich	Sehr	Extrem

Agitated Calm

Extremely	Very	Quite	Rather	Rather	Quite	Very	Extreme

Energiegeladen Energielos

Extrem	Sehr	Ziemlich	Eher	Eher	Ziemlich	Sehr	Extrem

Full of energy Without energy

Extremely	Very	Quite	Rather	Rather	Quite	Very	Extreme

Aufgeregt							Gelassen
Extrem	Sehr	Ziemlich	Eher	Eher	Ziemlich	Sehr	Extrem

Aroused							Composed
Extremely	Very	Quite	Rather	Rather	Quite	Very	Extreme

Gut							Schlecht
Extrem	Sehr	Ziemlich	Eher	Eher	Ziemlich	Sehr	Extrem

Good							Bad
Extremely	Very	Quite	Rather	Rather	Quite	Very	Extreme

Stress

Subjective stress since the last prompt is assessed using 3 items. The items were adapted from Reichenberger et al. [3] making 3 changes. The first item captures how stressed the participant was since the last prompt. Responses are rated on a visual analogue scale from 0% (not at all) to 100% (very stressed). The other 2 stress items from the Perceived Stress Scale (PSS [4]) assess whether the participants felt that they “could not cope with all the things they had to do” and whether they are “on top of things” on a visual analogue scale from 0 (not at all) to 100 (very much). McDonald’s Omegas for the 3 stress items are 0.658 (within) and 0.923 (between).

Participants are also asked if a stressor occurred since the last prompt. If affirmed, participants are requested to describe the stressor. An additional question assesses the changeability of the stressor and whether the participant performed an action to change it.

Since the main outcome measure of the Eat2beNICE-APPetite-study (impulsivity) is assessed for a time interval (since the last prompt), stress is assessed accordingly and not momentary as in Reichenberger et al. [3]. *The first item captures how stressed the participant was since the last prompt, excluding the adjective “nervous”. Responses are rated on a visual analogue scale from 0% (not at all) to 100% (very stressed). “Not” was added to the second stress item in accordance with the original Perceived Stress Scale (PSS* [4]*): “Do you feel that you could not cope with all the things you had to do?”.*

Wie gestresst waren Sie seit der letzten Abfrage?

0% überhaupt nicht gestresst 100% sehr gestresst

How stressed have you been since the last prompt?

0% not stressed at all 100% very stressed

Wie stark hatten Sie seit der letzten Abfrage das Gefühl, mit all den anstehenden Aufgaben und Problemen <u>nicht</u> richtig umgehen zu können?

0 gar nicht 100 sehr stark

Do you feel that you could <u>not</u> cope with all the things you had to do?

0 not at all 100 very much

Wie stark hatten Sie seit der letzten Abfrage das Gefühl, alles im Griff zu haben?

0 gar nicht 100 sehr stark

Do you feel that you're on top of things?

0 not at all 100 very much

Ist seit der letzten Abfrage ein stressiges/belastendes Ereignis eingetreten? Ein stressiges/belastendes Ereignis ist jedes Ereignis, möge es noch so geringfügig sein, das negative Auswirkungen auf Sie hat.

Ja Nein

Has a stressful event happened since the last prompt? A stressful event is any event, no matter how small, that has a negative impact on you.

Yes No

If "Yes" was selected:

Bitte beschreiben Sie das stressige Ereignis:

____

Please describe the stressful event:

____

If "Yes" was selected:

War das stressige Ereignis durch Sie selbst veränderbar?

Nein, ich konnte es nicht verändern.

Ja, ich konnte es verändern und habe dies getan.

Ja, ich hätte es verändern können. Meine Versuche, es zu verändern, waren jedoch nicht erfolgreich.

Ja, ich hätte es verändern können, habe es jedoch nicht versucht.

Was the stressful event modifiable for you?

No, I could not modify it.

Yes, I could modify it and did so.

Yes, I could have modified it. However, my attempts to modify it weren't successful.

Yes, I could have modified it, but did not try to.

Impulsivity

State impulsivity is measured with the Momentary Impulsivity Scale (MIS [5]). The MIS consists of 4 items which are rated on a 5-point scale on how well the statement describes an individual's behavior, cognition, and experiences since the last prompt. The MIS has a McDonald's Omega of 0.485 (within) and of 0.833 (between).

Ich habe Dinge gesagt, ohne vorher nachzudenken.

1 = nicht zutreffend; 2 = eher nicht zutreffend; 3 = teils-teils; 4 = eher zutreffend; 5 = zutreffend

"I said things without thinking"

1 = very slightly or not at all; 2 = a little; 3 = moderately; 4 = quite a bit; 5 = extremely

Ich habe mehr Geld ausgegeben als ich sollte.

1 = nicht zutreffend; 2 = eher nicht zutreffend; 3 = teils-teils; 4 = eher zutreffend; 5 = zutreffend

"I spent more money than I meant to"

1 = very slightly or not at all; 2 = a little; 3 = moderately; 4 = quite a bit; 5 = extremely

Ich war ungeduldig.

1 = nicht zutreffend; 2 = eher nicht zutreffend; 3 = teils-teils; 4 = eher zutreffend; 5 = zutreffend

"I have felt impatient"

1 = very slightly or not at all; 2 = a little; 3 = moderately; 4 = quite a bit; 5 = extremely

Ich habe eine unüberlegte Entscheidung getroffen.

1 = nicht zutreffend; 2 = eher nicht zutreffend; 3 = teils-teils; 4 = eher zutreffend; 5 = zutreffend

"I made a 'spur of the moment' decision"

1 = very slightly or not at all; 2 = a little; 3 = moderately; 4 = quite a bit; 5 = extremely

Food availability

Each prompt assesses the food availability since the last prompt on a visual analogue scale from 0 (not available at all) to 100 (easily available).

Wie leicht verfügbar war Essen für Sie seit der letzten Abfrage?

0 - gar nicht verfügbar 100 - sehr leicht verfügbar

How easily available was food for you since the last prompt?

0 – not available at all 100 – very easily available

1. Dunton GF, Liao Y, Intille S, Huh J, Leventhal A. Momentary Assessment of Contextual Influences on Affective Response During Physical Activity. Heal Psychol. 2015;34(12):1145–1153. PMID: 26053885
2. Wilhelm P, Schoebi D. Assessing mood in daily life: Structural validity, sensitivity to change, and reliability of a short-scale to measure three basic dimensions of mood. Eur J Psychol Assess. 2007;23(4):258–267.
3. Reichenberger J, Kuppens P, Liedlgruber M, Wilhelm FH, Tiefengrabner M, Ginzinger S, et al. No haste, more taste: An EMA study of the effects of stress, negative and positive emotions on eating behavior. Biol Psychol. 2018;131:54–62. PMID: 27654506
4. Cohen S, Kamarck T, Mermelstein R. A Global Measure of Perceived Stress. J Health Soc Behav. 1983;24(4):385–396. PMID: 6668417
5. Tomko RL, Solhan MB, Carpenter RW, Brown WC, Jahng S, Wood PK, et al. Measuring impulsivity in daily life: The momentary impulsivity scale. Psychol Assess. 2014;26(2):339–349. PMID: 24274047

Multimedia Appendix 4. Bland-Altman analysis of protein, fat, carbohydrate, sugar, and fiber intake.

Protein

The mean difference of protein intake was not normally distributed. Therefore, natural log transformation was performed, as recommended by Bland and Altman (1986) [1]. The antilogs of the limits were taken and multiplied by 100 to allow interpretation of the ln-transformed data as percentages (100%=ideal agreement).

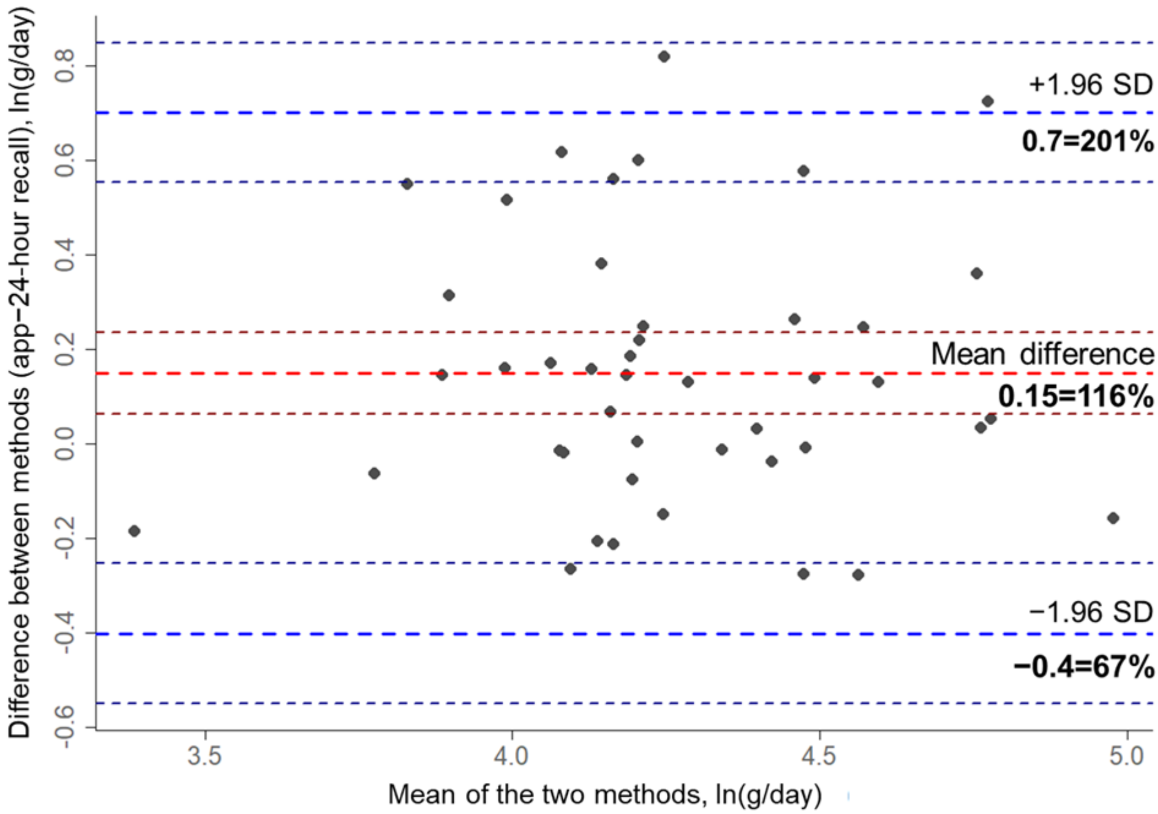


Figure 1. Bland-Altman plot assessing agreement between habitual protein intake in g per day captured by the APPetite-food record and the 24-hour recall (red line: mean difference=app –24-hour recall; dark red lines: 95% CI of mean difference; blue lines: lower and upper limits of agreement; dark blue lines: 95% CI of lower and upper limits of agreement).

Mean difference: 0.15=116% (95% CI 0.06 to 0.24=106.6% to 126.5%)
Lower limits of agreement: –0.4=67% (95% CI –0.55 to –0.25=57.7% to 77.6%)
Upper limits of agreement: 0.7=201% (95% CI 0.55 to 0.85=137.8% to 233.7%)

The limits of agreement transcended the predefined acceptable limits of agreement of 10% considerably.

Fat

The mean difference of fat intake was not normally distributed. The same procedure as described for protein intake was used.

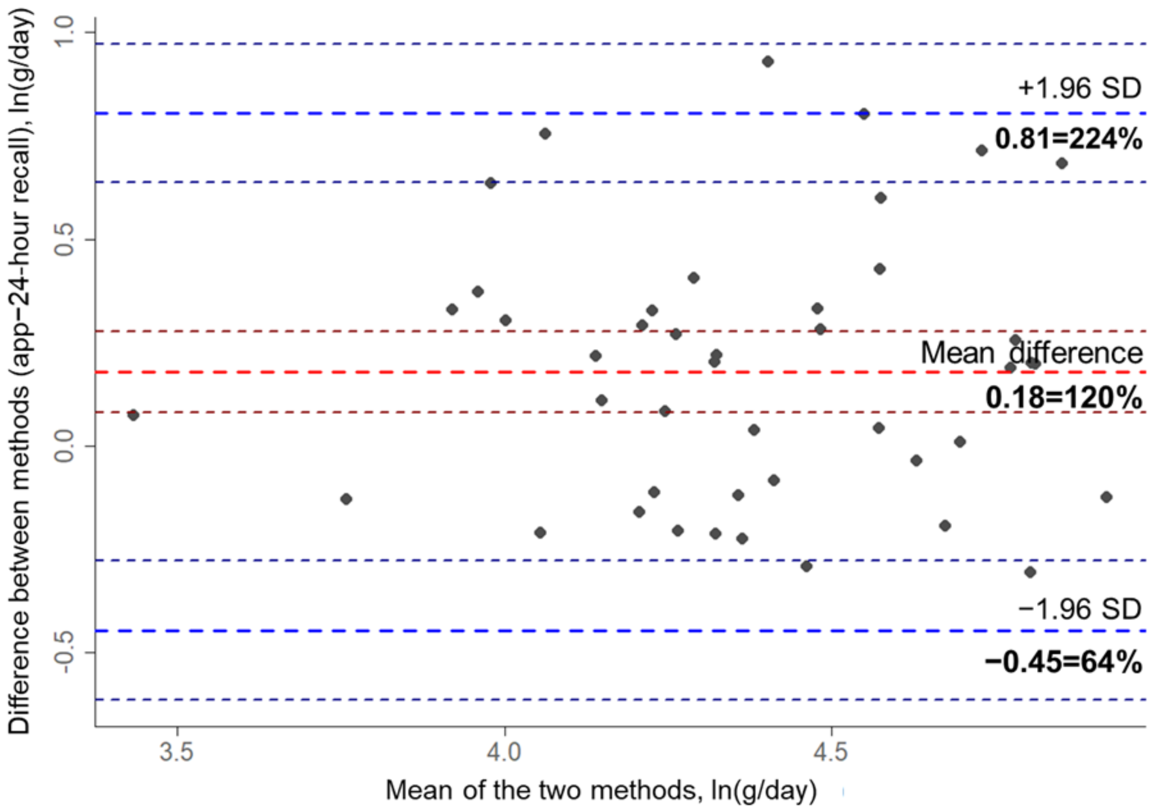


Figure 2. Bland-Altman plot assessing agreement between habitual fat intake in g per day captured by the APPetite-food record and the 24-hour recall (red line: mean difference=app–24-hour recall; dark red lines: 95% CI of mean difference; blue lines: lower and upper limits of agreement; dark blue lines: 95% CI of lower and upper limits of agreement).

Mean difference: 0.18=120% (95% CI 0.08 to 0.28=108.6% to 131.8%)

Lower limits of agreement: –0.45=64% (95% CI –0.61 to –0.28=54.1% to 75.7%)

Upper limits of agreement: 0.81=224% (95% CI 0.64 to 0.97=189% to 264.5%)

The limits of agreement exceeded the predefined acceptable limits of agreement of 10% substantially.

Carbohydrates

The mean difference of carbohydrates intake was normally distributed.

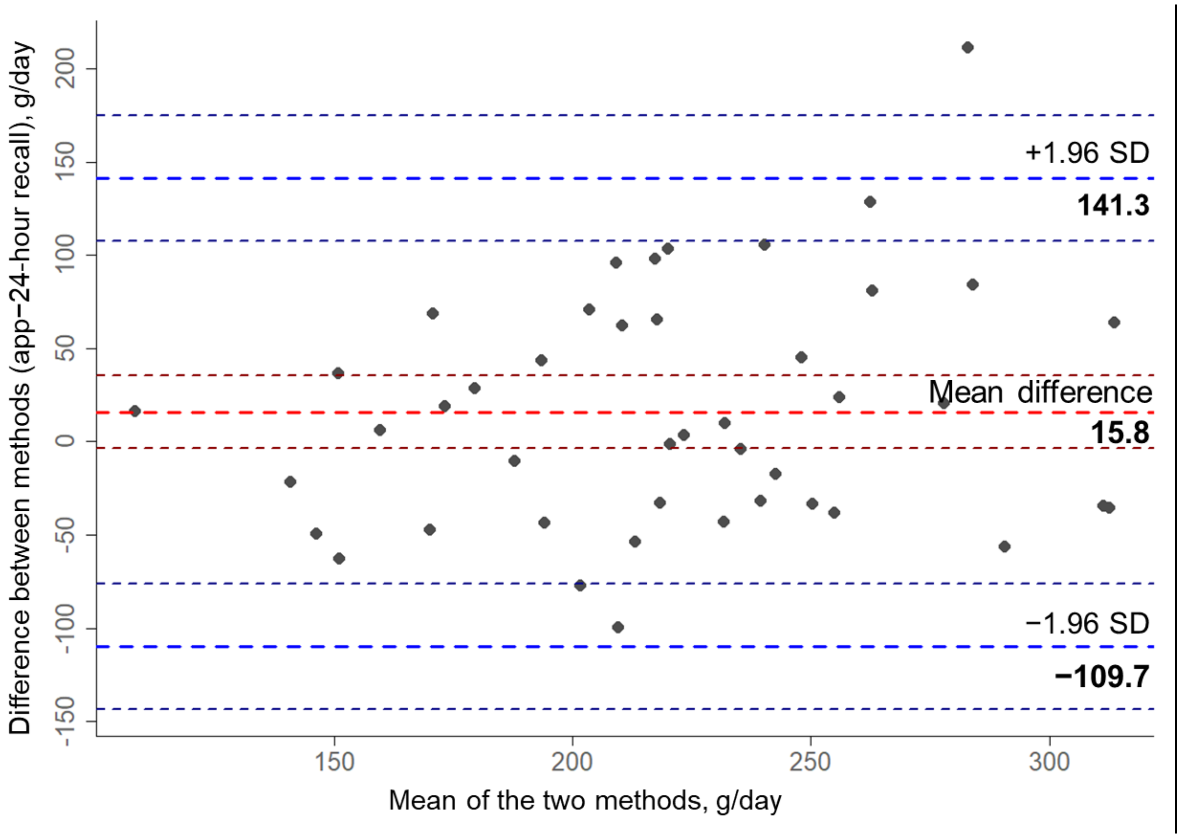


Figure 3. Bland-Altman plot assessing agreement between habitual carbohydrate intake in g per day captured by the APPetite-food record and the 24-hour recall (red line: mean difference=app–24-hour recall; dark red lines: 95% CI of mean difference; blue lines: lower and upper limits of agreement; dark blue lines: 95% CI of lower and upper limits of agreement).

Mean difference: 15.8 (95% CI –3.7 to 35.3)

Lower limits of agreement: –109.7 (95% CI –143.4 to –75.97)

Upper limits of agreement: 141.3 (95% CI 107.6 to 175.1)

The pre-established acceptable limits of agreement were 22.1 g, which were exceeded considerably.

Sugar

Normal distribution of the differences of sugar intake was given.

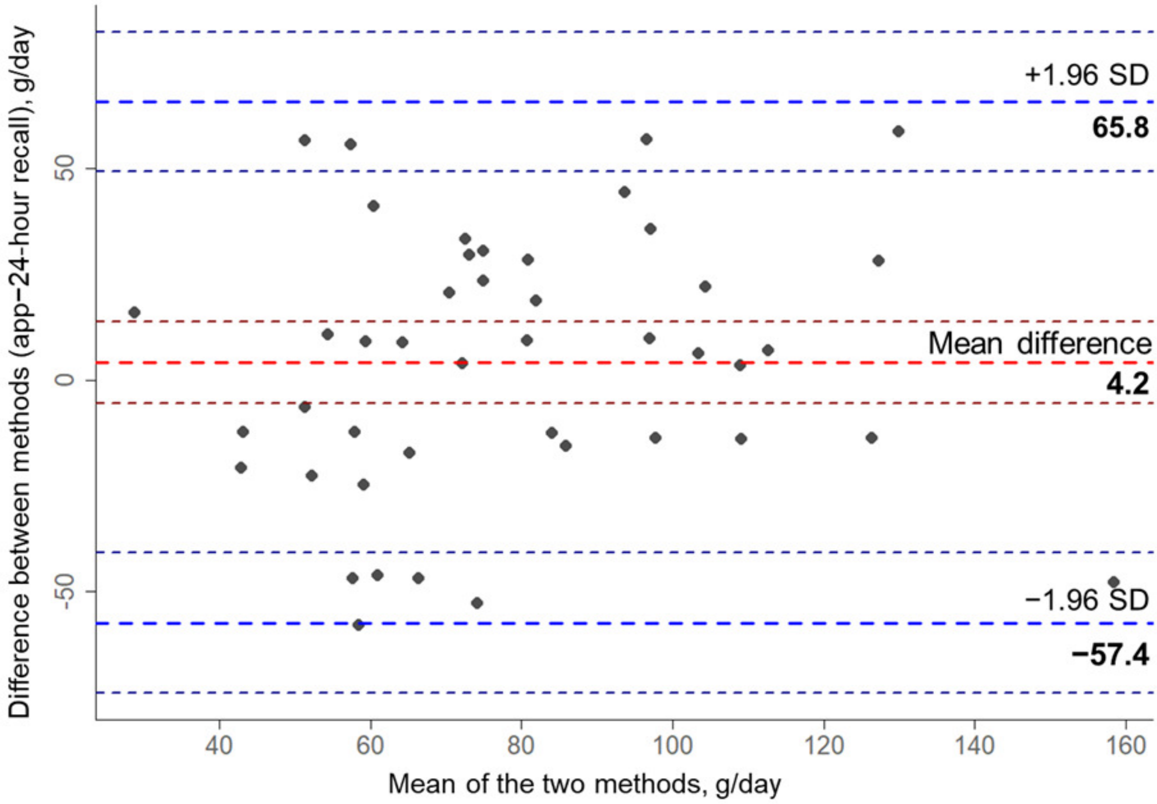


Figure 4. Bland-Altman plot assessing agreement between habitual sugar intake in g per day captured by the APPetite-food record and the 24-hour recall (red line: mean difference=app–24-hour recall; dark red lines: 95% CI of mean difference; blue lines: lower and upper limits of agreement; dark blue lines: 95% CI of lower and upper limits of agreement).

Mean difference: 4.2 (95% CI –5.3 to 13.8)

Lower limits of agreement: –57.4 (95% CI –73.9 to –40.8)

Upper limits of agreement: 65.8 (95% CI 49.3 to 82.4)

Disagreement of 7.9 g was predefined as acceptable, but was transcended substantially.

Fiber

The difference of fiber intake between the two methods was normally distributed.

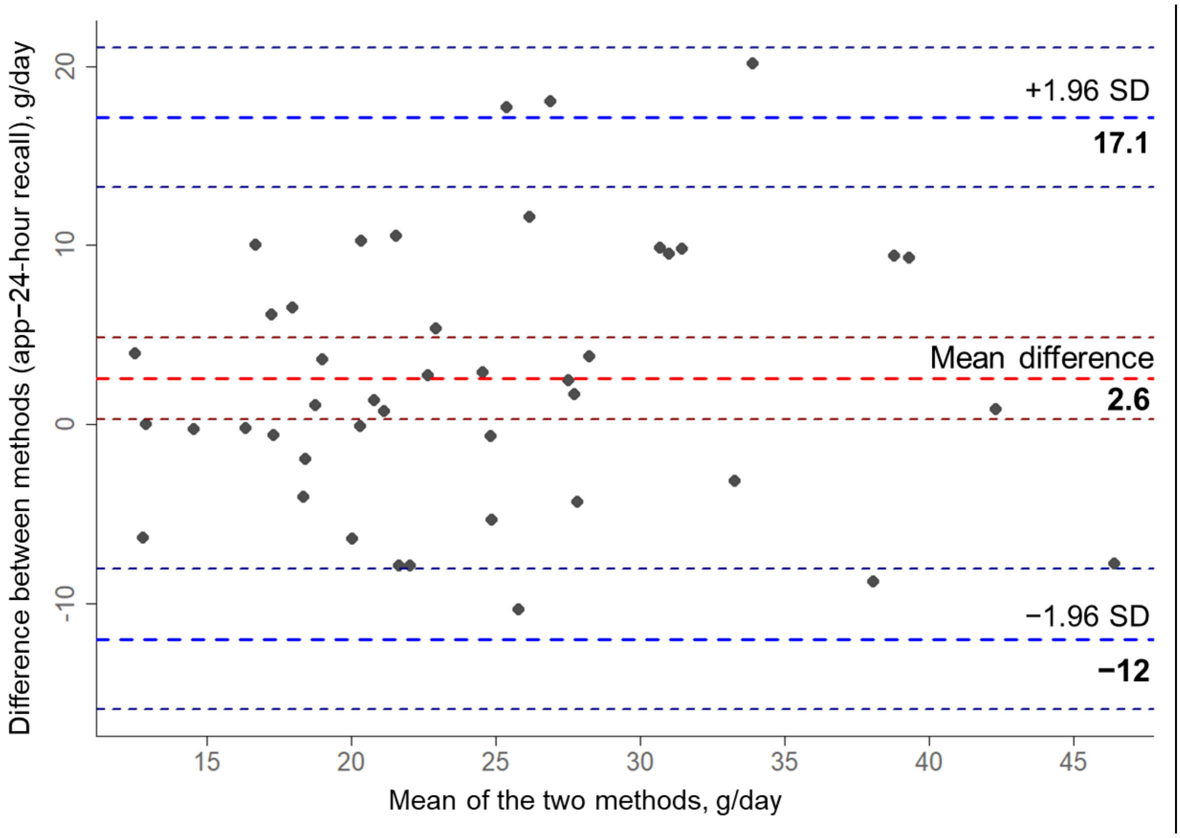


Figure 5. Bland-Altman plot assessing agreement between habitual fiber intake in g per day captured by the APPetite-food record and the 24-hour recall (red line: mean difference=app–24-hour recall; dark red lines: 95% CI of mean difference; blue lines: lower and upper limits of agreement; dark blue lines: 95% CI of lower and upper limits of agreement).

Mean difference: 2.6 (95% CI 0.3 to 4.8)

Lower limits of agreement: –12 (95% CI –15.9 to –8.1)

Upper limits of agreement: 17.1 (95% CI 13.2 to 21.1)

The limits of agreement exceeded the predefined acceptable disagreement of 2.5 g.

1. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307–310. PMID: 2868172

……………………………………………………………………………………………………...

Paper 2

Ruf, A., Neubauer, A.B., Ebner-Priemer, U., Reif, A., & Matura, S. (2021). Studying dietary intake in daily life through multilevel two-part modelling: a novel analytical approach and its practical application. *International Journal of Behavioral Nutrition and Physical Activity, 18*(130). https://doi.org/10.1186/s12966-021-01187-8

……………………………………………………………………………………………………...

Ruf *et al. Int J Behav Nutr Phys Act (2021) 18:130*
https://doi.org/10.1186/s12966-021-01187-8

International Journal of Behavioral Nutrition and Physical Activity

METHODOLOGY

Open Access

Studying dietary intake in daily life through multilevel two-part modelling: a novel analytical approach and its practical application

Alea Ruf[1*], Andreas B. Neubauer[2,3], Ulrich Ebner-Priemer[4,5], Andreas Reif[1] and Silke Matura[1]

Abstract

Background: Understanding which factors influence dietary intake, particularly in daily life, is crucial given the impact diet has on physical as well as mental health. However, a factor might influence whether but not how much an individual eats and vice versa or a factor's importance may differ across these two facets. Distinguishing between these two facets, hence, studying dietary intake as a dual process is conceptually promising and not only allows further insights, but also solves a statistical issue. When assessing the association between a predictor (e.g. momentary affect) and subsequent dietary intake in daily life through ecological momentary assessment (EMA), the outcome variable (e.g. energy intake within a predefined time-interval) is semicontinuous. That is, one part is equal to zero (i.e. no dietary intake occurred) and the other contains right-skewed positive values (i.e. dietary intake occurred, but often only small amounts are consumed). However, linear multilevel modelling which is commonly used for EMA data to account for repeated measures within individuals cannot be applied to semicontinuous outcomes. A highly informative statistical approach for semicontinuous outcomes is multilevel two-part modelling which treats the outcome as generated by a dual process, combining a multilevel logistic/probit regression for zeros and a multilevel (generalized) linear regression for nonzero values.

Methods: A multilevel two-part model combining a multilevel logistic regression to predict whether an individual eats and a multilevel gamma regression to predict how much is eaten, if an individual eats, is proposed. Its general implementation in R, a widely used and freely available statistical software, using the R-package brms is described. To illustrate its practical application, the analytical approach is applied exemplary to data from the Eat2beNICE-APPetite-study.

Results: Results highlight that the proposed multilevel two-part model reveals process-specific associations which cannot be detected through traditional multilevel modelling.

Conclusions: This paper is the first to introduce multilevel two-part modelling as a novel analytical approach to study dietary intake in daily life. Studying dietary intake through multilevel two-part modelling is conceptually as well as methodologically promising. Findings can be translated to tailored nutritional interventions targeting either the occurrence or the amount of dietary intake.

*Correspondence: alea.ruf@kgu.de
[1] Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital, Goethe University, Heinrich-Hoffmann-Straße 10, 60528 Frankfurt am Main, Germany
Full list of author information is available at the end of the article

Keywords: Multilevel two-part modelling, Semicontinuous, Longitudinal, Dietary intake, Ecological momentary assessment, R, Brms

Background

Which factors influence whether an individual eats? Which factors influence how much an individual eats? These two questions might be answered differently. For instance, a study found that inhibitory control predicted how much individuals snacked, whereas it did not predict whether individuals snacked [1]. These findings emphasize the dual character of dietary intake. Understanding which factors drive an individual to eat as well as which factors influence how much an individual eats, particularly in daily life, is crucial given the impact diet has on physical as well as mental health.

Diet is a repeated-occurrence health behaviour which is performed multiple times per day [2]. Studying eating behaviour through ecological momentary assessment (EMA) several times a day in natural environments when and "where the action takes place" [3] is a promising and increasingly popular approach [4, 5]. Dietary intake is influenced by a variety of dynamic factors and their interactions [6] which cannot be replicated reliably in a laboratory setting, highlighting the need for EMA.

Studying dietary intake as a dual process in daily life is conceptually promising and not only allows novel insights, but also solves a statistical issue.

Distributional characteristic of dietary data in EMA studies

EMA studies allow investigating whether individual and/or situational factors (e.g. momentary affect) assessed multiple times per day predict dietary intake (e.g. energy/sugar/fat intake) within a subsequent predefined time-interval (e.g. within the next 2 h). However, dietary intake typically does not occur within each predefined time-interval (e.g. no intake in 46% of 2-h-time-intervals [7]) or only a small amount is consumed (e.g. a snack). This results in an outcome that is zero-inflated (i.e. contains a large proportion of zeros) and right-skewed (i.e. contains a large proportion of small positive values concentrated on the left of the distribution) (see Fig. 1). This type of data is often referred to as semicontinuous.

Traditional statistical approach for EMA data

A common statistical approach to analyse EMA data is linear multilevel modelling (also known as linear mixed or linear hierarchical modelling). It accounts for dependency among longitudinal data due to repeated measures within the same participant and allows studying effects on the level of moments (within-person fluctuations) and individuals (between-person differences). However, traditional linear multilevel modelling cannot be applied to semicontinuous outcomes, as the assumption of normally distributed residuals is likely violated.[1]

Using traditional linear multilevel models without accounting for the large proportion of zeros can lead to incorrect inferences and conclusions and overlooks the dual character of semicontinuous data. For instance, Baldwin et al. [8] showed that a traditional linear multilevel model falsely indicated that there was no change in the semicontinuous outcome daily physical activity (PA) over the course of the study, not detecting that with time participants became less likely to engage in PA. An alternative approach to study semicontinuous data using traditional models is to exclude time-intervals with zeros, i.e. include only time-intervals in which dietary intake occurred to study how much but not whether dietary intake occurred. However, this causes loss of important information [9] and can cause bias in the parameter estimates [10, 11] (as outlined in the discussion). Furthermore, a common approach is to study whether but not how much dietary intake occurred through multilevel logistic regressions (e.g. [7, 12–14]). However, if the amounts of the consumed foods/drinks are also captured, available data with important implications are disregarded.

Statistical approach for semicontinuous outcomes

A generally less known, but highly informative statistical approach for semicontinuous outcomes is two-part modelling. It treats zeros and nonzeros of the outcome separately as generated by a dual process. The zero part (occurrence indicator – e.g. has an individual eaten in a given time-interval?) and continuous/positive part (intensity indicator – e.g. if an individual ate in a given time-interval, how much was eaten?) of the outcome follow different distributions. Two-part models combine these two distributions: a logistic or probit regression for zeros (e.g. to predict whether an individual eats) and a linear or generalized linear regression for positive values (e.g. to predict how much is eaten, if an individual eats).

Two-part modelling assesses these two parts (e.g. the probability of eating and the amount that is eaten) while

[1] It is important to note that a normal distribution of the dependent variable itself is not an assumption of linear multilevel models. Instead, the residuals which reflect the unexplained part of the dependent variable have to follow a normal distribution. However, in practice the distribution of the residuals typically looks similar to the distribution of the dependent variable.

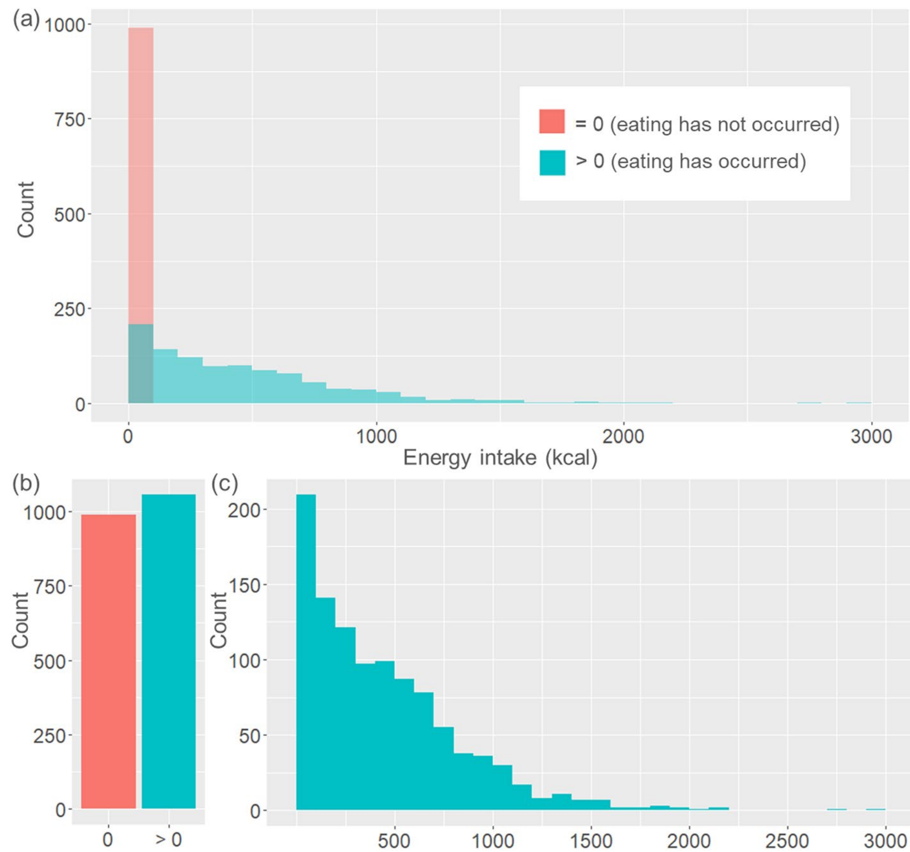


Fig. 1 Illustration of distributional characteristics of a semicontinuous dietary outcome (exemplary: Energy intake in kilocalories [kcal] of the data used in this paper), **a** overall distribution containing a large number of zeros as well as right-skewed positive values, **b** distribution of zeros vs. nonzero values, **c** distribution of the right-skewed positive values

accounting for the potential dependency between them. The importance of taking this potential dependency into account was highlighted by Olsen and Schafer who were the first to extend these models to longitudinal data [15].

Hence, multilevel two-part modelling not only allows studying dietary intake as a dual process, but also overcomes the challenges of semicontinuous data. It does not overlook relevant information and provides additional and novel insights. It differentiates between factors either influencing the occurrence or the amount of dietary intake or both. If both, it can be assessed whether a factor's importance differs across the two parts.

Even though the use of two-part models is less common in most research fields, it has become popular for example in the following fields: Medical costs [16, 17], substance use disorder [18–22] and PA [8, 23–25]. Two-part models have also been applied to nutritional data in order to estimate usual intake of episodically consumed foods [26]. However, to the best of our knowledge, multilevel two-part modelling has not yet been applied to studying dietary intake in daily life. Furthermore, most publications on multilevel two-part modelling used statistical software which is less common (e.g. WinBUGS [22]) or not free to use (e.g. SAS Proc NLMIXED [16], "gsem" command in Stata [8, 25]).

Objective

This paper is the first to introduce multilevel two-part modelling as a novel analytical approach to study dietary intake in daily life. We believe that the importance of multilevel two-part models in behavioural nutrition as well as other behavioural research fields (e.g. PA) is growing. Practical guidance is needed to facilitate the implementation of these rather complex models, particularly in commonly used and freely available software. For this reason, this paper proposes a multilevel two-part model combining a multilevel logistic and a multilevel gamma regression to study dietary intake in daily life using R [27], one of the most commonly used data software programs which is freely available and therefore accessible to everyone. In the present work, we use the R-package brms [28, 29] which is based on Bayesian inference. We chose this package because it allows great flexibility in this specific model. Furthermore, its syntax is very similar to the syntax of other and likely more widely used multilevel packages in R (nlme [30]; lme4 [31]). This has the benefit that readers familiar with multilevel modelling in R can more easily build upon prior experience. We assume that readers have basic knowledge of multilevel modelling (e.g. multilevel structure of the data, random effects). Readers not familiar with these basic concepts are referred to introductory literature on multilevel modelling (e.g. [32, 33]). To ensure readers who are new to Bayesian statistics are able to follow, the basic concept of Bayesian inference is briefly introduced in Additional file 1a.

The aim of this paper is to introduce multilevel two-part modelling as a novel analytical approach to study dietary intake in daily life and provide easy-to-follow guidance on its practical application. To do so, the methods section covers (1) general model specifications of the proposed model, (2) a brief overview of brms and the general implementation of the proposed model in brms and (3) the description of the data used in this paper. The results section outlines the results of the exemplary analyses in detail, in order to provide practical guidance on the model specification and interpretation. Data and R code are provided in Additional files 2 and 3.

Methods

Multilevel two-part model for semicontinuous dietary data

In order to study dietary intake in daily life, we propose a multilevel two-part model which combines a multilevel logistic regression for zeros to predict whether an individual eats and a multilevel gamma regression for right-skewed positive values to predict how much is eaten, if an individual eats. Here, repeated assessments (Level 1) of the semicontinuous variable dietary intake are nested within individuals (Level 2). We chose the multilevel gamma regression for positive values as it does not require data transformation (e.g. logarithmizing) and beyond that performed well for right-skewed continuous PA data in Baldwin et al. [8]. A gamma distribution is a continuous probability distribution which is commonly used to model continuous variables which can only be positive and follow a skewed distribution.

In the following we briefly introduce the model specifications. A more comprehensive introduction to the model specifications can be found in Additional file 4.

The variable y_{ij} represents the semicontinuous dietary intake response from subject j ($j=1, \ldots, m$) at time point i ($i=1, \ldots, n_i$). We are interested in two parts of this variable: (1) Did the participant eat? In other words, is $y_{ij}=0$ or $y_{ij}>0$ (illustrated in Fig. 1b)? (2) If the participant ate, how much was eaten? In other words, what is the expected value of y_{ij}, if $y_{ij}>0$ (illustrated in Fig. 1c)?

A multilevel logistic regression is used for part (1) of the semicontinuous variable. It predicts the log-odds of no eating for person j at time point i ($\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right)$).[2] Figure 2 shows that the log-odds of no eating can be predicted as a function of Level-1 and Level-2 covariates.

A multilevel gamma regression is used for part (2) of the semicontinuous variable. It predicts the expected log amount of dietary intake of person j at time point i ($log(\mu_{ij})$) when eating occurred. μ_{ij} is modelled on the log scale due to the fact that the gamma distribution only supports positive values. However, it is important to highlight that this does not change the data as would log-transforming the data before running the model. The metric of the variable remains unchanged and the slope coefficients can be interpreted through exponentiation (demonstrated in the results). Figure 2 shows that a function of Level-1 and Level-2 covariates can be used to predict the (log) amount of dietary intake.

$L1predictor_{ij}$ in Fig. 2 represents a Level-1 covariate assessed at time point i in person j, e.g. participant j's momentary affect at measurement occasion i. $L2predictor_j$ is a Level-2 covariate of person j, e.g. participant j's BMI. β_{00} and β_{10} are the overall intercepts. The coefficients β_{01} and β_{11} represents the expected change for a one-unit increase in *L1predictor*. The expected change for a one-unit increase in *L2predictor* is expressed by β_{02} and β_{12}. u_{0j} and u_{1j} represent the random intercepts of person j, i.e. person-specific deviation from the overall intercept. u_{01j} and u_{11j} denote the random effects of *L1predictor* in person j, i.e. person-specific differences in the effect of *L1predictor*. The error term ε_{ij} in the continuous part of the model denotes the Level-1 residual, i.e. difference between the predicted value and the observed value of person j at time point i. The first subscript 0 or 1 of the parameters indicates that the equation refers to the zero or the continuous part of the model, respectively. Part specific interpretations of the parameters can be found in Table 1.

The two processes modelled through the multilevel logistic and gamma regression are likely not independent. Therefore, an important consideration in two-part modelling, as highlighted by Olsen and Schafer [15] for longitudinal data, is whether an individual's average probability of eating is related to the individual's average amount consumed when the individual eats. In other words, the average proportion of occasions on which the participant does not eat may be related to the average (log) amount of dietary intake during eating occasions. To account for this potential relation, the correlation

[2] Note that the multilevel logistic regression predicts NO dietary intake (i.e. $y_{ij}=0$). Typically logistic regressions predict $y=1$. π_{ij} represents the probability not to eat for person j at time point i.

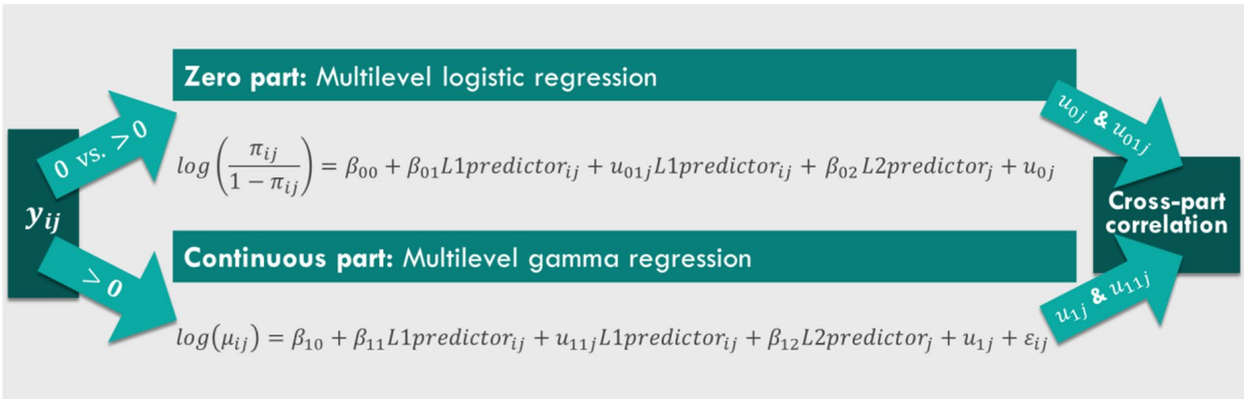


Fig. 2 Illustration of the proposed multilevel two-part model

Table 1 Overview of the most relevant parameters and their interpretation

Parameter	Description	Interpretation
Zero part		
β_{00}	overall intercept	mean of the log-odds of no eating across all participants when all predictors are equal to 0
β_{01}	fixed effect of *L1predictor*	expected change in log-odds of no eating for a one-unit increase in *L1predictor*
β_{02}	fixed effect of *L2predictor*	expected change in log-odds of no eating for a one-unit increase in *L2predictor*
u_{01j}	random effect of *L1predictor* in person *j*	person-specific differences in the effect of *L1predictor* on the log-odds of no eating
u_{0j}	random intercept of person *j*	person-specific differences in the log-odds of no eating
$\sqrt{\sigma^2{}_{u_0}}$	standard deviation of the random intercept u_0	variation of between-person differences in the log-odds of no eating
$\sqrt{\sigma^2{}_{u_{01}}}$	standard deviation of the random effect u_{01}	variation of between-person differences in the effect of *L1predictor* on the log-odds of no eating
Continuous part		
β_{10}	overall intercept	mean of the (log) amount consumed across all participants when all predictors are equal to 0 given that dietary intake occurred
β_{11}	fixed effect of *L1predictor*	expected change in the (log) amount consumed for a one-unit increase in *L1predictor*
β_{12}	fixed effect of *L2predictor*	expected change in the (log) amount consumed for a one-unit increase in *L2predictor*
u_{11j}	random effect of *L1predictor* in person *j*	person-specific differences in the effects of *L1predictor* on the (log) amount consumed
u_{1j}	random intercept of person *j*	person-specific differences in the (log) amount consumed
$\sqrt{\sigma^2{}_{u_1}}$	standard deviation of the random intercept u_1	variation of between-person differences in the expected (log) amount consumed
$\sqrt{\sigma^2{}_{u_{11}}}$	standard deviation of the random effect u_{11}	variation of between-person differences in the effect of *L1predictor* on the expected (log) amount consumed
Cross-part correlation		
$\rho_{u_0 u_1}$	correlation between the random intercepts u_0 and u_1 of the zero and continuous part	correlation between the person-specific differences in the log-odds of no eating and the person-specific difference in the (log) amount consumed

between the random effects across the two parts (e.g. $\rho_{u_0 u_1}$), often called cross-part correlation, is modelled (illustrated in Fig. 2). The number of modelled correlations is determined by the number of random effects included in the model (see Additional file 4 for details).

An overview of the most relevant parameters in the proposed multilevel two-part model is provided in Table 1.

More general overviews of (multilevel) two-part models can be found in the following literature: Neelon et al.

[34, 35] provide an overview as well as case studies on zero-modified count and semicontinuous data, marginally also covering longitudinal data. Liu et al. [36] discuss statistical analyses of semicontinuous data in the cross-sectional as well as longitudinal setting. Farewell et al. [37] provide a review on two-part and related regression models for longitudinal semicontinuous as well as longitudinal count data.

Multilevel two-part modelling in brms

brms

The R-package brms [28, 29] supports Bayesian multilevel modelling and is implemented via the probabilistic programming language Stan [38]. For readers who are new to Bayesian statistics, a brief introduction is provided in Additional file 1a. We recommend Depaoli et al. [39] as well as van de Schoot and Depaoli [40] to readers who are interested in a broader introduction to Bayesian statistics in the context of health psychology.

We chose brms for this paper for a number of reasons: Firstly and most importantly, the major advantage of brms is that it uses a lme4-like formula syntax. lme4 is one of the most commonly used R-packages for multilevel modelling which will facilitate the initial familiarization with brms for those readers who are familiar with multilevel modelling in R. Secondly, it does not require any data preprocessing (e.g. dividing the semicontinuous outcome into two variables, a dichotomous and a continuous variable) as other software programs do (e.g. gsem in Stata). Thirdly, it offers great flexibility in the model specification (see [28, 29] for details).

Multilevel two-part model in brms

The proposed multilevel two-part model combining a multilevel logistic and a multilevel gamma regression can be run in brms through the family *hurdle_gamma*.

The general syntax of the model looks as follows:

```
model_name <-
 brm(
   bf(
    semicontinuous_outcome ~ 1 + L1/L2-predictor + (1 + L1-predictor |x| ID),
    hu                     ~ 1 + L1/L2-predictor + (1 + L1-predictor |x| ID)
      ),
      data = data,
      family = hurdle_gamma()
    )
```

First of all, the name with which the fitted model will be saved in the R-Environment is specified. The *brm*-function indicates that a Bayesian generalized (non-) linear multilevel model is fitted. *bf* (short for *brmsformula*) is used for setting up the model formula. The upper part within *bf* represents the formula for the multilevel gamma regression for positive values.[3] The bottom part indicated by *hu* shows the formula for the multilevel logistic regression for zero values. The two parts of the model are specified after ~ through a formula almost identical to lme4-syntax. The initial *1* represents the overall intercepts which are followed by Level-1 and/or Level-2 predictors with fixed effects. Within parentheses, random effects of Level-1 predictors can be specified after the random intercept *1*. $|x|$ specifies random effects of the same participant to be correlated across the two parts of the model, i.e. cross-part correlations are modelled (denoted as the cross-part covariance matrix $\boldsymbol{\Sigma_{01}}$ in Additional file 4). x within | | was chosen arbitrarily and can be exchanged for any letter or digit. After $|x|$ the grouping variable is specified, in longitudinal data the variable indicating the participant ID. *data* indicates which data frame is used for the analysis. Bold parts of the syntax have to be customized.

Additional parameters can—and in some cases must—be specified within the *brm*-function to adapt the sampling algorithm (see Additional file 1a for a brief introduction to Bayesian sampling). brms runs 4 Markov chains with 2000 iterations each by default. The number of chains and iterations per chain can be customized through the arguments *chains* and *iter*. Unless otherwise specified through the argument *warmup*, half of the iterations are warm-up iterations (in the default setting: 2000/2 = 1000). If a model does not converge, brms provides a link to a website [41] with detailed information on recommended modifications (e.g. increase the number of iterations) to make the model converge. The argument *set_pior* can be used to incorporate prior information. However, due to a lack of prior information we exclusively use the default priors of brms in this paper which are very weakly informative and therefore influence the results as little as possible.

Data and material

The following research question is assessed exemplary within this paper: "Do momentary energetic arousal and gender predict the occurrence of energy intake and/or the amount of energy consumed within time-intervals in which energy intake occurred in daily life?" This question was chosen purely for illustrative purposes.

[3] Note that the upper formula predicts only positive values even though the variable *semicontinuous_outcome* contains all values of the semicontinuous outcome, including zeros.

We do not test specific a prior hypotheses with these analyses.

Data were collected within the Eat2beNICE-APPetite-study. This study captures dietary intake and related factors through EMA using the APPetite-mobile-app (details on the APPetite-mobile-app can be found in Ruf et al. [45]). Dietary intake was captured in an event-contingent fashion and used to quantify energy and nutrient intake. Momentary energetic arousal was assessed signal-contingent through 8 semi-random prompts per day. Participants used the app for three consecutive days. Hence, energetic arousal was assessed at up to 24 time points.

Each assessment of energetic arousal was matched to subsequent energy intake (in kcal). Subsequent energy intake was defined as the sum of any intake of energy within the time until the next assessment of energetic arousal or within the next 2 h if the time between two assessments was more than 2 h (e.g. because a prompt was missed) (see Fig. 3 for an illustration).

The dataset and the R code used in this paper can be found in Additional files 2 and 3. The dataset contains 2044 time points from 99 participants. 48.4% (989/2044) of time-intervals show no energy intake and are therefore equal to 0. The mean of non-zero values is 444.5 kcal. The dataset is in long-format (that is, repeated measurements for each participant are reported in separate lines of the dataset) and contains the variables shown in Table 2.

Analyses were run using version 4.0.5 of R, version 1.4.1106 of RStudio (RStudio Inc., Boston, MA, USA [42]), version 2.15 of brms and version 2.21.2 of rstan [43].

Results

Intercept only model

First of all, we specify and run an intercept only model (also called empty model or null model). As the name implies, it does not contain any predictors, only intercepts. The model syntax looks as follows:

```
m.null <-
  brm(
    bf(
      energy_intake ~ 1 + (1 |x| ID),
      hu            ~ 1 + (1 |x| ID)
    ),
    data = data,
    family = hurdle_gamma()
  )
```

When running the model, the following code appears progressively in the console:

```
Compiling Stan program...
Start sampling
Loading required namespace: rstudioapi
Warning message:
In system(paste(CXX, ARGS), ignore.stdout = TRUE, ignore.stderr = TRUE) :
  'C:/rtools40/usr/mingw_/bin/g++' not found
```

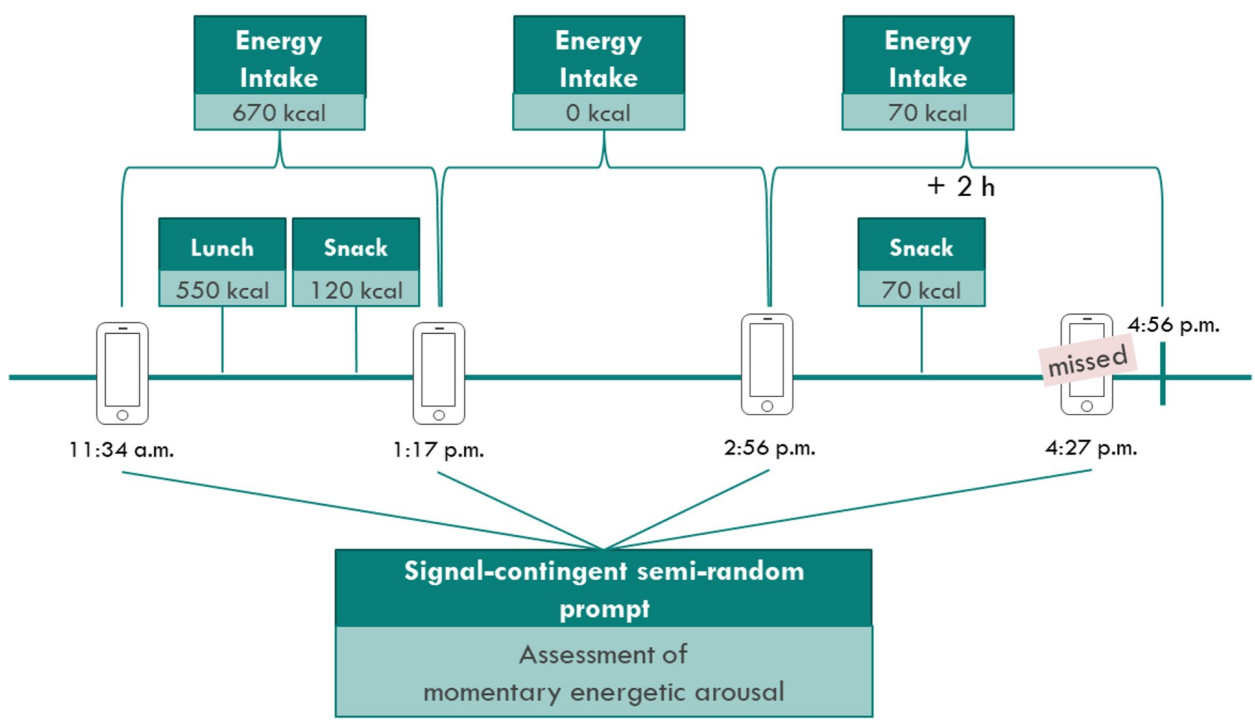


Fig. 3 Illustration of the data pre-processing procedure

Table 2 Variable overview

Name	Description	Coding
ID	subject ID	1, 2, 3, …, 99
alarm	number of prompt (maximum = 24)	1, 2, 3, …, 24
day	day 1 to 3	1, 2, 3
time	numeric time of random alarm	e.g. 8.5 for 8:30 a.m
energy_intake	energy intake in kcal	
gender	participants' gender	0 = male, 1 = female
EA	momentary energetic arousal, person-mean-centered	

First, it shows that Stan is being compiled. A few moments later, sampling is started and the viewer opens. By refreshing the viewer, the progress of the sampling can be monitored. When the model is fitted, a warning is printed. However, this warning can be ignored as it does not affect the model estimation and will be removed in the next release of rstan [44]. As we do not get any other warnings, the model seems to have converged. However, to reassure the quality of the parameter estimates, additional information regarding the construction of the posterior distribution should be obtained. To check convergence, we have a look at density and trace plots of the parameter estimates. These plots can be produced by running the command *plot(m.null)* and are shown in Fig. 4.

Density plots of model parameters should be clearly unimodal which seems to be the case in this model. Trace plots show each sampled parameter estimate from the first to the 1000th iteration of each of the four chains after warm-up. The estimates should circle around a single value to indicate convergence. The trace plots in Fig. 4 indicate convergence as the estimates hover around a single value. If the density and trace plots suggest that the model has not converged, the model should be run with more iterations. The potential scale reduction factor evaluates convergence through assessing differences between the chains (between-chain variance/within-chain variance) and should be close to 1. It is given for each parameter in the brms output in the column *Rhat* and is close to 1 if no warning is displayed. As the plots do not show any signs of nonconvergence and no relevant warnings are displayed, we can have a look at the model estimation. To do so, we run the command *summary(m.null)* and get the following results[4]:

```
Family: hurdle_gamma
  Links: mu = log; shape = identity; hu = logit
Formula: energy_intake ~ 1 + (1 | x | ID)
         hu ~ 1 + (1 | x | ID)
   Data: data (Number of observations: 2044)
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup samples = 4000

Group-Level Effects:
~ID (Number of levels: 99)
                              Estimate  Est.Error l-95% CI  u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)                     0.14 (√σ^2_u1)  0.05      0.04      0.24 1.00     1236     1141
sd(hu_Intercept)                  0.23 (√σ^2_u0)  0.07      0.06      0.37 1.00     1024      799
cor(Intercept,hu_Intercept)       0.77 (ρ_u1u0)   0.25      0.07      0.99 1.00     1140     1161

Population-Level Effects:
              Estimate  Est.Error l-95% CI u-95% CI  Rhat Bulk_ESS Tail_ESS
Intercept         6.10 (β10)   0.04     6.03     6.17  1.00     3965     2843
hu_Intercept     -0.06 (β00)   0.05    -0.16     0.04  1.00     3317     3222

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
shape     0.94      0.04     0.87     1.01 1.00     5631     2680

Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

First of all, we double-check that the values in the column *Rhat* are close to 1. All Rhats are equal to 1.00 in this model, so the parameter estimates can be deemed trustworthy.

In the section *Population-Level Effects* which comprises fixed effects, we get two intercepts, one for the gamma part of the model β_{10} (= *Intercept*) and one for the logistic part β_{00} (= *hu_Intercept*). In brms, point estimates of parameters represent the mean of the respective posterior distribution. Estimates of the gamma part are modelled on the log scale as the outcome can only be positive. Hence, to obtain the estimate of the intercept in the original metric (kcal), we calculate the exponential of β_{10} (*exp(6.1)* = 445.9). This indicates that in time-intervals in which energy intake occurred we expect an individual to consume on average 445.9 kcal. This value should be close to the mean of non-zero values in the original data as the group mean is the best estimate in models without predictors. In our data the mean of positive values is 444.5 which is very close to the model estimate.

Estimates of the logistic part are modelled on the logit scale which accommodates the restricted range of probabilities (between 0 and 1). The intercept β_{00} represents the average log-odds of no energy intake across all participants. To transform the log-odds to the probability of no energy intake, we can use the inverse logit function in Eq. (1) or alternatively the *plogis*-function in R.

$$\pi = \frac{\exp(\beta_{00})}{1 + \exp(\beta_{00})} = \frac{\exp(-0.06)}{1 + \exp(-0.06)} = plogis(-0.06) = 0.485 \quad (1)$$

We get a mean probability of no energy intake of 0.485 (= 48.5%). We can check whether this estimate is reasonable through looking at the percentage of time-intervals without energy intake within the original data. In 48.4% (989/2044) of time-intervals energy intake is equal to zero which is close to the estimate of the intercept. We

[4] Note that the notations marked in blue are inserted by us for illustrative reasons.

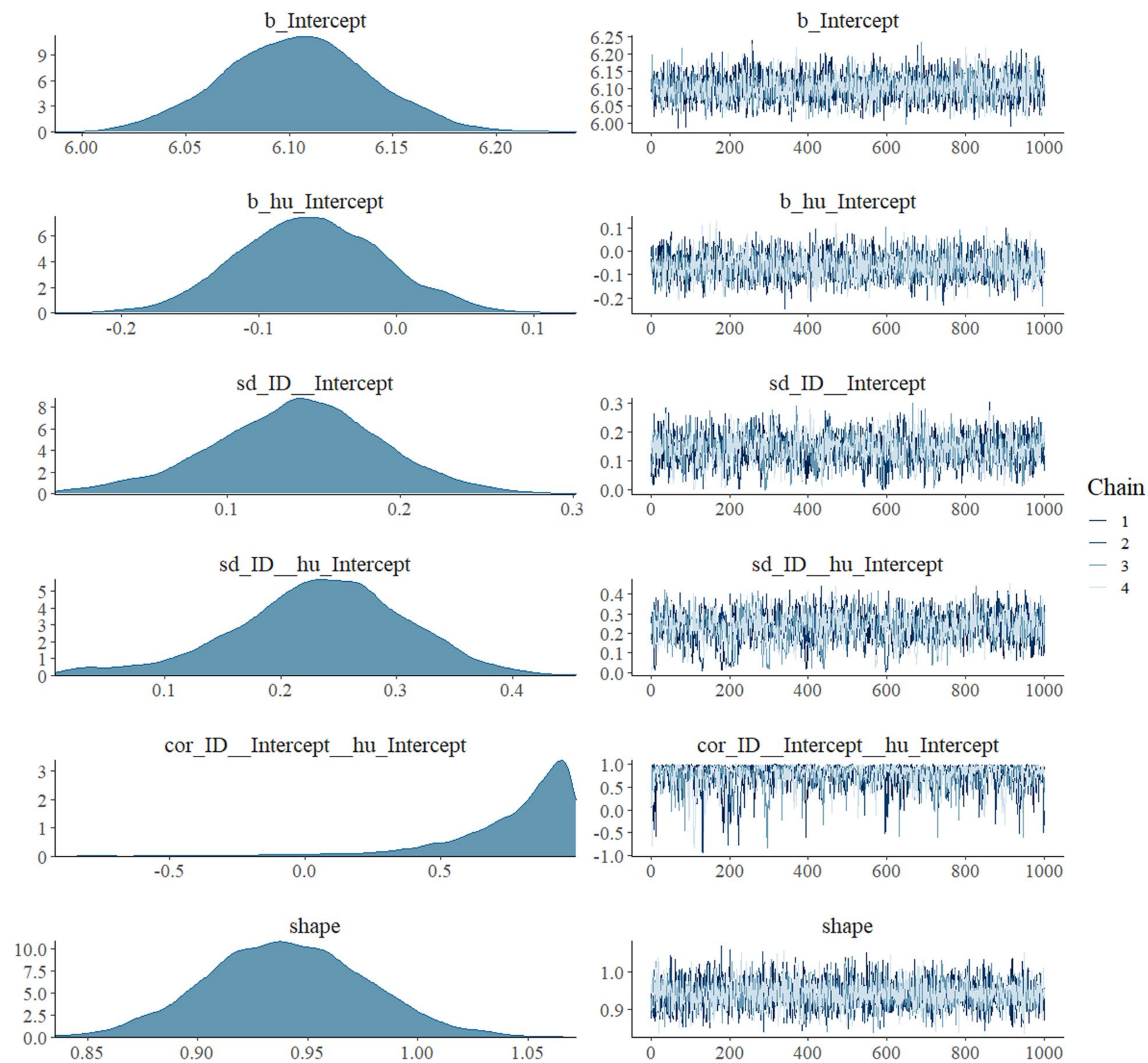


Fig. 4 Density and trace plots of each model parameter of the intercept only model

recommend always checking the model implied estimates from the intercept only model against the descriptive sample estimates to ascertain that the model was specified correctly and that the sample estimates could be reproduced with reasonable accuracy.

Now we have a look at the random effects (*Group-Level Effects* in brms). *Sd(Intercept)* ($\sqrt{\sigma^2{}_{u_1}}$) represents the variation of the random intercept u_1 of the gamma part, i.e. person-specific variation in the mean of (log) energy intake. Mean (log) energy intake varies between participants with a *SD* of 0.14 (95%-credible interval [CI] 0.04–0.24). As effects are assumed to be multivariate normally distributed, we can calculate the range in which the mean energy intake of 95% of participants is located (Intercept ± 1.96 *SDs*). The mean energy intake of 95% of participants is between 338.86 (*exp(6.1–0.14*1.96)*) and 586.63 kcal (*exp(6.1 + 0.14*1.96)*) in time-intervals in which energy intake occurred. Furthermore, participants differ in the mean log-odds of no energy intake with a *SD* of 0.23 (95%-CI 0.06–0.37) shown by the variation of the random intercept u_0 of the logistic part *sd(hu_Intercept)* ($\sqrt{\sigma^2{}_{u_0}}$). For 95% of participants the probability of no energy intake is between 0.38 (*plogis(-0.06–0.23*1.96)*) and 0.6 (*plogis(-0.06 + 0.23*1.96)*).

The fairly strong positive cross-part correlation between the random intercepts ($\rho_{u_0 u_1}$) of 0.77 indicates that participants who consume on average more energy within time-intervals in which energy intake occurs have on average a higher probability of no energy intake.

Random intercept model with Level-2 predictor

Now we want to include a fixed effect of the Level-2 predictor gender in both parts of the model by running the following code:

```
brm(
  bf(
    energy_intake ~ 1 + gender + (1 |x| ID),
    hu            ~ 1 + gender + (1 |x| ID)
  ),
  data = data,
  family = hurdle_gamma()
)
```

We do not get any warnings regarding nonconvergence and the density and trace plots do not indicate convergence problems, therefore we can interpret the model estimates[5]:

Group-Level Effects:
~ID (Number of levels: 99)

	Estimate		Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sd(Intercept)	0.12	$(\sqrt{\sigma^2{}_{u_1}})$	0.05	0.01	0.22	1.00	795	599
sd(hu_Intercept)	0.23	$(\sqrt{\sigma^2{}_{u_0}})$	0.08	0.04	0.38	1.00	831	507
cor(Intercept,hu_Intercept)	0.71	$(\rho_{u_1 u_0})$	0.34	-0.34	0.99	1.00	704	627

Population-Level Effects:

	Estimate		Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	6.27	(β_{10})	0.06	6.14	6.39	1.00	4889	3135
hu_Intercept	-0.05	(β_{00})	0.09	-0.23	0.13	1.00	4074	2942
gender	-0.25	(β_{11})	0.07	-0.39	-0.10	1.00	4855	3059
hu_gender	-0.02	(β_{01})	0.11	-0.24	0.20	1.00	4241	2749

In this model the intercept β_{10} of the gamma part of the model represents the mean log energy intake for men (*gender* = 0). Male participants consume on average 528.48 kcal ($exp(\beta_{10}) = exp(6.27)$) in time-intervals in which energy intake occurred. Results show that gender has a fixed effect on the mean log energy intake in time-intervals in which energy intake occurs as the 95%-CI of β_{11} does not include 0. To interpret the regression coefficient of the fixed effect of gender, we can get the rate decrease in energy intake associated with a one-unit increase in *gender* through exponentiation of β_{11}. Hence, women (*gender* = 1) consume on average around 22% less energy ($exp(\beta_{11}) = exp(-0.25) = 0.78$) in time-intervals in which energy intake occurred compared to men. However, women and men do not differ in the probability of no energy intake as the fixed effect of gender is not relevant for the prediction in the logistic part of the model (95%-CI of β_{01} includes 0).

[5] To keep this paper short, we do not show any further density and trace plots and only show the relevant parts of the brms output (Population-Level and Group-Level effects). However, density and trace plots can be found in Additional file 5 and complete model summaries in Additional file 6.

We get three estimates within the random effects. There is between-person variation in the log energy intake ($\sqrt{\sigma^2{}_{u_1}}$=0.12, 95%-CI 0.01–0.22) in time-intervals in which energy intake occurred as well as in the log-odds of no energy intake ($\sqrt{\sigma^2{}_{u_0}}$=0.23, 95%-CI 0.04–0.38). The cross-part correlation $\rho_{u_0 u_1}$ is still fairly strong (0.71), suggesting that participants who consume on average more energy in time-intervals in which energy intake occurs have on average a higher probability of no energy intake. However, what we see here is that the 95%-CI of $\rho_{u_0 u_1}$ includes 0 and is wider compared to the intercept only model indicating that the estimation is rather inaccurate (95%-CI -0.34–0.99).

Random slope model with Level-1 predictor

Now we want to include the Level-1 predictor momentary energetic arousal in both parts of the model as fixed and random effects. As there was no fixed effect of gender in the logistic part of the model, we only include gender in the gamma part. To do so, we fit the following model:

```
brm(
  bf(
   energy_intake ~ 1 + gender + EA + (1 + EA |x| ID),
   hu            ~ 1 +          EA + (1 + EA |x| ID)
  ),
  data = data,
  family = hurdle_gamma()
)
```

We do not get any warnings regarding nonconvergence and the density and trace plots do not indicate serious convergence problems, therefore we can interpret the model estimates:

Group-Level Effects:
~ID (Number of levels: 99)

	Estimate		Est.Error	l-95% CI	u-95% CI	Rhat	Bulk ESS	Tail ESS
sd(Intercept)	0.11	$(\sqrt{\sigma^2{}_{u_1}})$	0.05	0.01	0.21	1.00	801	611
sd(EA)	0.02	$(\sqrt{\sigma^2{}_{u_{12}}})$	0.02	0.00	0.06	1.00	2144	1789
sd(hu_Intercept)	0.22	$(\sqrt{\sigma^2{}_{u_0}})$	0.08	0.04	0.37	1.00	1098	950
sd(hu_EA)	0.11	$(\sqrt{\sigma^2{}_{u_{01}}})$	0.04	0.02	0.18	1.00	1154	1608
cor(Intercept,EA)	0.00	$(\rho_{u_1 u_{12}})$	0.45	-0.80	0.81	1.00	4486	2421
cor(Intercept,hu_Intercept)	0.56	$(\rho_{u_1 u_0})$	0.34	-0.36	0.95	1.00	843	899
cor(EA,hu_Intercept)	0.04	$(\rho_{u_{12} u_0})$	0.43	-0.78	0.80	1.00	1744	2524
cor(Intercept,hu_EA)	0.36	$(\rho_{u_1 u_{01}})$	0.37	-0.51	0.89	1.00	1177	1124
cor(EA,hu_EA)	-0.05	$(\rho_{u_{12} u_{01}})$	0.44	-0.81	0.78	1.00	1463	2423
cor(hu_Intercept,hu_EA)	0.24	$(\rho_{u_0 u_{01}})$	0.34	-0.50	0.82	1.00	2347	2418

Population-Level Effects:

	Estimate		Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	6.24	(β_{10})	0.06	6.12	6.36	1.00	4883	3222
hu_Intercept	-0.06	(β_{00})	0.05	-0.16	0.04	1.00	4046	2537
gender	-0.22	(β_{11})	0.07	-0.37	-0.08	1.00	4479	2810
EA	0.02	(β_{12})	0.02	-0.01	0.05	1.00	5552	3172
hu_EA	-0.04	(β_{01})	0.03	-0.09	0.01	1.00	5232	3301

Again we see the meaningful fixed effect of gender in the gamma part (β_{11}). However, there is no fixed effect of energetic arousal in either of the two parts (95%-CI include 0). That is, there is no evidence that participants were more likely not to eat when their energetic arousal was higher than usual, β_{01} = -0.04 (95%-CI -0.09–0.01). There was also no evidence that participants consumed more energy when their energetic arousal was higher than usual, β_{12} = 0.02 (95%-CI -0.01–0.05). Notice, however, that the random effect for energetic arousal in the logistic part suggests that the effect of energetic arousal on the log-odds of no energy intake varies across participants with a *SD* of 0.11 (95%-CI 0.02–0.18). Hence, for 95% of participants the effect of energetic arousal on the log-odds of no energy intake is between -0.26 (-0.04–0.11*1.96) and 0.18 (-0.04 + 0.11*1.96). This suggests that on average there is no association of energetic arousal with the probability not to eat. However, for some participants, higher arousal may be associated with a higher probability not to eat. Whereas for others, higher arousal may be associated with a lower probability not to eat. The random effect for energetic arousal in the gamma part was smaller and the lower bound of the 95%-CI was 0.00. Note that non-positive estimates for *SD* are not permitted, and the lower bound of the CI for this parameter will therefore always be positive. This suggests that inter-individual differences in the effect of energetic arousal on the amount of energy intake are small and possibly not statistically meaningful.

We get ten estimates within the random effects: 4 *SD*s and 6 correlations (as shown in expression (6) in Additional file 4). We see that the cross-part correlation $\rho_{u_0 u_1}$ between the random intercepts is weaker than in the previous models (0.56) and that the 95%-CI of all correlations is very wide indicating that it is not possible to get accurate estimates (see also the platykurtic posterior distributions in Additional file 5).

Discussion

Studying dietary intake through multilevel two-part modelling is a methodologically as well as conceptually promising approach. It accounts for the semicontinuous data structure and offers novel and distinct insights in terms of the occurrence as well as the amount of dietary intake. Results of this paper highlight that the differentiation between the two processes reveals process-specific associations which cannot be detected through traditional multilevel modelling. For instance, we found that gender is associated with the amount consumed during eating occasions, but not with the probability of eating. The model we propose overcomes a number of limitations of traditional modelling when analysing semicontinuous data: (1) accounts for the zero-inflation by introducing two model parts, a zero and a continuous part, which avoids incorrect inferences (as shown by Baldwin et al. [8]), (2) accommodates the skewness of the continuous part of the outcome by applying a gamma regression which does not rely on controversial transformation of the outcome and does not change the metric of the data, and (3) considers the dependency between the two model parts by modelling the cross-part correlation which prevents bias in parameter estimation as would running separate models (as outlined below). Despite its potential, multilevel two-part modelling is still missing in the statistical repertoire of most researchers. This may be due to the fact that these models are rather complex and therefore require initial training. However, we believe that multilevel two-part models are the most appropriate and valid method to study semicontinuous outcomes and therefore are worth the training. To facilitate the initial training and encourage other researchers to use these models, this paper offers an application-oriented introduction to multilevel two-part modelling.

The R-package brms used in this paper offers a user-friendly and freely available option for fitting multilevel two-part models. It is particularly intuitive for users familiar with lme4 and Bayesian statistics (see Additional file 1b for a brief overview of similarities and differences between Bayesian and frequentist-based two-part models). We believe that multilevel two-part models are of particular interest to those researchers who are familiar with traditional multilevel modelling.

For demonstration and simplicity purposes, we have focused on multilevel two-part models with fixed and random effects. However, extensions to the model (e.g. cross-level-interactions) are straightforward.

We found fairly strong to moderate positive cross-part correlations (0.77, 0.71, 0.56) indicating that participants who consume on average more energy during eating occasions have on average a higher probability not to eat. However, we have faced some estimation inaccuracies of the cross-part correlations: the more predictors we included in the model, the wider the 95%-CIs got. Nonetheless, we do not recommend fitting separate models as ignoring the cross-part correlation can induce bias in regression coefficients as well as variance components [10, 11]. Not accounting for the cross-part correlation can cause bias particularly in the continuous part of the model. This can be explained by the fact that the zero part determines the cluster size of the continuous part of the model (e.g. the number of observations with dietary intake within an individual). For instance, we found moderate to strong cross-part correlations. Hence, an individual less likely to eat will have fewer observations in the continuous part of the model but the few observations will contain larger amounts. An individual eating more frequently will have more observations in the continuous part which will contain smaller amounts. As a result, higher values of dietary intake will be underrepresented and smaller values will be overrepresented. Su et al. [10] outline that even when researchers are only interested in the continuous part of the semicontinuous outcome and therefore chose to fit a single model, the described bias will still be present.

To run the proposed multilevel two-part model, data on dietary intake as well as individual and/or situational factors have to be collected. Dietary intake can either be captured event- (i.e. when food is consumed [45]), signal- (e.g. time since the last prompt) or time-contingently (e.g. within the last hour [1]), while individual and/or situational factors have to be assessed either signal- or time-contingently. The proposed model cannot be applied to simple event-contingent sampling protocols (e.g. dietary intake and factors of interest assessed only when food is consumed).

While first empirical evidence [1] as well as results of this paper support the importance of distinguishing between the occurrence of eating and the amount that is eaten, future research is needed to verify the conceptual relevance of studying dietary intake as a dual process. We believe that multilevel two-part models will contribute to a better understanding of which situational and individual factors are associated with an increased probability of eating and/or with an increased amount of dietary intake. Findings in this area offer new perspectives and enable the development of tailored interventional strategies. For instance, in the context of preventing and treating overweight and obesity two types of interventions are needed: (1) interventions customized to reduce the probability of dietary intake and therefore reduce the number of eating occasions within a day, (2) interventions tailored to reduce the amount eaten within eating occasions to prevent overeating.

In this paper, we applied multilevel two-part modelling to study factors influencing energy intake. However, multilevel two-part models can also be employed to study macro-nutrient intakes which are also semicontinuous in the Eat2beNICE-APPetite data. Furthermore, multilevel two-part modelling can also be applied to studies which capture food categories (e.g. vegetable intake), provided that the consumed amounts are also assessed. Findings in the context of macro-nutrient intake and food categories can be translated to the promotion of healthy eating, e.g. reducing the occurrence of sugar intake or boosting vegetable consumption within eating occasions. Hence, there are numerous ways multilevel two-part modelling can be applied in the context of studying dietary intake in daily life.

Beyond that, the model proposed in this paper can also be applied to other research contexts in which a semicontinuous outcome is present, including PA data in which zeros are a common problem [46] (e.g. daily PA data [8] or PA data in EMA studies [25]). In fact, almost all behavioural outcomes are likely to show semicontinuous characteristics which can be traced back to dual processes: one process determining whether the behaviour is shown and the other determining how long/intensive/often the behaviour is shown, e.g. smoking behaviour (Has an individual smoked? If so, how many cigarettes have been smoked?), social interaction (Has an individual engaged in social interaction? If so, how many minutes did the individual interact socially?) and purchase behaviour (Has an individual purchased anything? If so, how much money was spent?)—to name only a few. The shorter time-intervals are in which a specific behaviour is studied (e.g. daily diary and EMA studies), the more likely it is that the outcome is zero-inflated, i.e. the behaviour of interest is not shown. Therefore, as the number of these studies is continuously growing, so will the need for multilevel two-part modelling to study predictors of specific behaviours. This paper addresses this need by providing guidance on the implementation and interpretation of these rather complex models.

Conclusions

To the best of our knowledge, this paper is the first to introduce multilevel two-part modelling as a novel analytical approach to study dietary intake in daily life. Distinguishing between factors influencing whether and how much is eaten is conceptually promising and offers new opportunities, particularly for customized nutritional interventions either targeting the occurrence of intake or the amount consumed during eating occasions. As we believe that the importance of EMA studies assessing factors influencing dietary intake in daily life is growing within the next years, this paper will help to establish an appropriate data analysis procedure that accounts for the dual character of dietary intake and the semicontinuous data structure.

Abbreviations

EMA: Ecological momentary assessment; kcal: Kilocalories; PA: Physical activity; *SD*: Standard deviation; CI: Credible interval.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12966-021-01187-8.

Additional file 1. Bayesian Statistics. **a.** Brief introduction to Bayesian statistics **b.** Similarities and differences between Bayesian and frequentist-based two-part models.

Additional file 2. Dataset.

Additional file 3. R Code.

Additional file 4. Model specifications of the proposed multilevel two-part model.

Additional file 5: Density and Trace Plots. **Figure 1.** Density and Trace Plots of the Random Intercept Model with Level-2 predictor *gender*. **Figure 2.** Density and Trace Plots of the Random Slope Model with Level-1 predictor energetic arousal (*EA*).

Additional file 6: Model Summaries. **Figure 1.** Model summary of the Random Intercept Model with Level-2 predictor *gender*. **Figure 2.** Model summary of the Random Slope Model with Level-1 predictor energetic arousal (*EA*).

Acknowledgements
Not applicable.

Authors' contributions
A Ruf, ABN and SM devised the analysis plan. A Ruf conducted the analyses and wrote the first draft of the manuscript. A Ruf and ABN contributed to data interpretation. A Ruf, ABN, UE-P, A Reif and SM critically reviewed, edited and approved the final manuscript.

Funding
This work was supported by the European Union's Horizon 2020 Research and Innovation Program under grant agreement No 728018. The funding source has had no involvement in the study design, data collection, interpretation of the findings, or writing of this manuscript. Open Access funding enabled and organized by Projekt DEAL

Availability of data and materials
All data generated or analysed during this study are included in this published article and its supplementary information files.

Declarations

Ethics approval and consent to participate
The local ethics committee of the faculty of medicine of the Goethe University Frankfurt (Ethikkommission des Fachbereichs Medizin der Goethe-Universität) approved the study (reference number: 192/18). All subjects declared that they understood the study procedure and signed a written informed consent.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital, Goethe University, Heinrich-Hoffmann-Straße 10, 60528 Frankfurt am Main, Germany. [2]DIPF | Leibniz Institute for Research and Information in Education, Frankfurt am Main, Germany. [3]Center for Research on Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt am Main, Germany. [4]Mental mHealth Lab, Institute of Sports and Sports Science, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany. [5]Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany.

Received: 17 May 2021 Accepted: 10 August 2021
Published online: 27 September 2021

References

1. Powell DJH, McMinn D, Allan JL. Does real time variability in inhibitory control drive snacking behavior? An intensive longitudinal study. Health Psychol. 2017;36:356–64.
2. Dunton GF. Sustaining health-protective behaviors such as physical activity and healthy eating. JAMA. 2018;320:639–40.
3. Hofmann W, Dohle S. Capturing eating behavior where the action takes place: a comment on McKee et al. Ann Behav Med. 2014;48:289–90.
4. Schembre SM, Liao Y, O'Connor SG, Hingle MD, Shen SE, Hamoy KG, et al. Mobile ecological momentary diet assessment methods for behavioral research: systematic review. JMIR mHealth uHealth. 2018;6:e11170.
5. Maugeri A, Barchitta M. A systematic review of ecological momentary assessment of diet: implications and perspectives for nutritional epidemiology. Nutrients. 2019;11:2696.
6. Hummel E, Hoffmann I. Complexity of nutritional behavior: capturing and depicting its interrelated factors in a cause-effect model. Ecol Food Nutr. 2016;55:241–57.
7. Maher JP, Harduk M, Hevel DJ, Adams WM, McGuirt JT. Momentary physical activity co-occurs with healthy and unhealthy dietary intake in African American College freshmen. Nutrients. 2020;12:1360.
8. Baldwin SA, Fellingham GW, Baldwin AS. Statistical models for multilevel skewed physical activity data in health research and behavioral medicine. Health Psychol. 2016;35:552–62.
9. Tooze JA, Grunwald GK, Jones RH. Analysis of repeated measures data with clumping at zero. Stat Methods Med Res. 2002;11:341–55.
10. Su L, Tom BDM, Farewell VT. Bias in 2-part mixed models for longitudinal semicontinuous data. Biostatistics. 2009;10:374–89.
11. Liu L, Ma JZ, Johnson BA. A multi-level two-part random effects model, with application to an alcohol-dependence study. Stat Med. 2008;27:3528–39.
12. Elliston KG, Ferguson SG, Schüz B. Personal and situational predictors of everyday snacking: an application of temporal self-regulation theory. Br J Health Psychol. 2017;22:854–71.
13. McNaughton SA, Pendergast FJ, Worsley A, Leech RM. Eating occasion situational factors and sugar-sweetened beverage consumption in young adults. Int J Behav Nutr Phys Act. 2020;17:71.
14. Ghosh Roy P, Jones KK, Martyn-Nemeth P, Zenk SN. Contextual correlates of energy-dense snack food and sweetened beverage intake across the day in African American women: an application of ecological momentary assessment. Appetite. 2019;132:73–81.
15. Olsen MK, Schafer JL. A two-part random-effects model for semicontinuous longitudinal data. J Am Stat Assoc. 2001;96:730–45.
16. Liu L, Strawderman RL, Cowen ME, Shih YCT. A flexible two-part random effects model for correlated medical costs. J Health Econ. 2010;29:110–23.
17. Liu L. Joint modeling longitudinal semi-continuous data and survival, with application to longitudinal medical cost data. Stat Med. 2009;28:972–86.
18. Atkins DC, Baldwin SA, Zheng C, Gallop RJ, Neighbors C. A tutorial on count regression and zero-altered count models for longitudinal substance use data. Psychol Addict Behav. 2013;27:166–77.
19. Bandyopadhyay D, DeSantis SM, Korte JE, Brady KT. Some considerations for excess zeroes in substance abuse research. Am J Drug Alcohol Abuse. 2011;37:376–82.
20. DeSantis SM, Bandyopadhyay D, Baker NL, Randall PK, Anton RF, Prisciandaro JJ. Modeling longitudinal drinking data in clinical trials: an application to the COMBINE study. Drug Alcohol Depend. 2013;132:244–50.
21. Liu L, Strawderman RL, Johnson BA, O'Quigley JM. Analyzing repeated measures semi-continuous data, with application to an alcohol dependence study. Stat Methods Med Res. 2012;25:133–52.
22. Xing D, Huang Y, Chen H, Zhu Y, Dagne GA, Baldwin J. Bayesian inference for two-part mixed-effects model using skew distributions, with application to longitudinal semicontinuous alcohol data. Stat Methods Med Res. 2015;26:1838–53.
23. Lee AH, Zhao Y, Yau KKW, Xiang L. How to analyze longitudinal multilevel physical activity data with many zeros? Prev Med. 2010;51:476–81.
24. Lee AH, Xiang L. Mixture analysis of heterogeneous physical activity outcomes. Ann Epidemiol. 2011;21:780–6.
25. Zink J, Belcher BR, Dzubur E, Ke W, O'Connor S, Huh J, et al. Association between self-reported and objective activity levels by demographic factors: ecological momentary assessment study in children. JMIR mHealth uHealth. 2018;6:e150.
26. Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, et al. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. J Am Diet Assoc. 2006;106:1575–87.
27. R Core Team. R: a language and environment for statistical computing. Vienna; 2020. Available from: https://www.r-project.org/.
28. Bürkner P-C. brms: an R package for Bayesian multilevel models using Stan. J Stat Softw. 2017;80:1–28.
29. Bürkner P-C. Advanced Bayesian multilevel modeling with the R package brms. R J. 2018;10:395–411.

30. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. nlme: linear and nonlinear mixed effects models. 2020. Available from: https://cran.r-project.org/package=nlme.
31. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw. 2015;67:1–46.
32. Hox JJ, Moerbeek M, van de Schoot R. Multilevel analysis: techniques and applications. 3rd ed. New York: Routledge; 2017.
33. Snijders TAB, Bosker RJ. Multilevel analysis: an introduction to basic and advanced multilevel modeling. 2nd ed. London: SAGE Publications Ltd; 2011.
34. Neelon B, O'Malley AJ, Smith VA. Modeling zero-modified count and semicontinuous data in health services research part 1: background and overview. Stat Med. 2016;35:5070–93.
35. Neelon B, O'Malley AJ, Smith VA. Modeling zero-modified count and semicontinuous data in health services research part 2: case studies. Stat Med. 2016;35:5094–112.
36. Liu L, Shih YCT, Strawderman RL, Zhang D, Johnson BA, Chai H. Statistical analysis of zero-inflated nonnegative continuous data: a review. Stat Sci. 2019;34:253–79.
37. Farewell VT, Long DL, Tom BDM, Yiu S, Su L. Two-part and related regression models for longitudinal data. Annu Rev Stat Appl. 2017;4:283–315.
38. Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, et al. Stan: a probabilistic programming language. J Stat Softw. 2017;76:1–32.
39. Depaoli S, Rus HM, Clifton JP, van de Schoot R, Tiemensma J. An introduction to Bayesian statistics in health psychology. Health Psychol Rev. 2017;11:248–64.
40. van de Schoot R, Depaoli S. Bayesian analyses: where to start and what to report. Eur Health Psychol. 2014;16:75–84.
41. Brief guide to Stan's warnings. http://mc-stan.org/misc/warnings.html. Accessed 17 Apr 2021.
42. RStudio Team. RStudio: integrated development environment for R. Boston; 2020. Available from: http://www.rstudio.com/.
43. Stan Development Team. RStan: the R interface to Stan. 2020. Available from: http://mc-stan.org/.
44. RStan getting started. https://github.com/stan-dev/rstan/wiki/RStan-Getting-Started. Accessed 17 Apr 2021.
45. Ruf A, Koch ED, Ebner-Priemer U, Knopf M, Reif A, Matura S. Studying microtemporal, within-person processes of diet, physical activity, and related factors using the APPetite-Mobile-App: feasibility, usability, and validation study. J Med Internet Res. 2021;23:e25850.
46. Rasmussen CL, Palarea-Albaladejo J, Johansson MS, Crowley P, Stevens ML, Gupta N, et al. Zero problems with compositional data of physical behaviors: a comparison of three zero replacement methods. Int J Behav Nutr Phys Act. 2020;17:126.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Additional file 1. Bayesian statistics

a. Brief introduction to Bayesian statistics

Contrary to frequentist methods in which population parameter estimates are based on the sample data only, Bayesian statistics allow accounting for prior information. Based on the sample data as well as provided information on the prior distribution, a posterior distribution of each parameter is computed. However, the Bayesian approach can be used even if there is no reliable prior information available. In this case, instead of setting informative priors, noninformative priors are set which impact the posterior distribution as little as possible. Information from the posterior distribution can be summarized in the form of a point estimate of the respective parameter (the mean, median or mode of the posterior distribution) as well as a 95% credible interval (CI; 2.5 and 97.5 percentiles). The CI can be used to assess whether a regression coefficient is likely to be non-zero and hence relevant for the prediction of the outcome. If the CI does not include zero, it is reasonable to assume that the regression coefficient is different from zero (i.e. statistically significant). Note, however, that any type of CI could be used to determine whether or not the parameter is different from zero (this is similar to frequentist approaches in which the $p < 0.05$ criterion is ultimately arbitrary as well).

In contrast to frequentist methods, Bayesian estimation of the model proposed in this paper relies on Markov Chain Monte Carlo (MCMC) sampling, an approach that is based on simulation methods (see [1] for an overview on MCMC). It is necessary to understand the basic concept of MCMC to evaluate whether the model estimation has worked and the parameter estimates are reliable. MCMC combines the prior distribution and the information from the actual data through an iterative process obtaining a posterior distribution. Within this process, parameter values are sampled and used to update the posterior distribution. This procedure of sampling and updating is repeated many times (specified through the number of iterations) and for multiple runs (specified through the number of chains). The first iterations of each chain are discarded and not used for inference, in order to reduce the influence of the starting values. These initial discarded iterations are often referred to the warm-up. To check whether the estimation process for each parameter estimate has converged, density and trace plots can be inspected (presented within the results). Another important criterion is the potential scale reduction factor which evaluates convergence through assessing differences between the chains. It is calculated for each parameter (between-chain variance/within-chain variance) and should be close to 1 (see [2–4] for details).

b. Similarities and differences between Bayesian and frequentist-based two-part models

The model we propose is based on Bayesian inference which raises the question how similar or different Bayesian two-part models are compared to frequentist-based models. Bayesian multilevel two-part models with non-informative priors, as used in this paper, are generally expected to yield similar estimates as maximum likelihood based approaches in the traditional frequentist framework [5]. However, it remains unclear whether or in which cases estimates of the Bayesian approach proposed in this paper are comparable to likelihood based estimates. Systematic simulation studies are needed to answer this question conclusively. Nonetheless, there are two distinct advantages of the Bayesian approach over the frequentist approach which should be taken into consideration: (1) Confidence intervals in the frequentist approach are often based on normality assumptions and are therefore defined as symmetrical (point estimate +/- 1.96 SE). As a result confidence intervals of random effects in the frequentist approach can include negative values. However, random effects are expressed by measures of variability (variance or SD) which cannot be negative. The Bayesian approach avoids this by incorporating prior distributions with only non-negative values which prevent negative estimates for random effect variances in the posterior distribution. (2) Maximum likelihood estimation can cause computational challenges and convergence difficulties, particularly when models include complex random effect structures [5]. In general, Bayesian approaches have been shown to be more computational efficient [5,6] and are therefore suited particularly well for more complex data [5]. Nevertheless, future research is needed to examine if and when Bayesian approaches are more efficient than frequentist approaches for the two-part models used in this work.

1. van Ravenzwaaij D, Cassey P, Brown SD. A simple introduction to Markov Chain Monte–Carlo sampling. Psychon Bull Rev. 2018;25:143–54.

2. Gelman A, Rubin DB. Inference from Iterative Simulation Using Multiple Sequences. Stat Sci. 1992;7:457–72.

3. Brooks SP, Gelman A. General Methods for Monitoring Convergence of Iterative Simulations. J Comput Graph Stat. 1998;7:434–55.

4. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis. 3rd ed. Chapman and Hall/CRC; 2013.

5. Smith VA, Neelon B, Preisser JS, Maciejewski ML. A marginalized two-part model for longitudinal semicontinuous data. Stat Methods Med Res. 2015;26:1949–68.

6. Ghosh P, Albert PS. A Bayesian analysis for longitudinal semicontinuous data with an application to an acupuncture clinical trial. Comput Stat Data Anal. 2009;53:699–706.

Additional file 2. Dataset.

Available at https://doi.org/10.1186/s12966-021-01187-8

Additional file 3. R Code.

Available at https://doi.org/10.1186/s12966-021-01187-8

Additional file 4. Model specifications of the proposed multilevel two-part model

In the following we use the notation from Baldwin et al. [1] to introduce the model specifications. However, we extend the model by introducing random effects in both parts of the model.

The variable y_{ij} represents the semicontinuous dietary intake response from subject j (j = 1, …, m) at time point i (i = 1, …, n_i). We are interested in two parts of this variable: (1) Did the participant eat? In other words, is $y_{ij}=0$ or $y_{ij}>0$? (2) If the participant ate, how much was eaten? In other words, what is the expected value of y_{ij}, if $y_{ij}>0$? To approach these questions, the semicontinuous variable y_{ij} is split into two parts:

$$y_{ij} \sim \begin{cases} \pi_{ij} & if\ y_{ij} = 0 \\ (1-\ \pi_{ij})h(y_{ij}) & if\ y_{ij} > 0. \end{cases} \qquad (1)$$

The upper part of expression (1) shows the probability π_{ij} for person j not to eat at time point i.[1] The bottom part represents the conditional probability distribution $h(y_{ij})$ for positive values, i.e. the expected amount eaten by person j at time point i if eating occurred. $h(y_{ij})$ is weighted by the probability that a person did eat which translates to subtracting the probability of no eating from 1 ($1 - \pi_{ij}$). A gamma distribution is used for $h(y_{ij})$.[2]

In the zero part of the model a multilevel logistic regression predicts the log-odds of no eating for person j at time point i ($\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right)$).[3] Equation (2) shows that the log-odds of no eating can be predicted as a function of Level-1 and Level-2 covariates:

$$log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_{00} + \beta_{01}L1predictor_{ij} + u_{01j}L1predictor_{ij} +\ \beta_{02}L2predictor_j + u_{0j} \quad (2)$$

[1] Note that the multilevel logistic regression predicts NO dietary intake (i.e. $y_{ij}=0$). Typically logistic regressions predict $y=1$.
[2] The gamma distribution is assumed for Level-1 residuals (i.e. the part of the dependent variable which is not explained by the predictors of the model).
[3] Technically, the probability π_{ij} from expression (1) is not directly predicted, instead the log-odds of no eating are predicted, as the multilevel logistic regression is modelled on the logit scale to accommodate the restricted range of probabilities (between 0 and 1).

$L1predictor_{ij}$ in equation (2) represents a Level-1 covariate assessed at time point i in person j, e.g. participant j's momentary affect at measurement occasion i. $L2predictor_j$ is a Level-2 covariate of person j, e.g. participant j's BMI. β_{00} is the overall intercept, i.e. the mean of the log-odds of no eating across all participants when all predictors are equal to 0. The coefficient β_{01} represents the expected change in log-odds of no eating for a one-unit increase in *L1predictor*. The expected change in log-odds of no eating for a one-unit increase in *L2predictor* is expressed by β_{02}. u_{0j} represents the random intercept of person j, i.e. person-specific differences in the log-odds of no eating. u_{01j} denotes the random effect of *L1predictor* in person j, i.e. person-specific differences in the effect of *L1predictor* on the log-odds. The first subscript 0 of the parameters indicates that the equation refers to the zero part of the model (i.e. the logistic regression predicting if the participant did not eat).

Turning to the continuous part of the semicontinuous variable (see lower part of Equation 1), a multilevel gamma regression is used to predict the expected log amount of dietary intake of person j at time point i ($log(\mu_{ij})$) when eating occurred. μ_{ij} is modelled on the log scale due to the fact that the gamma distribution only supports positive values.

Equation (3) shows that a function of Level-1 and Level-2 covariates can be used to predict the (log) amount of dietary intake:

$$log(\mu_{ij}) = \beta_{10} + \beta_{11}L1predictor_{ij} + u_{11j}L1predictor_{ij} + \beta_{12}L2predictor_j + u_{1j} + \varepsilon_{ij} \quad (3)$$

$L1predictor_{ij}$ represents a Level-1 covariate of person j at time point i, $L2predictor_j$ a Level-2 covariate of person j and β_{10} the overall intercept, i.e. conditional mean of the (log) amount consumed across all participants when all predictors are equal to 0 given that the response is non-zero (i.e. dietary intake occurred). The regression coefficients β_{11} and β_{12} represent the expected change in the (log) amount consumed for a one-unit increase in *L1predictor* and *L2predictor*, respectively. The parameter u_{1j} reflects the random intercept of person j, i.e. person-specific differences in the expected (log) amount consumed. u_{11j} is the random effect of predictor *L1predictor* in person *j*, i.e. person-specific difference in the effect of *L1predictor* on the expected (log) amount consumed. The error term ε_{ij} denotes the Level-1 residual, i.e. difference between the predicted value and the observed value of person *j* at time point *i*. The first subscript 1 of the parameters denotes the positive part of the model.

The two processes modelled through the multilevel logistic and gamma regression are likely not independent. Therefore, an important consideration in two-part modelling, as highlighted by Olsen and Schafer [2] for longitudinal data, is to account for this potential relation. To do so, the correlation between the random effects across the two parts (often called cross-part correlation) is modelled. Random effects are assumed to be jointly normal and possibly correlated as illustrated in expression (4). The random effects of the logistic part (summarized in the vector $\boldsymbol{u_{0j}}$) and of the gamma part (summarized in the vector $\boldsymbol{u_{1j}}$) are assumed to come from a multivariate normal distribution with a mean vector of 0 and an unknown covariance matrix $\boldsymbol{\Sigma}$:

$$\begin{bmatrix} \mathbf{u_{0j}} \\ \mathbf{u_{1j}} \end{bmatrix} \sim \mathrm{MVN}(\mathbf{0}, \boldsymbol{\Sigma}), \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma_0} & \\ \boldsymbol{\Sigma_{01}} & \boldsymbol{\Sigma_1} \end{bmatrix} \tag{4}$$

$\boldsymbol{\Sigma_0}$ and $\boldsymbol{\Sigma_1}$ are variance-covariance matrices of the random effects within the logistic and the gamma part of the model, respectively. $\boldsymbol{\Sigma_{01}}$ denotes the covariance matrix of the random effects across the two model parts, i.e. cross-part covariance matrix. The size of $\boldsymbol{\Sigma_0}, \boldsymbol{\Sigma_1}$ and $\boldsymbol{\Sigma_{01}}$ is determined by the number of random effects included in the model. Expression (5) shows exemplary the random effect variance-covariance matrix of a multilevel two-part model with the random intercept vectors $\boldsymbol{u_{0j}}$ and $\boldsymbol{u_{1j}}$.

$$\begin{bmatrix} \mathbf{u_{0j}} \\ \mathbf{u_{1j}} \end{bmatrix} \sim \mathrm{MVN}(\mathbf{0}, \boldsymbol{\Sigma}), \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2{}_{u_0} & \\ \sigma_{u_0 u_1} & \sigma^2{}_{u_1} \end{bmatrix} \tag{5}$$

$\boldsymbol{\Sigma}$ contains $\sigma^2{}_{u_0}$, the person-to-person variability in the log-odds of no eating, $\sigma^2{}_{u_1}$, the person-to-person variability in the expected (log) amount consumed, and $\sigma_{u_0 u_1}$, the covariance between these two random intercepts. Hence, three parameters (2 variances, 1 covariance) are estimated.

Expression (6) shows the random effects variance-covariance matrix of a multilevel two-part model with the random intercept vectors $\boldsymbol{u_{0j}}$ and $\boldsymbol{u_{1j}}$ as well as one random effect in the logistic (summarized in the vector $\boldsymbol{u_{01j}}$) and one in the gamma (summarized in the vector $\boldsymbol{u_{11j}}$) part of the model:

$$\begin{bmatrix} \mathbf{u_{0j}} \\ \mathbf{u_{1j}} \\ \mathbf{u_{01j}} \\ \mathbf{u_{11j}} \end{bmatrix} \sim \mathrm{MVN}(\mathbf{0}, \boldsymbol{\Sigma}), \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2{}_{u_0} & & & \\ \sigma_{u_0u_1} & \sigma^2{}_{u_1} & & \\ \sigma_{u_0u_{01}} & \sigma_{u_1u_{01}} & \sigma^2{}_{u_{01}} & \\ \sigma_{u_0u_{11}} & \sigma_{u_1u_{11}} & \sigma_{u_{01}u_{11}} & \sigma^2{}_{u_{11}} \end{bmatrix} \tag{6}$$

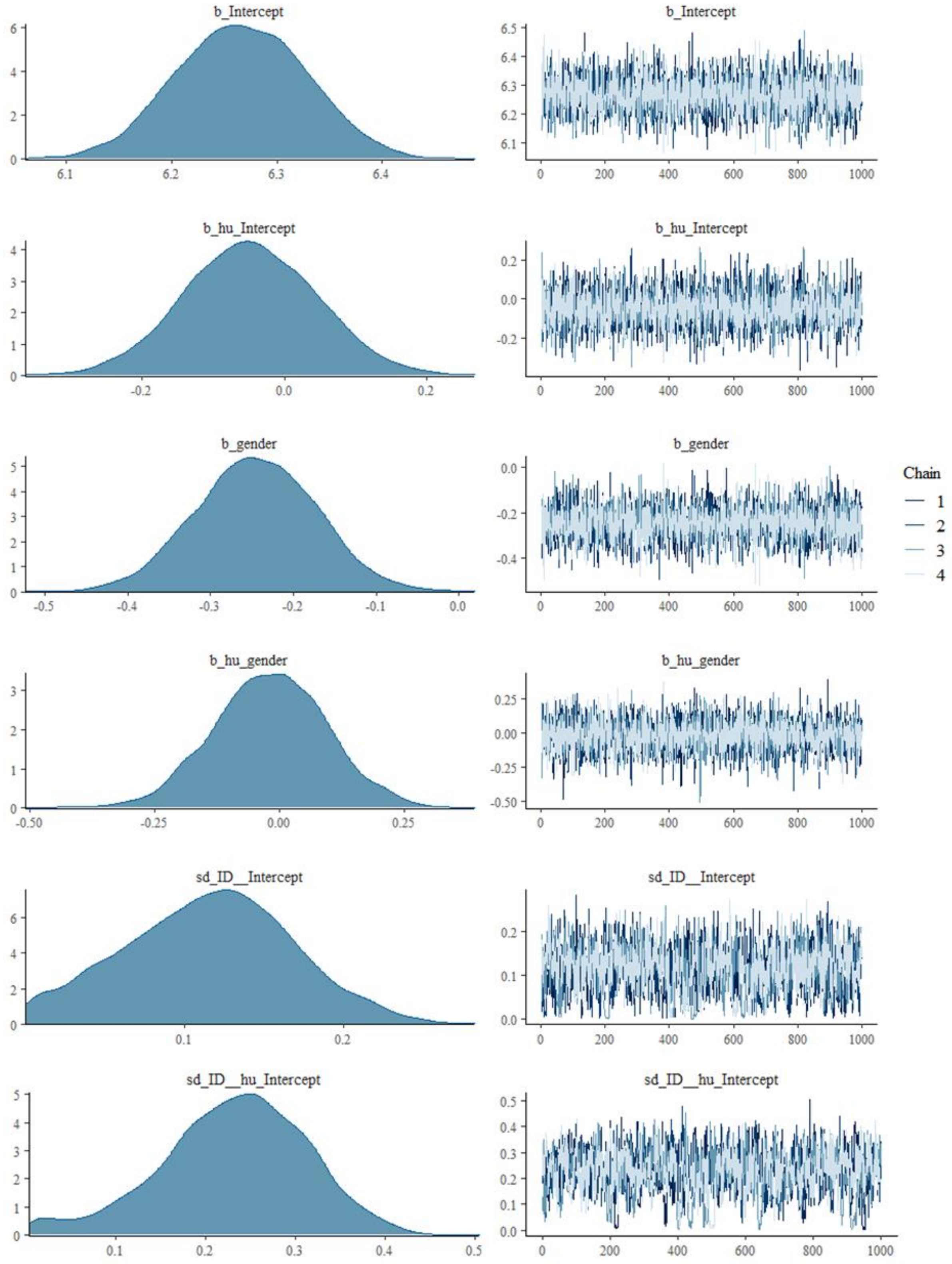
In this case, ten parameters (4 variances, 6 covariances) are estimated. $\sigma^2{}_{u_0}$, $\sigma^2{}_{u_{01}}$ and $\sigma_{u_0u_{01}}$ are elements of the matrix $\boldsymbol{\Sigma_0}$, $\sigma^2{}_{u_1}$, $\sigma^2{}_{u_{11}}$ and $\sigma_{u_1u_{11}}$ of $\boldsymbol{\Sigma_1}$ and $\sigma_{u_0u_1}$, $\sigma_{u_0u_{11}}$, $\sigma_{u_1u_{01}}$ and $\sigma_{u_{01}u_{11}}$ of $\boldsymbol{\Sigma_{01}}$. Note that brms provides standard deviations and correlations instead of variances and covariances.

1. Baldwin SA, Fellingham GW, Baldwin AS. Statistical models for multilevel skewed physical activity data in health research and behavioral medicine. Heal Psychol. 2016;35:552–62.

2. Olsen MK, Schafer JL. A Two-Part Random-Effects Model for Semicontinuous Longitudinal Data. J Am Stat Assoc. 2001;96:730–45.

Additional file 5. Density and Trace Plots

Figure 1. Density and Trace Plots of the Random Intercept Model with Level-2 predictor *gender*



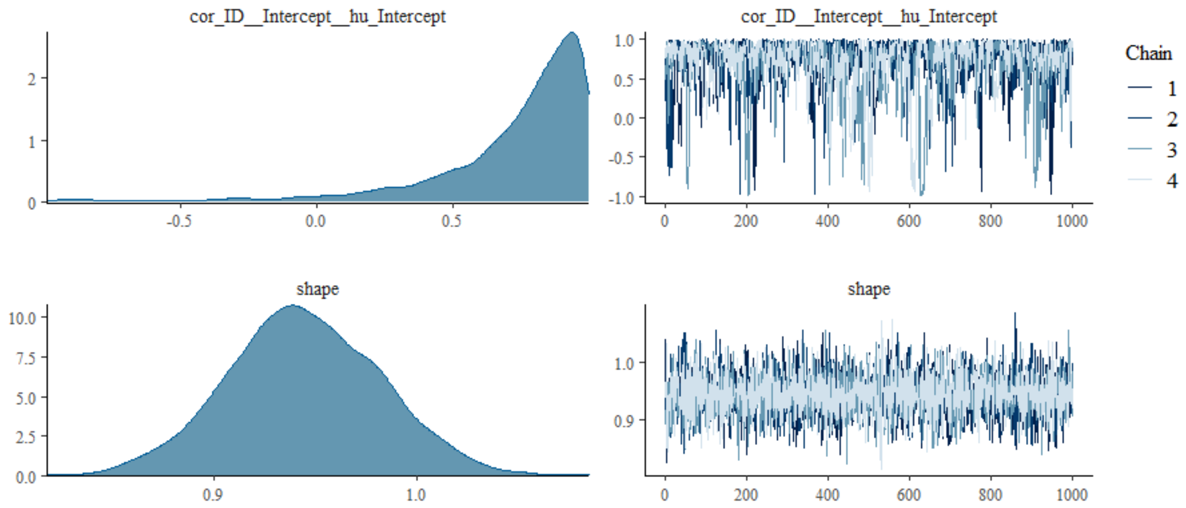
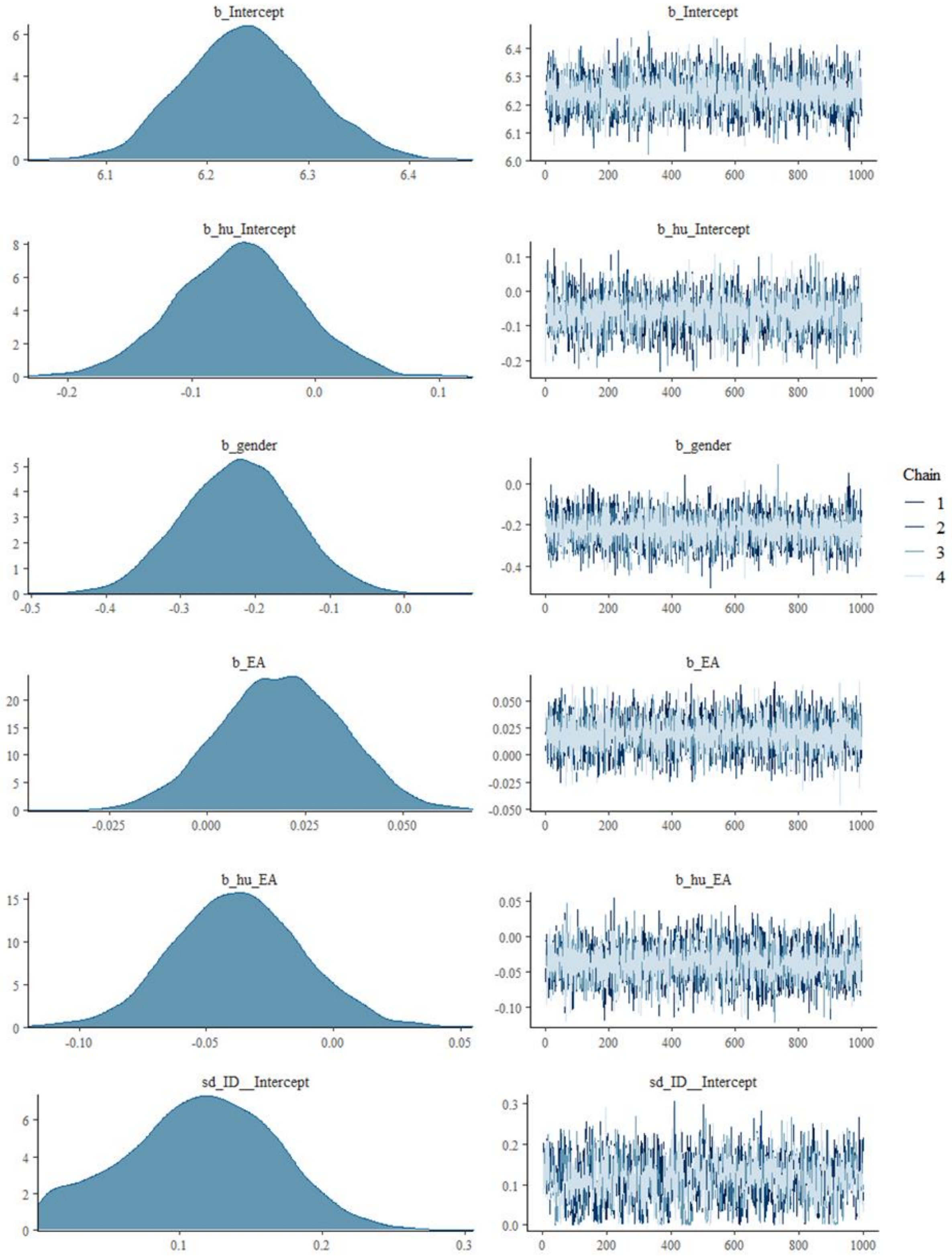
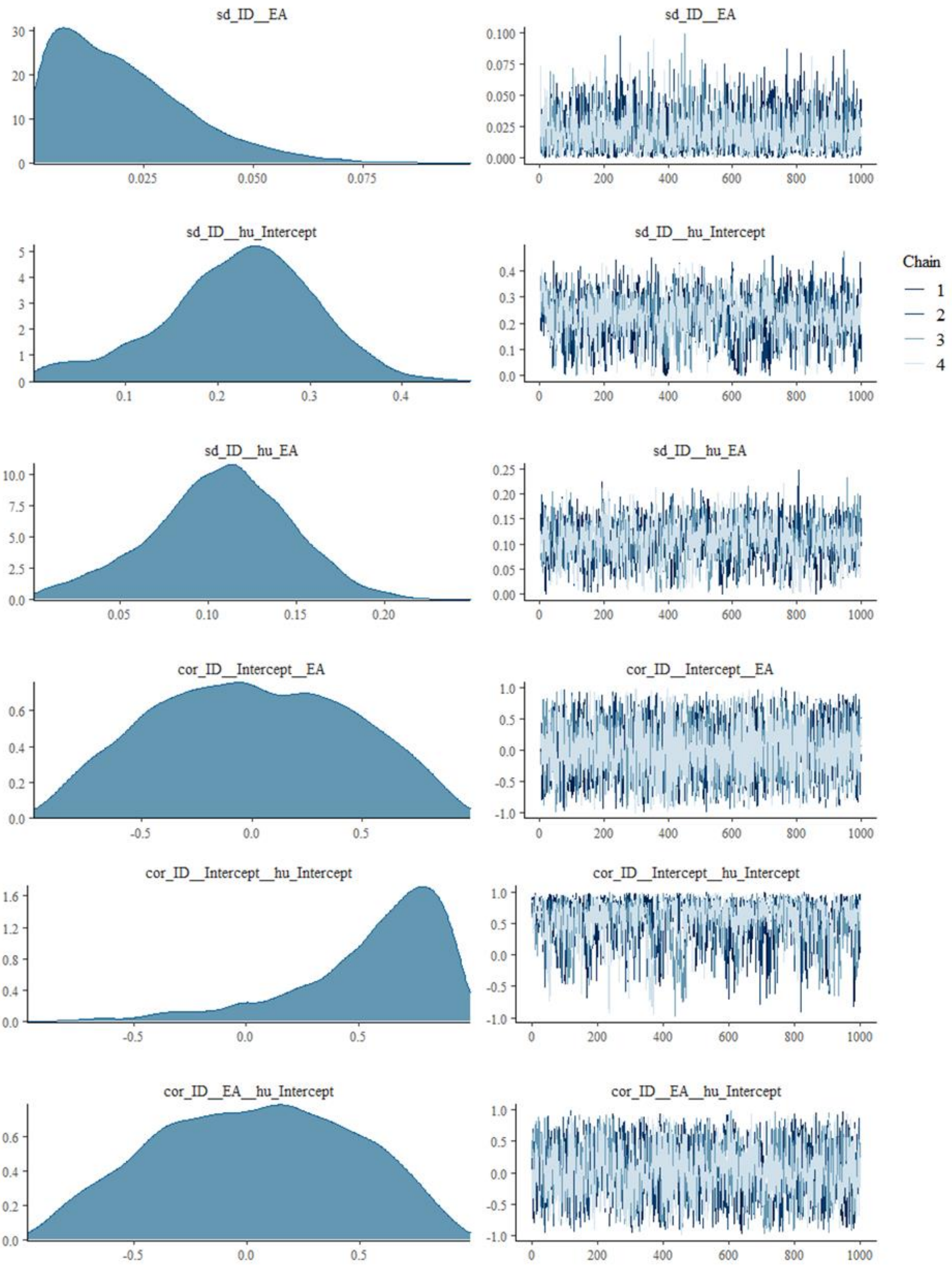
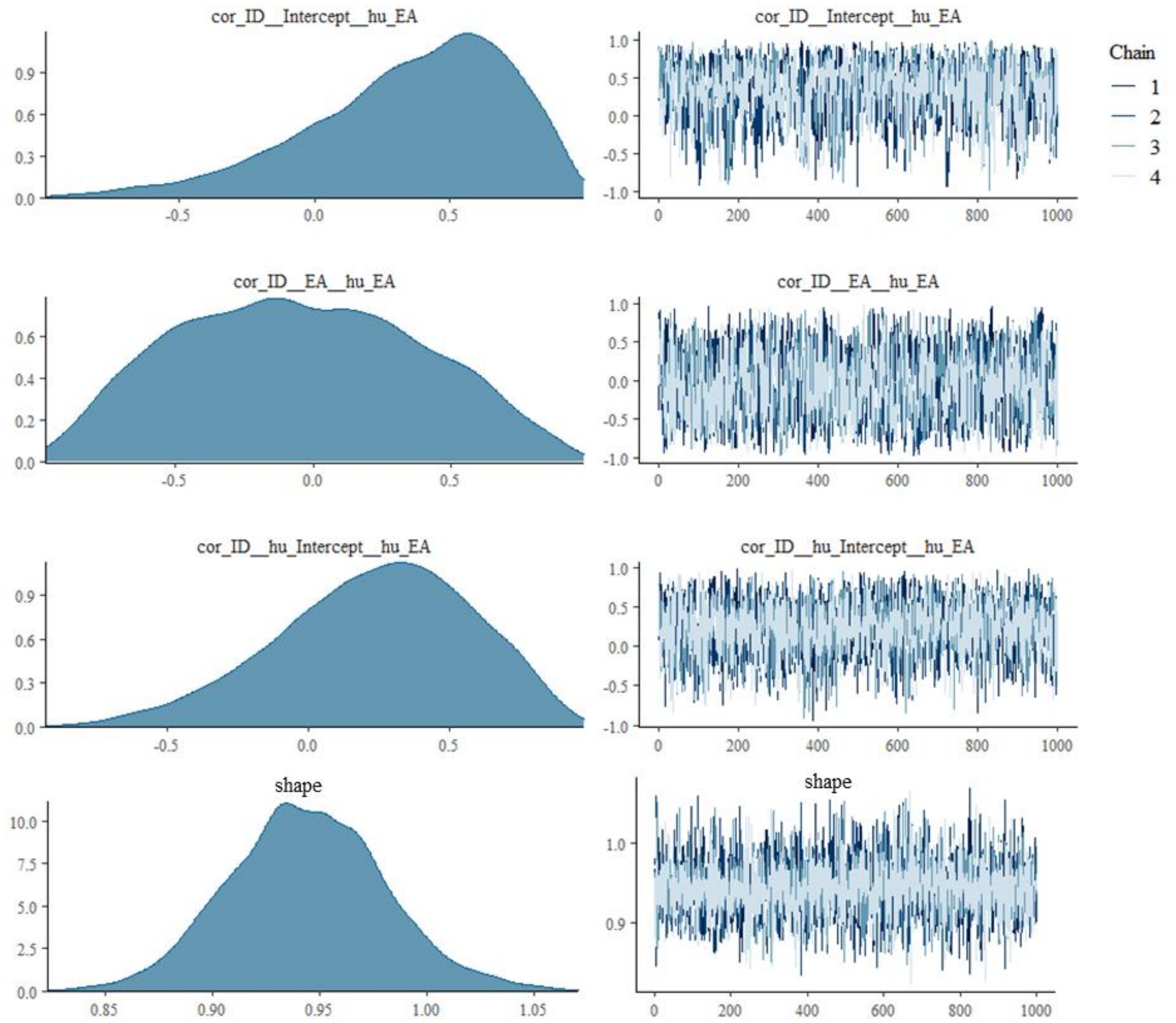

cor_ID__Intercept__hu_Intercept
cor_ID__Intercept__hu_Intercept
Chain
1
2
3
4
shape
shape

Figure 2. Density and Trace Plots of the Random Slope Model with Level-1 predictor energetic arousal (*EA*)




sd_ID__EA
sd_ID__EA
sd_ID__hu_Intercept
sd_ID__hu_Intercept
Chain
1
2
3
4
sd_ID__hu_EA
sd_ID__hu_EA
cor_ID__Intercept__EA
cor_ID__Intercept__EA
cor_ID__Intercept__hu_Intercept
cor_ID__Intercept__hu_Intercept
cor_ID__EA__hu_Intercept
cor_ID__EA__hu_Intercept


cor_ID__Intercept__hu_EA
Chain
1
2
3
4
cor_ID__EA__hu_EA
cor_ID__hu_Intercept__hu_EA
shape

Additional file 6. Model Summaries

Figure 1. Model summary of the Random Intercept Model with Level-2 predictor *gender*

```
Family: hurdle_gamma
  Links: mu = log; shape = identity; hu = logit
Formula: energy_intake ~ 1 + gender + (1 | x | ID)
         hu ~ 1 + gender + (1 | x | ID)
   Data: data (Number of observations: 2044)
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup samples = 4000
```

Group-Level Effects:
~ID (Number of levels: 99)

	Estimate		Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sd(Intercept)	0.12	$(\sqrt{\sigma^2_{u_1}})$	0.05	0.01	0.22	1.00	795	599
sd(hu_Intercept)	0.23	$(\sqrt{\sigma^2_{u_0}})$	0.08	0.04	0.38	1.00	831	507
cor(Intercept,hu_Intercept)	0.71	$(\rho_{u_1 u_0})$	0.34	-0.34	0.99	1.00	704	627

Population-Level Effects:

	Estimate		Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	6.27	(β_{10})	0.06	6.14	6.39	1.00	4889	3135
hu_Intercept	-0.05	(β_{00})	0.09	-0.23	0.13	1.00	4074	2942
gender	-0.25	(β_{11})	0.07	-0.39	-0.10	1.00	4855	3059
hu_gender	-0.02	(β_{01})	0.11	-0.24	0.20	1.00	4241	2749

Family Specific Parameters:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
shape	0.94	0.04	0.87	1.02	1.00	5706	2843

```
Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

Figure 2. Model summary of the Random Slope Model with Level-1 predictor energetic arousal (*EA*)

```
Family: hurdle_gamma
  Links: mu = log; shape = identity; hu = logit
Formula: energy_intake ~ 1 + gender + EA + (1 + EA | x | ID)
         hu ~ 1 + EA + (1 + EA | x | ID)
   Data: data (Number of observations: 2025)
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup samples = 4000

Group-Level Effects:
~ID (Number of levels: 99)
```

	Estimate		Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sd(Intercept)	0.11	$(\sqrt{\sigma^2{}_{u_1}})$	0.05	0.01	0.21	1.00	801	611
sd(EA)	0.02	$(\sqrt{\sigma^2{}_{u_{12}}})$	0.02	0.00	0.06	1.00	2144	1789
sd(hu_Intercept)	0.22	$(\sqrt{\sigma^2{}_{u_0}})$	0.08	0.04	0.37	1.00	1098	950
sd(hu_EA)	0.11	$(\sqrt{\sigma^2{}_{u_{01}}})$	0.04	0.02	0.18	1.00	1154	1608
cor(Intercept,EA)	0.00	$(\rho_{u_1 u_{12}})$	0.45	-0.80	0.81	1.00	4486	2421
cor(Intercept,hu_Intercept)	0.56	$(\rho_{u_1 u_0})$	0.34	-0.36	0.95	1.00	843	899
cor(EA,hu_Intercept)	0.04	$(\rho_{u_{12} u_0})$	0.43	-0.78	0.80	1.00	1744	2524
cor(Intercept,hu_EA)	0.36	$(\rho_{u_1 u_{01}})$	0.37	-0.51	0.89	1.00	1177	1124
cor(EA,hu_EA)	-0.05	$(\rho_{u_{12} u_{01}})$	0.44	-0.81	0.78	1.00	1463	2423
cor(hu_Intercept,hu_EA)	0.24	$(\rho_{u_0 u_{01}})$	0.34	-0.50	0.82	1.00	2347	2418

Population-Level Effects:

	Estimate		Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	6.24	(β_{10})	0.06	6.12	6.36	1.00	4883	3222
hu_Intercept	-0.06	(β_{00})	0.05	-0.16	0.04	1.00	4046	2537
gender	-0.22	(β_{11})	0.07	-0.37	-0.08	1.00	4479	2810
EA	0.02	(β_{12})	0.02	-0.01	0.05	1.00	5552	3172
hu_EA	-0.04	(β_{01})	0.03	-0.09	0.01	1.00	5232	3301

```
Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
shape     0.94      0.04     0.88     1.02 1.00     5073     2594

Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

……………………………………………………………………………………………...

Paper 3

Ruf, A., Neubauer, A.B., Koch, E.D., Ebner-Priemer, U., Reif, A., & Matura, S. (2022). Individual differences in the dietary response to stress in ecological momentary assessment: Does the individual-difference model need expansion? *Applied Psychology: Health & Wellbeing,* 1-21. https://doi.org/10.1111/aphw.12400

……………………………………………………………………………………………...

Received: 21 June 2022 | Accepted: 30 August 2022

DOI: 10.1111/aphw.12400

ORIGINAL ARTICLE

APPLIED PSYCHOLOGY: Health and Well-Being IAAP

Individual differences in the dietary response to stress in ecological momentary assessment: Does the individual-difference model need expansion?

Alea Ruf[1] | **Andreas B. Neubauer**[2,3] | **Elena D. Koch**[4] | **Ulrich Ebner-Priemer**[4,5] | **Andreas Reif**[1] | **Silke Matura**[1]

[1]Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, Goethe University Frankfurt, University Hospital, Frankfurt, Germany

[2]DIPF|Leibniz Institute for Research and Information in Education, Frankfurt, Germany

[3]Center for Research on Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt, Germany

[4]Mental mHealth Lab, Institute of Sports and Sports Science, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

[5]Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

Correspondence
Alea Ruf, Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital, Goethe University, Heinrich-Hoffmann-Straße 10, 60528 Frankfurt, Germany.
Email: alea.ruf@kgu.de

Abstract

According to the individual-difference model, individuals differ in the way stress changes their eating behaviour. Research shows that some increase, some decrease, and others show no change in food intake. Despite numerous efforts to identify moderating variables that explain these individual (i.e., between-person) differences, evidence remains inconclusive. The present study aims at deepening the understanding of the stress and eating relationship by applying ecological momentary assessment to study (1) the influence of stress on whether and how much individuals eat and (2) the moderating role of gender, age, BMI, trait stress-eating, and eating styles. The APPetite-mobile-app was used for 3 days to capture actual food intake (event-contingent) and perceived stress (signal-contingent). Data of 154 healthy adults suggest that stress is not associated with whether but how much individuals eat. Only gender moderated the relationship between stress and the amount of food intake. Individual differences were small indicating that an individual's dietary response to stress might not be as stable as yet

Funding Information
This work was supported by the European Union's Horizon 2020 Framework Programme under grant agreement No 728018. The funding source has had no involvement in the study design, data collection, interpretation of the findings, or writing of this manuscript.

assumed. Moreover, a study suggests that time-varying factors (e.g., food availability) moderate the stress and eating relationship. Hence, intraindividual (i.e., within-person) variability may be relevant. Therefore, we propose an expansion of the individual-difference model, which accounts for time-varying factors.

KEYWORDS
diet, ecological momentary assessment, food intake, individual differences, stress, time-varying moderators

INTRODUCTION

Human health is substantially and directly influenced by diet and stress. Poor dietary habits and elevated levels of stress are linked to numerous negative health outcomes, such as cardiovascular diseases (Kivimäki & Steptoe, 2018; Micha et al., 2017). Beyond that, stress has an indirect impact on health through changes in health-related behaviours, including diet (O'Connor et al., 2021). A substantial body of research has shown that stress is associated with changes in dietary intake (for overviews, see Araiza & Lobel, 2018; Hill et al., 2021). Even though people commonly associate stress with overeating, studies assessing the link between diet and stress have produced mixed results. While some studies found that stress increases food intake (e.g., Wardle et al., 2000), others found decreases in food consumption (e.g., Stone & Brownell, 1994). The inconsistency in findings is highlighted by a recent meta-analysis, which found only a small positive effect size for the relationship between stress and overall food intake due to considerable heterogeneity across subgroup analyses and across the 54 studies overall (Hill et al., 2021). To some extent, differences in study design and in the measurement of stress and diet might have contributed to these heterogeneous findings. However, individual (i.e., between-person) differences in the dietary response to stress seem to be the primary cause of the observed heterogeneity. As early as 1994, a review concluded that there is strong evidence for the individual-difference model—as opposed to a general effect model (Greeno & Wing, 1994). The individual-difference model is based on the assumption that the effect of stress on eating is determined by individual differences in learning history, attitudes, or biology.

Individual differences in the dietary response to stress

Studies have shown that individuals differ in the way stress changes their eating behaviour. Some individuals increase, some decrease food intake, whereas others do not change food consumption when experiencing stress. Estimates derived from self-reports indicate that about 36–42% of individuals report eating more, 26–38% less and the remaining report no consistent change as a response to stress (Epel et al., 2004; Oliver & Wardle, 1999). Despite various efforts to identify person-level factors that underlie individual differences in the stress and eating relationship, the evidence is inconclusive. For instance, some studies found gender differences, with men decreasing food intake under stress and women showing some increases in eating

(Grunberg & Straub, 1992). However, other studies did not find gender differences (e.g., Conner et al., 1999). Weight and eating styles (e.g., dietary restraint and emotional eating) have also been studied widely as potential moderators of the stress and eating relationship. Some evidence suggests that individuals with higher body weight (e.g., Cotter & Kelly, 2018; O'Connor et al., 2008) and individuals higher in emotional eating and dietary restraint (e.g., O'Connor et al., 2008; Wallis & Hetherington, 2004; Wardle et al., 2000) are more likely to increase food intake when experiencing stress. Nevertheless, inconsistencies are also present here as other studies found no moderating effect of emotional eating (Conner et al., 1999) as well as restrained eating (Conner et al., 1999; Pollard et al., 1995). The impact of potential moderators of the stress and eating relationship, such as gender, age, BMI, and eating style (i.e., dietary restraint), was also explored in the meta-analysis by Hill et al. (2021). However, none of these variables significantly moderated the relationship between stress and overall food intake. Based on the findings from the meta-analysis, Hill et al. highlight the need for (1) more detailed measures of the nature of the stressors, (2) more accurate assessments of food consumption, such as energy intake, (3) more studies that test key moderating variables of the stress and eating relationship, (4) assessment of eating styles, and (5) accurate measures of weight, height and diet status. Furthermore, they emphasise the importance of taking dispositional stress-related eating (i.e., self-reported tendency to eat more, less or the same in response to stress) into account.

Ecological momentary assessment of the stress and eating relationship

The influence of stress on eating behaviour is highly complex (Hill et al., 2021). Ecological momentary assessment (EMA) allows studying complex psychological, behavioural, and physiological processes through the repeated assessment of behaviours (e.g., food intake), experiences (e.g., perceived stress), and physiological parameters multiple times a day in real life (Smyth & Stone, 2003). Given that eating is a repeated-occurrence health behaviour that is performed several times a day (Dunton, 2018), EMA seems particularly suited to study the complex relationship between stress and food intake when and where it naturally occurs. It circumvents disadvantages of traditional approaches (e.g., retrospective self-reports and laboratory tasks), by minimizing recall bias, maximizing ecological validity and capturing within-person processes and variation over time and across settings (Shiffman et al., 2008). Furthermore, Araiza and Lobel (2018) point out that a closer study of the stress and eating relationship could be achieved and reliability and validity could be increased through novel and sophisticated methodological approaches, such as EMA.

Despite its potential, the number of studies using EMA to investigate the stress and eating relationship is limited. One EMA study assessed the relationship between daily hassles and snack intake in African American women (Zenk et al., 2014). Participants were more likely to consume snack foods on days they experienced more daily hassles. However, no association between experiencing a stressful event and concurrent as well as subsequent snack food intake was identified on the momentary level (i.e., within-day level). Reichenberger et al. (2018) studied the effect of stress on taste- and hunger-driven eating in an EMA setting. While hunger-driven eating refers to eating in response to physiological feelings of hunger, taste-eating describes food intake that is driven by the anticipated pleasure associated with the taste of foods. They found that stress decreased taste-eating. This relationship was not moderated by gender, BMI, and eating styles (emotional, external, and restrained eating). Hunger-eating was not significantly influenced by stress. Again, gender, BMI, and eating styles did not moderate the relationship between stress and hunger-eating. A recent study by Reichenberger et al. (2021) used EMA to study the moderating role of trait stress-eating (i.e., an individual's self-reported tendency to eat more and less or the same in response to stress) in the relationship between stress and eating. When data collected throughout the day were aggregated (i.e., to the day level), they found that trait stress-eating moderated the relationship between stress and food intake. Individuals with high trait stress-eating reported more food intake on days with higher stress. No effect of stress on food intake was found in individuals with low trait stress-eating. Contrary to the day level, trait stress-eating did not moderate the stress and food intake relationship on the within-day level. There was also no main effect of stress on food intake. It should be noted, however, that only perceived food intake was assessed. That is, for each eating episode since the last prompt, participants reported how much they had eaten on a scale from 0 (eaten too little) to 100 (eaten too much). Even though the importance of assessing actual food intake, such as energy intake, has been highlighted (Araiza & Lobel, 2018; Hill et al., 2021), to the best of our knowledge, there are no EMA studies available that assess the association between stress and actual food intake in healthy adults. Only one small EMA study assessed stress and calorie intake in nine patients with type 2 diabetes and found a positive association between stress and calorie intake from snacks as well as a negative association between stress and calorie intake from lunch and dinner (Inada et al., 2019). Presumably, due to the small sample size, individual differences of the stress and food intake relationship were not taken into account.

Previous research has either studied if stress is associated with whether individuals eat (e.g., Zenk et al., 2014) or how much they eat (e.g., Reichenberger et al., 2021). What has been overlooked so far is that the occurrence and the amount of food intake are likely not independent. When the association between stress and the amount of food intake is studied, time intervals in which no eating is reported are excluded (e.g., Reichenberger et al., 2021: study 1—2318 out of 4656). This causes loss of important information (Tooze et al., 2002) and can cause bias in the parameter estimates (Liu et al., 2008; Su et al., 2009). However, including intervals in which participants did not consume any food yields a zero-inflated outcome (i.e., one considerable part of the outcome is equal to zero). This type of data can be challenging as traditional linear multilevel modelling cannot be applied. A promising statistical approach to analyse this type of data is multilevel two-part modelling, which accounts for the potential dependency between the occurrence and the amount of food intake (see Ruf, Neubauer, et al., 2021, for a detailed description of this approach).

The present study

The present study addresses the need for research that assesses the stress and eating relationship in an EMA setting based on accurate dietary assessments. Hence, the present study uses an EMA tool, which showed good validity to capture actual food intake (Ruf, Koch, et al., 2021). Following the recommendations by Hill et al. (2021), the present study assesses the moderating effect of gender, age, weight, eating styles, and trait stress-eating on the stress and eating relationship. It is examined (1) whether individuals differ in the stress and eating relationship, (2) whether individual differences in the stress and eating relationship can be explained by person characteristics (i.e., gender, age, BMI, trait stress-eating, and eating styles), and (3) whether these findings support the individual-difference model of stress-eating (Greeno & Wing, 1994). Furthermore, the present study is the first to use multilevel two-part modelling to assess the stress and eating relationship. This offers novel and distinct insights in terms of the occurrence as well as the amount of food intake.

Given the inconsistent body of evidence, we hope that the novel approach of our study—assessment of actual food intake in an EMA setting combined with sophisticated multilevel two-part modelling—will allow us to deepen the understanding of the stress and eating relationship. Understanding which individuals are more likely to eat or prone to overeating when experiencing stress in daily life is crucial in order to identify individuals at higher risk for diet-related negative health outcomes.

METHODS

Procedure

Data were collected within the Eat2beNICE-APPetite-study (parts of the data have been used to study different research questions, see Ruf, Koch, et al., 2021, and Ruf, Neubauer, et al., 2021). Participants completed two in-person sessions as well as an EMA period. Body weight and body height were measured in the first in-person session and were used to calculate BMI. Furthermore, participants completed questionnaires and received detailed training to familiarize with the APPetite-mobile-app used for the EMA assessment (for further details see Ruf, Koch, et al., 2021). The local ethics committee approved the study. All subjects declared that they understood the study procedure and signed a written informed consent.

EMA protocol

Participants received a study smartphone to complete the EMA protocol of the APPetite-mobile-app for three consecutive days (two weekdays and one weekend day). Between 8 a.m. and 10 p.m. participants received eight semirandom signal-contingent prompts (at least 1 h in-between prompts). Each prompt assessed stress and food availability. Food intake was captured event-contingent through the incorporated APPetite-food record. Hence, food intake could be recorded at any time. At 9 p.m. there was a time-contingent prompt asking if all foods and drinks of the day have been recorded. Further details on the APPetite-mobile-app are available in Ruf, Koch, et al. (2021).

Sample

Participants from the Longitudinal Resilience Assessment-study (inclusion criteria described in Chmitorz et al., 2021) were invited to the study ‘APPetite: the influence of diet and physical activity on impulsivity and resilience’. In total, 185 healthy adults participated in the study. Four participants dropped out before starting the EMA assessment due to personal reasons (e.g., spontaneous trip abroad) or because they realized they were unable to respond to prompts (e.g., due to work commitments). Data of one participant had to be excluded as they proved to be untrue. Data of 26 participants were excluded as dietary intake was recorded poorly (e.g., only breakfast recorded). The final sample includes 154 participants (112 female, 42 male) with an average age of 28.91 years ($SD = 7.75$). The sample has a mean BMI of 24.20 ($SD = 4.09$).

Measures

Food intake

The APPetite-mobile-app (Ruf, Koch, et al., 2021) was used to capture dietary intake. This mobile application comprises a food record. Participants were asked to enter all foods and drinks as soon as possible after consuming them. Foods and drinks were recorded through a six-step process: (1) Selection of meal type, (2) entry of time of intake, (3) selection of consumed foods and drinks, (4) specification of consumed amounts, (5) presentation of reminder for commonly forgotten foods, and (6) indication of predominant reason for eating or drinking. The obtained dietary data were transferred to myfood24-Germany, an online 24-h dietary recall (Koch et al., 2020), by trained staff in order to generate nutritional values, such as the exact energy intake in kilocalories (kcal), which is the outcome in the present study. The APPetite-mobile-app was subject to a feasibility, usability and validation study. Results indicated that the APPetite-mobile-app is a feasible EMA tool and a valid dietary assessment method that is is likely more precise than 24-h recalls (Ruf, Koch, et al., 2021).

Stress

Three items (adapted from Reichenberger et al., 2018) were used to capture perceived stress. The first item assessed how stressed participants felt since the last prompt. Responses were rated on a visual analogue scale from 0% (*not at all*) to 100% (*very stressed*). Two stress items, based on the Perceived Stress Scale (Cohen et al., 1983), assessed whether participants felt that they 'could not cope with all the things they had to do' and whether they were 'on top of things' since the last prompt on a visual analogue scale from 0 (*not at all*) to 100 (*very much*). In the first prompt per day, participants were instructed to rate stress since waking up instead of since the last prompt. McDonald's Omegas (Geldhof et al., 2014) for the three stress items were 0.648 (within) and 0.895 (between) in the present sample. Based on the three items (third item reversed), a mean stress score was calculated for each prompt.

Food availability

Because the effect of stress on food intake can only be reliably studied in time intervals in which food was actually available, food availability was assessed on a visual analogue scale from 0 (*not available at all*) to 100 (*easily available*) since the last prompt.

Trait stress-eating

The Salzburg Stress Eating Scale (SSES) (Meule et al., 2018) was used to capture trait stress-eating. Each of the 10 items describes a stressful situation. Participants were asked whether they eat a lot less (1), less (2), the same (3), more (4), or a lot more (5) than usual in this situation. A mean score was calculated. Mean SSES scores below 3 indicate an individual reports to decrease food intake when experiencing stress, above 3 that an individual reports to increase food intake under stress. A mean SSES score of 3 suggests that an individual reports to not change food intake when feeling stressed. Internal consistency was $\alpha = .86$ in the present sample.

Eating styles

The German version of the Three-Factor-Eating-Questionnaire (Stunkard & Messick, 1985), the questionnaire 'Fragebogen zum Ernährungsverhalten' (FEV; Pudel & Westenhöfer, 1989), was used to assess three eating styles: cognitive restraint of eating, disinhibition, and hunger. The questionnaire was chosen as its reliability and validity was evaluated in three large German samples (total $N > 80{,}000$; Pudel & Westenhöfer, 1989). The subscale cognitive restraint consists of 21 items, disinhibition of 16 items, and hunger of 14 items, all coded as 0 or 1. A sum score was calculated for each subscale. Higher subscale values indicate stronger cognitive restraint, greater disinhibition, and more pronounced feelings of hunger, respectively. In the present sample, internal consistency was $\alpha = .85$ for cognitive restraint, $\alpha = .81$ for disinhibition, and $\alpha = .70$ for hunger.

Data preprocessing

Due to poor or biased dietary data, single days of the EMA assessment of some participants had to be excluded (13 days in total).

Each time interval for which stress was assessed (i.e., time between current prompt and previous prompt/waking up) was paired with concurrent energy intake in kcal. Concurrent energy intake was defined as the sum of any intake of energy within the respective time interval.

The length of each time interval varied due to the semirandom sampling protocol. Furthermore, the assessment of stress 'since waking up' in the first prompt as well as the postponement of prompts yield either shorter time intervals than standardized (minimum time between two prompts = 1 h) or rather long time intervals. Time intervals shorter than 15 min ($n = 144$) and longer than 3 h ($n = 135$) were excluded. 314 time intervals were excluded as the stress items had not been completed. In addition, time intervals in which food availability was rated 10 or lower ($n = 191$) were excluded. However, the assessment of food availability was added to the study a few months after data collection started. Hence, we were unable to exclude time intervals of the first 33 participants based on this criterion. The final sample includes 2779 time intervals.

The Level-1 predictor stress was divided by 10 to avoid estimation problems (due to large differences in variance of the predictor and the outcome) and centred on the person-mean to generate unbiased estimates of the within-person effect (Wang & Maxwell, 2015). We centred the Level-2 predictor age around 30 years and BMI around 25 (i.e., the constant 30 or 25 were subtracted from participants' age or BMI respectively as recommended by Viechtbauer, 2022, to make the model intercept more interpretable). For the same reason, Level-2 predictor trait stress-eating was centred around 3. The Level-2 predictors dietary restraint, disinhibition, and hunger were centred on the grand-mean. Gender was coded as 0 (male) and 1 (female).

Data analysis

Due to the nested data structure (time intervals [Level 1] nested within individuals [Level 2]) and the zero-inflated, right-skewed outcome energy intake in kcal, multilevel two-part models were used for analysis. More specifically, the model we applied combines a multilevel logistic regression for the zero part of the outcome (to predict whether an individual eats in a given time interval) and a multilevel gamma regression for the continuous part of the outcome (to predict how much is eaten, if an individual eats in a given time interval), while accounting for the potential dependency between the two outcome components. This approach allows differentiating between stress influencing either the occurrence or the amount of food intake (or both). Therefore, findings are separately reported for the occurrence (zero part of the model) and the amount of food intake (continuous part of the model) in the results section. While logistic regressions typically predict the outcome to be 1, the multilevel logistic regression of our two-part model predicts no food intake (outcome = 0), that is, the probability not to eat for a given individual in a given time interval. These models were run using the R-package brms (Bürkner, 2017, 2018), which supports Bayesian multilevel modelling. Details on this type of analysis (e.g., implementation and interpretation) can be found in Ruf, Neubauer, et al. (2021).

To examine individual (i.e., between-person) differences in the within-person effect of stress on energy intake, a model with the Level-1 predictor stress in both parts of the model (i.e., the logistic regression as well as the gamma regression) was run. Next, seven separate models were run to test the association between the Level-1 predictor stress in interaction with the Level-2 predictor (1) gender, (2) age, (3) BMI, (4) trait stress-eating, (5) dietary restraint, (6) disinhibition, or (7) hunger (cross-level interaction), and energy intake in both model parts. If more than one of the cross-level interactions was significant, a combined model with all significant moderators was run. All models included a random intercept (i.e., we expect individuals to differ in their average probability not to eat and the average amount of energy intake) and a random slope for stress to examine whether the effect of stress differs between individuals.

Model parameters were estimated based on 4000 iterations. All other sampling and prior parameters were maintained as brms defaults. Analyses were performed using R version 4.0.5, RStudio version 1.4.1106 (RStudio Team, 2020), brms version 2.15, and rstan version 2.21.2 (Stan Development Team, 2020). The data and R code that support the findings of this study are available in Data S1 and S2 of this article.

RESULTS

Descriptive findings

Descriptive statistics of the variables can be found in Table 1. In the trait questionnaire, 27 participants reported not to change (SSES mean score = 3), 72 to decrease (SSES mean score <3), and 55 to increase (SSES mean score >3) food intake when experiencing stress.

In 1201 time intervals, no food intake was reported. Within time intervals in which participants ate ($n = 1578$), on average 466 kcal ($SD = 381$) were consumed. Mean compliance with the signal-contingent prompts (i.e., percentage of complete prompts within received prompts) was 89.3 ($SD = 12.2$) (not including participants and days that were excluded as a whole from final analyses due to poor or biased dietary data, but including time intervals that were excluded from final analyses based on interval length and food availability).

TABLE 1 Descriptive statistics of the Level-1 predictor stress ($N = 2779$) and Level-2 predictors trait stress-eating and eating styles ($N = 154$)

	M	SD	Range
Level-1			
Stress	18.65	12.38 (within) 12.35 (between)	0–100
Level-2			
Trait stress-eating	2.93	0.56	1.4–4.5
Cognitive restraint	6.91	4.53	0–19
Disinhibition	5.50	3.09	0–15
Hunger	5.09	2.92	1–12

Findings from the multilevel two-part models

Stress

Estimates of the zero part of the multilevel two-part model are modelled on the logit scale. The intercept of the zero part represents the average log-odds of no energy intake in time intervals with average stress (stress = 0). To transform the log-odds to the probability of no energy intake, we use the plogis-function in R. In model 1 (see Table 2), the mean probability of no energy intake in time intervals with an average stress level is 0.43. Credible intervals (95% CI) of fixed effects that do not include 0 indicate a significant effect. Hence, there is no significant fixed effect of stress on the probability not to eat. Note that nonpositive estimates for standard deviations (*SD*) are not allowed, and the lower limit of the CI for random effects will therefore always be positive. Accordingly, lower limit of the CI of random effects that are equal to 0.00 suggest that individual differences in the intercept or the effect of stress are small and possibly not statistically meaningful. Hence, as the lower limit of the 95% CI of the *SD* of the intercept is above 0.00, participants differ in the probability of no energy intake with an *SD* of 0.32. However, the random effect for stress in the zero part suggests that the effect of stress on the probability of no energy intake does not vary across participants. Consequently, individual/between-person differences in the within-person effect of stress on the probability of no energy intake are small and negligible (illustrated in Figure 1a).

Estimates of the continuous part are modelled on the log scale. The intercept of the continuous part represents the mean log energy intake in time intervals with average stress (stress = 0) in which eating occurred. To obtain the estimate of the intercept in the original metric (kcal), we calculate the exponential of the estimates. Participants consume on average 468.7 kcal in time intervals with average stress in which energy intake occurred. There is between-person variation in the log energy intake in time intervals with average stress in which energy intake occurred (see *SD* (intercept)). However, the effect of stress on the (log) energy intake does not vary across participants (see *SD* (stress)). This suggests that individual/between-person differences in the within-person effect of stress on the (log) energy intake are minor (illustrated in Figure 1b).

TABLE 2 Model estimates of the multilevel two-part model with the Level-1 predictor stress

	Zero part				Continuous part			
			95% CI				95% CI	
	Estimate	SE	LL	UL	Estimate	SE	LL	UL
	Model 1: Stress							
Fixed effects								
Intercept	−0.28	0.05	−0.38	−0.19	6.15	0.03	6.09	6.20
Stress	0.03	0.03	−0.04	0.09	−0.04	0.02	−0.08	0.00
Random effects								
SD (intercept)	0.32	0.06	0.19	0.44	0.11	0.04	0.04	0.19
SD (stress)	0.06	0.04	0.00	0.16	0.03	0.02	0.00	0.09

Note: CI = credible interval; LL = lower limit; UL = upper limit.

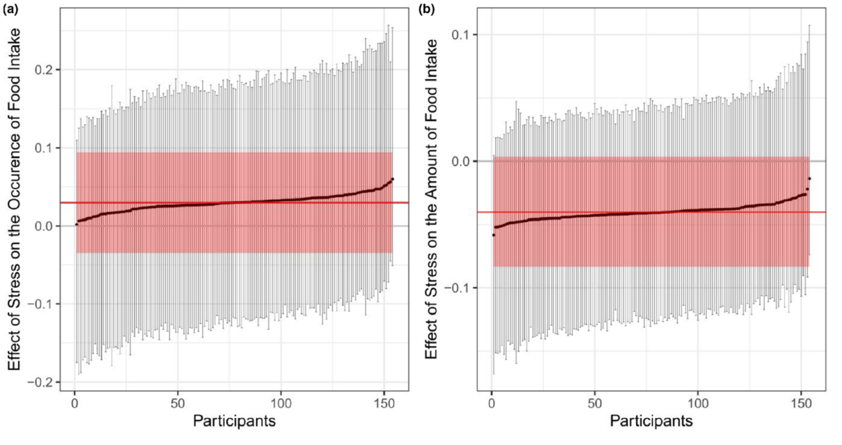


FIGURE 1 Individual/between-person differences in the within-person effect of stress on (a) the occurrence and (b) the amount of food intake. Note: Black dots represent estimates of the within-person effect for each participant. Vertical lines indicate the 95% credible interval of each within-person effect. The red horizontal line represents the average within-person effect. The shaded area around the red line indicates the 95% credible interval of the average within-person effect.

Gender, age, and BMI

Results of the three models assessing the effect of stress on energy intake as well as the moderating effect of gender, age, and BMI on the stress and food intake relationship in the zero and continuous part of the model are shown in Table 3.

There is no significant fixed effect of stress, gender, or their interaction in the zero part of the model. This suggests that (1) men and women do not differ in the probability not to eat, (2) stress is not associated with the likelihood that an individual eats, and (3) the relationship between stress and the probability not to eat is not moderated by gender.

The intercept of the continuous part represents the mean log energy intake for men (gender = 0) in time intervals with average stress (stress = 0) in which eating occurred. Male

TABLE 3 Model estimates of the multilevel two-part models of the moderating effect of gender, age, and BMI

	Zero part				Continuous part			
			95% CI				95% CI	
	Estimate	SE	LL	UL	Estimate	SE	LL	UL
	Model 2: Gender							
Fixed effects								
Intercept	−0.25	0.09	−0.43	−0.08	6.28	0.05	6.18	6.37
Stress	0.05	0.08	−0.10	0.21	−0.13	0.05	−0.23	−0.03
Gender	−0.03	0.10	−0.24	0.17	−0.18	0.06	−0.30	−0.07
Stress * gender	−0.03	0.09	−0.19	0.15	0.12	0.06	0.01	0.22
Random effects								
SD (intercept)	0.32	0.06	0.20	0.44	0.10	0.04	0.02	0.17
SD (stress)	0.06	0.04	0.00	0.16	0.03	0.02	0.00	0.08
	Model 3: Age							
Fixed effects								
Intercept	−0.28	0.05	−0.37	−0.18	6.15	0.03	6.10	6.21
Stress	0.03	0.03	−0.04	0.10	−0.04	0.02	−0.08	0.00
Age	0.00	0.01	−0.01	0.01	0.00	0.00	−0.00	0.01
Stress * age	0.00	0.00	−0.00	0.01	0.00	0.00	−0.00	0.01
Random effects								
SD (intercept)	0.32	0.06	0.20	0.45	0.11	0.04	0.03	0.19
SD (stress)	0.06	0.04	0.00	0.16	0.03	0.02	0.00	0.09
	Model 4: BMI							
Fixed effects								
Intercept	−0.27	0.05	−0.36	−0.18	6.16	0.03	6.11	6.21
Stress	0.03	0.04	−0.04	0.10	−0.04	0.02	−0.09	0.01
BMI	0.02	0.01	−0.01	0.04	0.01	0.01	0.00	0.03
Stress * BMI	0.00	0.01	−0.01	0.02	0.00	0.00	−0.01	0.01
Random effects								
SD (intercept)	0.31	0.06	0.18	0.43	0.10	0.04	0.02	0.18
SD (stress)	0.06	0.04	0.00	0.16	0.03	0.02	0.00	0.09

Note: CI = credible interval; LL = lower limit; UL = upper limit.

participants consume on average 534 kcal in time intervals with average stress in which energy intake occurred. Gender has a fixed effect on the (log) energy intake. Through exponentiation, we get the rate decrease in the amount of energy intake associated with the female gender. Hence, women consume on average 16.5% less in time intervals in which energy intake occurred compared with men (e^{-18} = 83.5%). There is also a fixed effect of stress. On average 12.2% less energy is consumed in time intervals in which stress is one-unit higher than usual

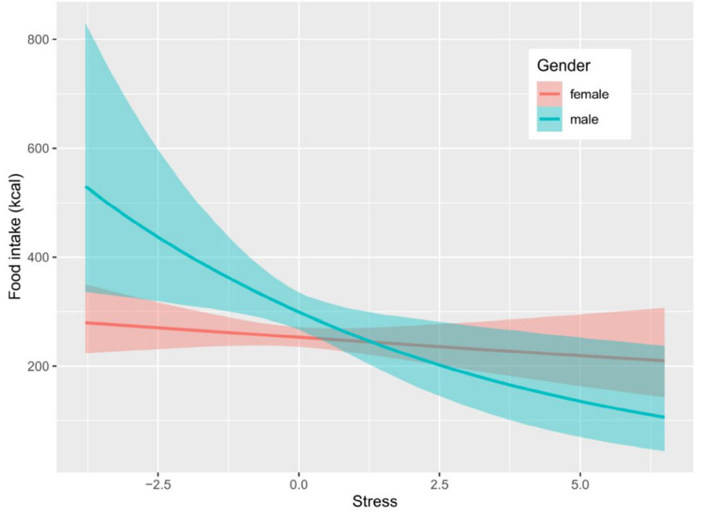


FIGURE 2 Relationship between stress and the amount of food intake moderated by gender. Note: The linear effect of stress on (log) food intake translates to an exponential effect of stress on food intake in the original metric kcal.

(10 points on the original 0 to 100 scale). The cross-level interaction between stress and gender is significant indicating that gender moderates the relationship between stress and the amount of energy intake as illustrated in Figure 2. Increased stress is associated with a decrease in the amount of energy intake in men, whereas no association between stress and the amount of energy intake is observed in women.

Age and BMI did not moderate the stress and eating relationship in either of the two model parts and did not have a statistically meaningful main effect on the probability not to eat. Only BMI, not age, had a small fixed effect on the (log) amount consumed in time intervals in which eating occurs (a one-unit increase in BMI is associated with a 1% increase in the amount of food intake).

Trait stress-eating

Trait stress-eating did not significantly moderate the relationship between stress and the probability not to eat as well as the (log) amount consumed in time intervals in which eating occurred (see Table 4). There was no fixed effect of trait stress-eating in either of the two model parts.

Eating styles

No eating style significantly moderated the relationship between stress and the probability not to eat as well as the (log) amount consumed in time intervals in which eating occurred (see Table 5). The three eating styles had no fixed effects in either of the two model parts.

TABLE 4 Model estimates of the multilevel two-part model of the moderating effect of trait stress-eating

	Zero part				Continuous part			
			95% CI				95% CI	
	Estimate	SE	LL	UL	Estimate	SE	LL	UL
	Model 5: SSES							
Fixed effects								
Intercept	−0.28	0.05	−0.38	−0.19	6.15	0.03	6.09	6.20
Stress	0.03	0.03	−0.04	0.09	−0.04	0.02	−0.08	0.00
SSES	−0.10	0.08	−0.26	0.07	−0.02	0.05	−0.12	0.07
Stress * SSES	−0.00	0.05	−0.11	0.10	0.01	0.04	−0.06	0.09
Random effects								
SD (intercept)	0.32	0.06	0.19	0.44	0.11	0.04	0.04	0.19
SD (stress)	0.06	0.04	0.00	0.16	0.03	0.02	0.00	0.09

Note: CI = credible interval; LL = lower limit; UL = upper limit.

DISCUSSION

Following a novel approach—assessing actual food intake in an EMA setting combined with sophisticated multilevel two-part modelling—the present study assessed (1) whether individuals differ in the stress and eating relationship, (2) whether individual differences in the stress and eating relationship can be explained by person characteristics (i.e., gender, age, BMI, trait stress-eating, and eating styles), and (3) whether these findings support the individual-difference model.

The results of the present study indicate that stress was not related to whether individuals eat. The relationship between stress and the occurrence of eating was not moderated by gender, age, BMI, trait stress-eating, and eating styles. Stress had a significant effect on the amount of food intake in men, but not in women. That is, increased stress was associated with decreased amounts of food intake in men. BMI, age, trait stress-eating, and eating styles did not moderate the relationship between stress and the amount of food intake. Stress had no significant random effect. This indicates that individual differences in the stress and eating relationship were minor.

The present study provides first evidence that stress is not associated with whether individuals eat in daily life. The effect of stress seems to manifest primarily in the amount of food intake. Accordingly, stress may not be related to individuals being more or less likely to eat, whereas in men (but not in women), it may be associated with how much individuals eat when they eat. As this is the first study that differentiates between effects of stress on the occurrence and the amount of food intake, further studies are needed to verify these findings.

Given the large heterogeneity in findings across studies, which assess moderating effects of person characteristics in the stress and eating relationship, it is not surprising that our results are in line with some, but contradict others. For instance, contrary to O'Connor et al. (2008), Wallis and Hetherington (2004), and Wardle et al. (2000), who found that individuals higher in dietary restraint are more likely to increase food intake when experiencing stress, restrained eating did not moderate the stress and eating relationship in the present study as well as in

TABLE 5 Model estimates of the multilevel two-part models of the moderating effect of eating styles

	Zero part				Continuous part			
			95% CI				95% CI	
	Estimate	SE	LL	UL	Estimate	SE	LL	UL
	Model 6: Cognitive restraint							
Fixed effects								
Intercept	−0.28	0.05	−0.37	−0.19	6.15	0.03	6.10	6.20
Stress	0.02	0.03	−0.04	0.09	−0.04	0.02	−0.08	0.00
Restraint	0.02	0.01	−0.00	0.04	−0.00	0.01	−0.01	0.01
Stress * restraint	0.01	0.01	−0.01	0.02	0.00	0.00	−0.01	0.01
Random effects								
SD (intercept)	0.32	0.06	0.19	0.44	0.12	0.04	0.03	0.19
SD (stress)	0.06	0.04	0.00	0.16	0.03	0.02	0.00	0.09
	Model 7: Disinhibition							
Fixed effects								
Intercept	−0.28	0.05	−0.37	−0.18	6.15	0.03	6.10	6.20
Stress	0.03	0.03	−0.04	0.09	−0.04	0.02	−0.08	0.00
Disinhibition	0.00	0.02	−0.03	0.03	−0.01	0.01	−0.02	0.01
Stress * disinhibition	0.00	0.01	−0.02	0.03	0.00	0.01	−0.01	0.02
Random effects								
SD (intercept)	0.33	0.06	0.21	0.45	0.12	0.04	0.03	0.19
SD (stress)	0.06	0.04	0.00	0.16	0.03	0.02	0.00	0.09
	Model 8: Hunger							
Fixed effects								
Intercept	−0.28	0.05	−0.37	−0.18	6.15	0.03	6.10	6.20
Stress	0.03	0.03	−0.04	0.10	−0.03	0.02	−0.08	0.01
Hunger	−0.02	0.02	−0.05	0.01	−0.00	0.01	−0.02	0.01
Stress * hunger	−0.00	0.01	−0.02	0.02	−0.01	0.01	−0.02	0.01
Random effects								
SD (intercept)	0.31	0.06	0.19	0.43	0.11	0.04	0.03	0.19
SD (stress)	0.06	0.04	0.00	0.16	0.03	0.03	0.00	0.09

Note: CI = credible interval; LL = lower limit; UL = upper limit.

previous studies (Conner et al., 1999; Pollard et al., 1995). Furthermore, comparability of findings across studies is low due to differences in study design (e.g., daily diary—O'Connor et al., 2008; within-subject experimental design—Wallis & Hetherington, 2004; quasi-experimental approach—Pollard et al., 1995). Nevertheless, our findings are of importance as they provide novel evidence on the role of gender, age, BMI, trait stress-eating, and eating styles in the stress and eating relationship in daily life.

The results of the present study highlight the need to account for gender differences when studying the stress and eating relationship in daily life. Increased stress was associated with decreases in the amount of food intake in men, while stress was not related to changes in consumed amounts in women. Hence, men's eating behaviour seems to be affected by stress more intensely compared with women. While this is in line with a study by Grunberg and Straub (1992), which showed that men significantly decreased food consumption in the stress condition, it contradicts a study by Conner et al. (1999), which found no gender differences. Our finding could, to some extent, be explained by gender differences in compliance. Systematic noncompliance due to stress is a potential source of bias in EMA studies. Participants might be less likely to report all consumed foods when experiencing stress, whereby it appears as if food intake decreases as a response to stress. This effect might be particularly relevant in male participants as a recent meta-analysis indicates that women are more compliant compared with men, especially in EMA studies with many assessments (Wrzus & Neubauer, 2022). Even though the number of signal-contingent assessments was rather low in the present study, keeping record of food intake in daily life can be highly burdensome. However, to reduce the risk for bias due to systematic noncompliance in food recording when experiencing stress, we rigorously excluded participants and days with poor dietary records.

Reichenberger et al. (2021) found a relationship between stress and food intake (moderated by trait stress-eating) only on the day level, but not on the within-day level. They conclude that stress might have only prolonged or cumulative effects on food intake. A similar explanation was presented by Zenk et al. (2014), who outline that daily hassles might not influence snack food intake in small windows of time during a day, rather when daily hassles accumulate throughout the day. In contrast, the present study found that stress was associated with food intake on the within-day level in men, emphasizing the relevance of short-term effects of stress on food intake. This is in line with laboratory studies that found effects of stress during or shortly after stress-induction (e.g., Epel et al., 2001—within the subsequent 30 min; Grunberg & Straub, 1992—during a 14-min stress-inducing film). More EMA studies that assess actual food intake are needed to specify the time window in which stress affects food intake.

No moderating effect of trait stress-eating on the stress and food intake relationship was identified in the present study. Hence, trait stress-eating may not reliably predict if an individual eats more, less or the same in response to stress in daily life. This questions the ecological validity of self-reported trait stress-eating. Similar questions have been raised in the context of emotional eating (Adriaanse et al., 2011; Bongers & Jansen, 2016). Self-report emotional eating questionnaires seem not to measure what they intend to measure (i.e., increased food intake when experiencing negative emotions) and therefore lack predictive and discriminative validity (Bongers & Jansen, 2016). Further research is needed to assess the ecological validity of trait stress-eating questionnaires.

To our surprise, stress had no significant random effect indicating that individual differences in the stress and eating relationship were minor. This could be due to the fact that participants showed relatively low levels of stress. Only in 195 time intervals (out of 2779) stress was rated above 50 (with the highest score being 100). Again, systematic noncompliance could be a reason for this, as participants might be less likely to respond to prompts when experiencing stress and therefore higher levels of stress might be underrepresented in the data. However, this bias is most likely small in the present study given the high degree of compliance. Another explanation for the lack of individual differences in the dietary response to stress might be intraindividual (i.e., within-person) variability. Individuals might not always show the same dietary response to stress (as outlined below) and therefore intraindividual variability might mask individual differences.

We found fairly strong cross-part correlations in the multilevel two-part models indicating that individuals, who consume on average more energy when they eat, eat less often. Not accounting for this (i.e., running separate models) can cause bias particularly in the continuous part, as higher values of food intake will be underrepresented and smaller values overrepresented (see Ruf, Neubauer, et al., 2021, for a detailed description of this problem). This bias is still present when one is only interested in the continuous part and therefore choses to fit a single model (Su et al., 2009). This highlights the need for multilevel two-part modelling when studying the stress and eating relationship in an EMA setting.

Most research on stress-eating is based on the individual-difference model and thereby on the assumption that the dietary response to stress is stable within an individual (i.e., a trait). Hence, individuals are grouped into different stress-eater types. Even though research has been trying for decades to identify variables that moderate the stress and eating relationship and thereby explain individual differences in the dietary response to stress, no final conclusions can be drawn as findings are highly inconsistent. This poses the question whether stress-eating is as stable as yet assumed. While there is some evidence that the dietary response to stress is rather stable within individuals (Stone & Brownell, 1994), temporal and situational factors (e.g., location, social context, affective, and physical states) that change over short periods of time play an important role in shaping eating behaviour (Dunton, 2018). Instead of trying to understand between-person effects of time-invariant explanatory factors, such as traits and sociodemographic characteristics, on behaviour, there is a need to understand microtemporal processes underlying eating behaviour (Dunton, 2018). Furthermore, Huh et al. (2015) highlight the importance of taking time-varying relationship patterns into account in order to contribute to a deeper understanding of the effects of stress on eating behaviours. First evidence suggests that the stress and eating relationship might be influenced by time-varying factors, such as easy food availability (Zenk et al., 2014). Consequently, it may be time to expand the individual-difference model (Greeno & Wing, 1994) to a dynamic individual-difference model that accounts for dynamic, time-varying factors (external as well as internal) that may moderate the stress and eating relationship (see Figure 3). According to the extended model, time-varying factors might alter an individual's dominant dietary response to stress. Hence, the aim should not only be to identify individuals at greater risk for stress-related changes in food intake but also situations with increased risk for these changes. Research is needed to verify the relevance of the extended model.

Strengths and limitations

The present study has two main strengths: (1) The assessment of actual food intake based on a validated tool in an EMA setting and (2) the data analysis through multilevel two-part modelling, which prevents bias in parameter estimates (as outlined above) and allows new and distinct insights. EMA has great potential to advance the understanding of the stress and eating relationship. It provides more valid and more detailed data about real-world behaviour and experience and sheds light on the dynamics of behaviour in individuals' natural environments (Shiffman et al., 2008). Many authors have expressed the need for studies that assess food intake more accurately, such as energy intake (Araiza & Lobel, 2018; Hill et al., 2021). The present study used a validated EMA tool (Ruf, Koch, et al., 2021) for the assessment of complex dietary

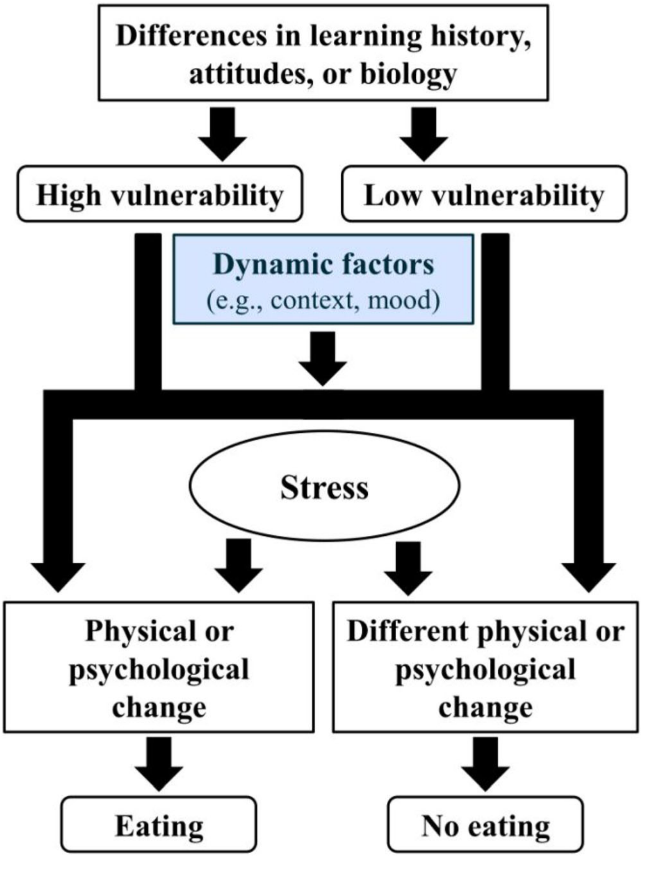


FIGURE 3 Expansion of the individual-difference model of stress-eating by Greeno and Wing (1994) to a dynamic individual-difference model

intake. In doing so, it is the first EMA study investigating the link between stress and actual food intake in daily life in a larger sample. Nevertheless, the food record relied on self-reports, which can cause bias (e.g., systematic noncompliance with reporting foods due to stress as described above). However, using EMA to capture food intake in real time or near real time instead of retrospectively as in traditional dietary assessment methods (e.g., 24-h recalls) seems to provide improved reporting accuracy (Ruf, Koch, et al., 2021).

Asking participants how stressed they were since the last prompt instead of momentarily while having semirandom prompts produced stress measurements for time intervals of different lengths. This is problematic given that it requires participants to average their level of stress over the duration of the time interval. Hence, shorter time intervals reflect more recent measurements of stress compared with longer intervals. For instance, a time interval of 1 h provides a more recent stress measure compared with a 2-h interval. As the recency of the stress assessment likely confounds the effect of stress on subsequent food intake, we decided to assess the effect of stress on concurrent food intake. This limits our findings as temporal associations cannot be established. A further limitation concerns the relatively low level of stress found in the present sample (as outlined above). The findings of the present study may therefore be limited to a restricted range of the stress continuum.

Recommendation for future studies

More studies of high methodological quality assessing intraindividual and interindividual processes in real time are needed to study the relationship between stress and actual food intake in daily life. However, studying stress and actual food intake in daily life can be challenging. Capturing actual food intake through self-reports is prone to bias and burdensome. Yet, only if compliance is high, the link between stress and food intake can be studied meaningfully. Additional to the event-contingent assessment of food intake, asking participants whether all eating occasions have been recorded since the last prompt within the signal-contingent prompts, could reduce the risk of food intake not being reported and thereby further improve the assessment of food intake. Beyond that, advances in the dietary assessment (e.g., passive detection of eating episodes and reliable photo-based dietary assessments) are needed in order to decrease participants' burden, particularly if studies plan to assess food intake over a longer period of time. Furthermore, systematic noncompliance as a response to stress (as outlined above) has to be carefully taken into account. Using passive sensing of physiological stress responses (e.g., heart rate variability) in addition to self-reports can help to circumvent this problem in future studies. Furthermore, individuals are less willing to take part in EMA studies during stressful times (e.g., when work is demanding) resulting in a selective sample. Targeted efforts are needed in order to include individuals into a study when the assessment period is representative.

The distinct findings regarding the occurrence and the amount of food intake as well as the strong cross-part correlations in the present study highlight the importance of multilevel two-part models when examining the stress and eating relationship in daily life. For this reason, future studies should incorporate multilevel-two part modelling (practical guidance on this type of analysis can be found in Ruf, Neubauer, et al., 2021).

CONCLUSION

Individual differences in the dietary response to stress might not be as stable as has been assumed so far. We suggest that the dietary response to stress might not only differ between individuals but also within individuals (i.e., between situations). First evidence indicates that time-varying factors (such as food availability) moderate the stress and eating relationship. For this reason, we propose an expansion of the individual-difference model: a dynamic individual-difference model that accounts for time-varying factors as potential moderators of the stress and eating relationship. Research is needed to verify the extended model.

ACKNOWLEDGEMENTS

We thank Simone Demmel, Hiba El Jomaa, Jeanne Julia Flemming, Hendrik Friedrichsen, Maryam Golzarnia, Krystyna Kotthaus, Emina Ricciardi, Annebirth Steinmann, Söri Wenz, Maike Wisker, and Luise Wortmann for their contribution to data collection of the APPetite study. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ETHICS STATEMENT

The local ethics committee of the faculty of medicine of the Goethe University Frankfurt (Ethikkommission des Fachbereichs Medizin der Goethe-Universität) approved the study (reference number: 192/18). All subjects declared that they understood the study procedure and signed a written informed consent.

DATA AVAILABILITY STATEMENT

The data and R code that support the findings of this study are available in the supporting information of this article.

ORCID

Alea Ruf https://orcid.org/0000-0001-6325-3249
Andreas B. Neubauer https://orcid.org/0000-0003-0515-1126
Elena D. Koch https://orcid.org/0000-0001-8755-4409
Ulrich Ebner-Priemer https://orcid.org/0000-0002-2769-5944
Andreas Reif https://orcid.org/0000-0002-0992-634X
Silke Matura https://orcid.org/0000-0001-7666-9534

REFERENCES

Adriaanse, M. A., de Ridder, D. T. D., & Evers, C. (2011). Emotional eating: Eating when emotional or emotional about eating? *Psychology & Health*, *26*(1), 23–39. https://doi.org/10.1080/08870440903207627

Araiza, A. M., & Lobel, M. (2018). Stress and eating: Definitions, findings, explanations, and implications. *Social and Personality Psychology Compass*, *12*(4), e12378. https://doi.org/10.1111/spc3.12378

Bongers, P., & Jansen, A. (2016). Emotional eating is not what you think it is and emotional eating scales do not measure what you think they measure. *Frontiers in Psychology*, *7*, 1932. https://doi.org/10.3389/fpsyg.2016.01932

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395. https://doi.org/10.32614/RJ-2018-017

Chmitorz, A., Neumann, R. J., Kollmann, B., Ahrens, K. F., Öhlschläger, S., Goldbach, N., Weichert, D., Schick, A., Lutz, B., Plichta, M. M., Fiebach, C. J., Wessa, M., Kalisch, R., Tüscher, O., Lieb, K., & Reif, A. (2021). Longitudinal determination of resilience in humans to identify mechanisms of resilience to modern-life stressors: The longitudinal resilience assessment (LORA) study. *European Archives of Psychiatry and Clinical Neuroscience*, *271*, 1035–1051. https://doi.org/10.1007/s00406-020-01159-2

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, *24*(4), 385–396. https://doi.org/10.2307/2136404

Conner, M., Fitter, M., & Fletcher, W. (1999). Stress and snacking: A diary study of daily hassles and between-meal snacking. *Psychology & Health*, *14*(1), 51–63. https://doi.org/10.1080/08870449908407313

Cotter, E. W., & Kelly, N. R. (2018). Stress-related eating, mindfulness, and obesity. *Health Psychology*, *37*(6), 516–525. https://doi.org/10.1037/hea0000614

Dunton, G. F. (2018). Sustaining health-protective behaviors such as physical activity and healthy eating. *Jama*, *320*(7), 639–640. https://doi.org/10.1001/jama.2018.6621

Epel, E., Jimenez, S., Brownell, K., Stroud, L., Stoney, C., & Niaura, R. (2004). Are stress eaters at risk for the metabolic syndrome? *Annals of the new York Academy of Sciences*, *1032*(1), 208–210. https://doi.org/10.1196/annals.1314.022

Epel, E., Lapidus, R., McEwen, B., & Brownell, K. (2001). Stress may add bite to appetite in women: A laboratory study of stress-induced cortisol and eating behavior. *Psychoneuroendocrinology*, *26*(1), 37–49. https://doi.org/10.1016/S0306-4530(00)00035-4

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

Greeno, C. G., & Wing, R. R. (1994). Stress-induced eating. *Psychological Bulletin*, *115*(3), 444–464. https://doi.org/10.1037/0033-2909.115.3.444

Grunberg, N. E., & Straub, R. O. (1992). The role of gender and taste class in the effects of stress on eating. *Health Psychology*, *11*(2), 97–100. https://doi.org/10.1037/0278-6133.11.2.97

Hill, D., Conner, M., Clancy, F., Moss, R., Wilding, S., Bristow, M., & O'Connor, D. B. (2021). Stress and eating behaviours in healthy adults: A systematic review and meta-analysis. *Health Psychology Review*, *1–25*, 280–304. https://doi.org/10.1080/17437199.2021.1923406

Huh, J., Shiyko, M., Keller, S., Dunton, G., & Schembre, S. M. (2015). The time-varying association between perceived stress and hunger within and between days. *Appetite*, *89*, 145–151. https://doi.org/10.1016/j.appet.2015.02.001

Inada, S., Iizuka, Y., Ohashi, K., Kikuchi, H., Yamamoto, Y., Kadowaki, T., & Yoshiuchi, K. (2019). Preceding psychological factors and calorie intake in patients with type 2 diabetes: Investigation by ecological momentary assessment. *BioPsychoSocial Medicine*, *13*(1), 20. https://doi.org/10.1186/s13030-019-0161-4

Kivimäki, M., & Steptoe, A. (2018). Effects of stress on the development and progression of cardiovascular disease. *Nature Reviews Cardiology*, *15*(4), 215–229. https://doi.org/10.1038/nrcardio.2017.189

Koch, S. A. J., Conrad, J., Hierath, L., Hancock, N., Beer, S., Cade, J. E., & Nöthlings, U. (2020). Adaptation and evaluation of myfood24-Germany: A web-based self-administered 24-h dietary recall for the german adult population. *Nutrients*, *12*(1), 1–15. https://doi.org/10.3390/nu12010160

Liu, L., Ma, J. Z., & Johnson, B. A. (2008). A multi-level two-part random effects model, with application to an alcohol-dependence study. *Statistics in Medicine*, *27*(18), 3528–3539. https://doi.org/10.1002/sim.3205

Meule, A., Reichenberger, J., & Blechert, J. (2018). Development and preliminary validation of the Salzburg Stress Eating Scale. *Appetite*, *120*, 442–448. https://doi.org/10.1016/j.appet.2017.10.003

Micha, R., Shulkin, M. L., Peñalvo, J. L., Khatibzadeh, S., Singh, G. M., Rao, M., Fahimi, S., Powles, J., & Mozaffarian, D. (2017). Etiologic effects and optimal intakes of foods and nutrients for risk of cardiovascular diseases and diabetes: Systematic reviews and meta-analyses from the Nutrition and Chronic Diseases Expert Group (NutriCoDE). *PLoS ONE*, *12*(4), e0175149. https://doi.org/10.1371/journal.pone.0175149

O'Connor, D. B., Jones, F., Conner, M., McMillan, B., & Ferguson, E. (2008). Effects of daily hassles and eating style on eating behavior. *Health Psychology*, *27*(1S), S20–S31. https://doi.org/10.1037/0278-6133.27.1.S20

O'Connor, D. B., Thayer, J. F., & Vedhara, K. (2021). Stress and health: A review of psychobiological processes. *Annual Review of Psychology*, *72*(1), 663–688. https://doi.org/10.1146/annurev-psych-062520-122331

Oliver, G., & Wardle, J. (1999). Perceived effects of stress on food choice. *Physiology & Behavior*, *66*(3), 511–515. https://doi.org/10.1016/S0031-9384(98)00322-9

Pollard, T. M., Steptoe, A., Canaan, L., Davies, G. J., & Wardle, J. (1995). Effects of academic examination stress on eating behavior and blood lipid levels. *International Journal of Behavioral Medicine*, *2*(4), 299–320. https://doi.org/10.1207/s15327558ijbm0204_2

Pudel, V., & Westenhöfer, J. (1989). *Fragebogen zum Eßverhalten*. Hogrefe.

Reichenberger, J., Kuppens, P., Liedlgruber, M., Wilhelm, F. H., Tiefengrabner, M., Ginzinger, S., & Blechert, J. (2018). No haste, more taste: An EMA study of the effects of stress, negative and positive emotions on eating behavior. *Biological Psychology*, *131*, 54–62. https://doi.org/10.1016/J.BIOPSYCHO.2016.09.002

Reichenberger, J., Pannicke, B., Arend, A.-K., Petrowski, K., & Blechert, J. (2021). Does stress eat away at you or make you eat? EMA measures of stress predict day to day food craving and perceived food intake as a function of trait stress-eating. *Psychology & Health*, *36*(2), 129–147. https://doi.org/10.1080/08870446.2020.1781122

RStudio Team. (2020). RStudio: Integrated development environment for R. http://www.rstudio.com/

Ruf, A., Koch, E. D., Ebner-Priemer, U., Knopf, M., Reif, A., & Matura, S. (2021). Studying microtemporal, within-person processes of diet, physical activity, and related factors using the APPetite-Mobile-app: Feasibility, usability, and validation study. *Journal of Medical Internet Research*, *23*(7), e25850. https://doi.org/10.2196/25850

Ruf, A., Neubauer, A. B., Ebner-Priemer, U., Reif, A., & Matura, S. (2021). Studying dietary intake in daily life through multilevel two-part modelling: A novel analytical approach and its practical application.

International Journal of Behavioral Nutrition and Physical Activity, *18*(130), 1–14. https://doi.org/10.1186/s12966-021-01187-8

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

Smyth, J. M., & Stone, A. A. (2003). Ecological momentary assessment research in behavioral medicine. *Journal of Happiness Studies*, *4*(1), 35–52. https://doi.org/10.1023/A:1023657221954

Stan Development Team. (2020). RStan: The R interface to Stan. http://mc-stan.org/

Stone, A. A., & Brownell, K. D. (1994). The stress-eating paradox: Multiple daily measurements in adult males and females. *Psychology & Health*, *9*(6), 425–436. https://doi.org/10.1080/08870449408407469

Stunkard, A. J., & Messick, S. (1985). The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. *Journal of Psychosomatic Research*, *29*(1), 71–83. https://doi.org/10.1016/0022-3999(85)90010-8

Su, L., Tom, B. D. M., & Farewell, V. T. (2009). Bias in 2-part mixed models for longitudinal semicontinuous data. *Biostatistics*, *10*(2), 374–389. https://doi.org/10.1093/biostatistics/kxn044

Tooze, J. A., Grunwald, G. K., & Jones, R. H. (2002). Analysis of repeated measures data with clumping at zero. *Statistical Methods in Medical Research*, *11*(4), 341–355. https://doi.org/10.1191/0962280202sm291ra

Viechtbauer, W. (2022). Statistical methods for ESM data. In I. Myin-Germeys & P. Kuppens (Eds.), *The open handbook of experience sampling methodology: A step-by-step guide to designing, conducting, and analyzing ESM studies* (2nd ed.) (pp. 153–183). Center for Research on Experience Sampling and Ambulatory Methods Leuven.

Wallis, D. J., & Hetherington, M. M. (2004). Stress and eating: The effects of ego-threat and cognitive demand on food intake in restrained and emotional eaters. *Appetite*, *43*(1), 39–46. https://doi.org/10.1016/j.appet.2004.02.001

Wang, L., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods*, *20*(1), 63–83. https://doi.org/10.1037/met0000030

Wardle, J., Steptoe, A., Oliver, G., & Lipsey, Z. (2000). Stress, dietary restraint and food intake. *Journal of Psychosomatic Research*, *48*(2), 195–202. https://doi.org/10.1016/S0022-3999(00)00076-3

Wrzus, C., & Neubauer, A. B. (2022). Ecological momentary assessment: A meta-analysis on designs, samples, and compliance across research fields. *Assessment*, 107319112110675. https://doi.org/10.1177/10731911211067538

Zenk, S. N., Horoi, I., McDonald, A., Corte, C., Riley, B., & Odoms-Young, A. M. (2014). Ecological momentary assessment of environmental and personal factors and snack food intake in African American women. *Appetite*, *83*, 333–341. https://doi.org/10.1016/j.appet.2014.09.008

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ruf, A., Neubauer, A. B., Koch, E. D., Ebner-Priemer, U., Reif, A., & Matura, S. (2022). Individual differences in the dietary response to stress in ecological momentary assessment: Does the individual-difference model need expansion? *Applied Psychology: Health and Well-Being*, 1–21. https://doi.org/10.1111/aphw.12400

Supporting information 1. Open data.

Available at https://doi.org/10.1111/aphw.12400

Supporting information 2. R Code.

Available at https://doi.org/10.1111/aphw.12400