

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Messreaktivität im digitalen Ernährungstracking: Die Rolle von Ernährungszielen
und wahrgenommener Forscherbeobachtung

verfasst von | submitted by

Hannah Förg BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien | Vienna, 2026

Studienkennzahl lt. Studienblatt |
Degree programme code as it appears on the
student record sheet:

UA 066 840

Studienrichtung lt. Studienblatt | Degree
programme as it appears on the student
record sheet:

Masterstudium Psychologie

Betreut von | Supervisor:

Univ.-Prof. Dr. Laura Maria König B.Sc. M.Sc.

Abstract

Messreaktivität beschreibt Veränderungen in Verhalten, Emotionen und Kognitionen, die durch die Teilnahme an einer Studie ausgelöst werden können und dadurch Studienergebnisse verzerren. In der Ernährungsforschung gilt digitales Ernährungstracking als vielversprechende Erhebungsmethode. Gleichzeitig ist bislang unklar, ob und in welchem Ausmaß hierbei messreaktive Effekte auftreten. Ziel dieser Studie war es daher, Messreaktivität im digitalen Ernährungstracking zu untersuchen und damit zu einer besseren Einschätzung der Validität dieser Erhebungsmethode beizutragen. Dabei wurde Messreaktivität im Forschungskontext, isoliert vom Messkontext, untersucht. Weiters soll Aufschluss über mögliche Entstehungsmechanismen gewonnen werden, indem die Rolle von Ernährungszielen, wahrgenommener Forscherbeobachtung und einer Bewusstheit über Verhaltensveränderung untersucht wird. Die Stichprobe bestand aus $N = 86$ Personen, die ihre Ernährung bereits vor Studienteilnahme digital erfassten. Mittels Multi-Level Analysen (Random-Intercept/Random-Slope) wurde untersucht, ob sich die berichtete Kalorienaufnahme zwischen der Woche vor Studienbeginn und den ersten beiden Tagen der Studienteilnahme veränderte. Es zeigte sich kein signifikanter Effekt der Woche auf die Kalorienaufnahme ($b = 45.46$; $t(87.09) = 1.46$, $p = .147$). Somit ergeben sich keine Hinweise auf Messreaktivität in dieser Studie. Darüber hinaus wurden Cross-Level-Modelle berechnet, um Moderationseffekte der Ernährungsziele, der wahrgenommenen Forscherbeobachtung und der Bewusstheit über Verhaltensveränderungen zu testen. Es zeigten sich keine signifikanten Interaktionen der Woche mit Intention ($t(86.24) = 0.63$, $p = .532$), wahrgenommener Forscherbeobachtung Woche ($t(87.06) = 0.39$, $p = .696$) oder Bewusstheit ($t(89.66) = 0.39$, $p = .695$). Damit fanden sich keine Hinweise auf Moderationen. Insgesamt liefert die Studie keine Hinweise auf messreaktive Effekte im digitalen Ernährungstracking unter den vorliegenden Bedingungen. Da Messreaktivität kontextabhängig zu sein scheint, sollten die spezifischen Entstehungsbedingungen für Messreaktivität systematisch untersucht werden.

Abstract

Measurement reactivity describes changes in behavior, emotions, and cognitions due to study participation and may thereby bias study results. In nutrition research, digital dietary self-report is considered a promising method of data collection. At the same time, it remains unclear whether and to what extent measurement reactivity occurs in this context. The aim of this study was therefore to examine measurement reactivity in digital dietary self-report and thereby contribute to a more accurate assessment of the validity of this method.

Measurement reactivity was investigated within the research context, isolated from the measurement context. In addition, potential mechanisms underlying measurement reactivity were explored by examining the role of dietary goals, perceived researcher observation, and awareness of behavioral change. The sample consisted of $N = 86$ individuals who had already been digitally tracking their diet prior to study participation. Multilevel analyses (random intercept/random slope) were used to test whether reported caloric intake differed between the week prior to study start and the first two days of study participation. No significant effect of week on caloric intake was found ($b = 45.46$; $t(87.09) = 1.46$, $p = .147$), indicating no evidence of measurement reactivity in this study. In addition, cross-level models were computed to test moderation effects of dietary goals, perceived researcher observation, and awareness of behavioral change. No significant interactions between week and intention ($t(86.24) = 0.63$, $p = .532$), perceived researcher observation ($t(87.06) = 0.39$, $p = .696$), or awareness ($t(89.66) = 0.39$, $p = .695$) were observed. Thus, there was no evidence of moderation effects. Overall, the study provides no indication of reactive measurement effects in digital dietary self-report under the present conditions. Given that measurement reactivity appears to be context-dependent, the specific conditions under which it emerges should be examined systematically.

Inhaltsverzeichnis

Einleitung	6
Erhebungsmethoden in der Ernährungsforschung	6
Methodische Herausforderungen	7
Digitales Ernährungstracking.....	7
Messreaktivität.....	8
Relevanz einer Differenzierung zwischen Studienaspekten.....	10
Mechanismen der Messreaktivität.....	11
Kognitive Dissonanz.....	11
Soziale Erwünschtheit	13
Bewusstheit über Verhaltensveränderung	15
Ziel der Studie.....	16
Methodik.....	19
Stichprobe	19
Materialien und Maße	19
Stichprobenbeschreibende Variablen	19
Tägliche Kalorieneinnahme kodiert nach Woche	20
Kalorienziel.....	20
Forscherbeobachtung	20
Bewusste Ernährungsveränderung aufgrund der Studienteilnahme	20
Design und Ablauf	20
Datenaufbereitung	21
Deskriptive Statistik	22
Inferenzstatistische Analysen	22
Ergebnisse	24
Deskriptive Statistik der Soziodemographischen Daten	24
Deskriptive Statistik der App-Nutzung.....	26
Deskriptive Statistik der hypothesenspezifischen Daten	27
Hypothese 1: Veränderung der Kalorienaufnahme zwischen den Wochen.....	27
Hypothese 2: Interaktion von Woche und Intention.....	28

Hypothese 3: Interaktion von Woche und wahrgenommener Forscherbeobachtung	29
Explorative Analyse: Interaktion von Woche und Bewusstheit über Ernährungsveränderungen	29
Diskussion	31
Hypothese 1	31
Untersuchtes Verhalten	32
Verwendung des eigenen Messinstruments	32
Moderatoren der Messreaktivität	35
Hypothese 2	36
Hypothese 3	37
Exploration	38
Zusammenfassende Einordnung der Befunde	38
Stärken	39
Limitationen	40
Implikationen	41
Conclusio	42
Literaturverzeichnis	44

Einleitung

Ernährung stellt einen fundamentalen Faktor in der Prävention und Entstehung zahlreicher Krankheiten dar. So kann ungesundes Ernährungsverhalten ein bedeutender Risikofaktor bei der Entwicklung von Krankheiten mit hoher Morbiditäts- und Mortalitätsrate sein, darunter Herz-Kreislauferkrankungen, Diabetes oder bestimmte Krebsarten (Cena & Calder, 2020; Koene et al., 2016). Das Fördern eines gesunden Ernährungsverhaltens zählt daher zu den wichtigen Aufgaben der Gesundheitspsychologie (Sanderman & Morgan, 2025). Um wirksame gesundheitspsychologische Interventionen entwickeln und evaluieren zu können, bedarf es jedoch verlässlicher empirischer Grundlagen im Bereich des Ernährungsverhaltens (König et al., 2022; Naska et al., 2017). Eine vielversprechende und zunehmend eingesetzte Methode zur Ernährungserhebung stellt das Smartphone-basierte Ernährungstracking dar (König et al., 2022). In diesem können Studienteilnehmende zum Zeitpunkt des Konsums in einer App eintragen, was sie zu sich nehmen (König et al., 2022). Während diese Methode zwar einige Schwächen analoger Erhebungsmethoden adressiert, besteht dennoch die Gefahr, dass Daten allein durch die Studienteilnahme verzerrt werden (French & Sutton, 2010; Illner et al., 2012; König et al., 2022; Rollo et al., 2016; Sharp & Allman-Farinelli, 2014). Dieser Effekt wird als Messreaktivität bezeichnet und beschreibt Veränderungen in den Studienteilnehmenden durch die Studienteilnahme selbst (French & Sutton, 2010). Um eine umfassende Einschätzung der Validität der Ernährungserhebung durch digitales Ernährungstracking gewährleisten zu können, besteht daher die Notwendigkeit mögliche Messreaktivitätseffekte zu erforschen. Daraus ergibt sich das Ziel der vorliegenden Studie. Konkreter soll der Effekt von Messreaktivität im Forschungskontext, abseits des Messkontexts, untersucht werden, sowie die Rolle selbstgesetzter Ernährungsziele und der wahrgenommenen Forscherbeobachtung. Weiters wird untersucht, ob mögliche Verhaltensveränderungen durch Messreaktivität bewusst geschehen.

Erhebungsmethoden in der Ernährungsforschung

In der ernährungswissenschaftlichen Forschung kommen häufig Selbstberichtsinstrumente zum Einsatz (Burrows et al., 2019; Naska et al., 2017; Thompson & Subar, 2017). Diese können einerseits erinnerungsbasiert sein, wie 24-Stunden-Ernährungsprotokolle (24-hour dietary recall) und Lebensmittel-Häufigkeits-Fragebögen (Food Frequency Questionnaire), oder momentbasiert wie Ernährungstagebücher (Self-reported dietary record), bei denen zum Zeitpunkt des Konsums das zu sich Genommene dokumentiert wird (Naska et al., 2017). Diese Methoden gelten als praktikabel und ökonomisch, sind jedoch anfällig für verschiedene Verzerrungen (Burrows et al., 2019; Naska et al., 2017; Thompson & Subar, 2017). So wurde im Vergleich zur Energieaufnahmeerfassung mittels des Biomarkers Doubly Labeled Water, welcher eine objektive Messung der Energiezufuhr ermöglicht, in Selbstberichten eine Kalorienuntererfassung von 4.6% bis 42% berichtet (Burrows et al., 2019).

Methodische Herausforderungen

Dies lässt sich unter anderem auf methodisch bedingte Fehlerquellen zurückführen, zum Beispiel durch Erinnerungsfehler, einen hohen Dokumentationsaufwand oder Fehleinschätzungen der Nährwerte (Naska et al., 2017; Thompson & Subar, 2017). So müssen sich beim 24-Stunden-Ernährungsprotokoll Teilnehmende etwa daran erinnern, was sie am Vortag konsumiert haben, während Lebensmittel-Häufigkeits-Fragebögen das Essverhalten über einen länger vergangenen Zeitraum wie ein paar Monate abfragen. Dadurch sind diese beiden Methoden zwar ökonomisch, jedoch auch anfällig für Erinnerungsfehler (Naska et al., 2017; Thompson & Subar, 2017).

Ernährungstagebücher vermeiden dies, indem Teilnehmende über mehrere Tage hinweg zum Zeitpunkt des Konsums eintragen, was und wie viel zu sich genommen wird. Dadurch entsteht allerdings ein hoher Aufwand für die Teilnehmenden (Naska et al., 2017; Thompson & Subar, 2017). Das kann einerseits zu einem inkonsistenten Aufzeichnen führen. Andererseits können Unterschiede zwischen Personen die Teilnahme an der Studie bzw. eine konsistente Dokumentation wahrscheinlicher oder unwahrscheinlicher machen. Dies kann systematische Verzerrungen zur Folge haben, welche die Generalisierbarkeit der Ergebnisse einschränken (Burrows et al., 2019; Thompson & Subar, 2017). Außerdem können Fehler in der Einschätzung der Nährwerte entstehen, beispielsweise durch eine Fehleinschätzung der Portionsgröße oder der beinhaltenen Nährstoffe. Dem kann zwar durch Schulungen für Teilnehmende oder durch Einschätzungen von Experten entgegengewirkt werden, beides bedeutet jedoch einen erhöhten Aufwand sowie Kosten (Thompson & Subar, 2017).

Digitales Ernährungstracking

Eine digitale Anwendung von Ernährungstagebüchern durch Smartphone-basierte Apps, im Weiteren auch als digitales Ernährungstracking bezeichnet, bietet hier eine vielversprechende Erweiterung (Illner et al., 2012; Rollo et al., 2016). Zusätzlich zu den Vorteilen, die klassische Ernährungstagebücher im Vergleich zu erinnerungsbasierten Methoden mit sich bringen, reduziert eine digitale Anwendung den Aufwand für Teilnehmende sowie Forschende, und ermöglicht Hilfestellungen für eine genauere Nährwerteinschätzung (Rollo et al., 2016; Thompson & Subar, 2017). Da Smartphones meist ständig mitgeführt werden, ist es möglich, Mahlzeiten direkt zum Zeitpunkt des Konsums zu dokumentieren (König et al., 2022; Rollo et al., 2016). Das Verwenden digitaler Methoden erleichtert die Dokumentation, beispielsweise durch Lebensmitteldatenbanken oder der Verwendungsmöglichkeit von Fotos (Illner et al., 2012; Rollo et al., 2016; Thompson & Subar, 2017). In dem die Auswertung der Daten durch Software erfolgen kann, wird auch der Auswertungsaufwand für die Forschenden reduziert (Burrows et al., 2019; Thompson & Subar, 2017). Darüber hinaus erlauben digitale Funktionen wie Lebensmitteldatenbanken, Barcode-Scanner oder Referenzbilder von Portionsgrößen, eine erleichterte Einschätzung

der tatsächlichen Energiezufuhr und kann somit die Genauigkeit der Angaben erhöhen (Illner et al., 2012; Thompson & Subar, 2017; Turner-McGrievy et al., 2019). Dadurch stellt digitales Ernährungstracking eine moderne und praktikable Methode zur Ernährungserhebung dar, die einige Schwächen klassischer Selbstberichte adressiert (Illner et al., 2012; Sharp & Allman-Farinelli, 2014).

Messreaktivität

Verzerrungen in der Verhaltensforschung, und somit auch der Ernährungsforschung, können nicht nur durch Schwächen der Erhebungsmethode selbst entstehen, sondern auch durch die Tatsache, dass Personen an einer wissenschaftlichen Untersuchung teilnehmen (Arigo & König, 2024; French & Sutton, 2010; König et al., 2022; McCambridge et al., 2014). Dieses Phänomen wird unter dem Begriff Messreaktivität gefasst und beschreibt Veränderungen in der untersuchten Person, die allein durch die Teilnahme an einer Studie ausgelöst werden (French & Sutton, 2010). Messreaktivität kann somit dazu führen, dass das durch die Studie erhobene Verhalten nicht das tatsächliche Verhalten widerspiegelt (French & Sutton, 2010).

Eine unzureichende Berücksichtigung dieses Effekts kann die Validität von Forschung durch potenziell verzerrte Schlussfolgerungen von Studienergebnisse gefährden. In der Praxis könnte so beispielsweise ein beobachteter Effekt fälschlicherweise auf eine wirksame Intervention zurückgeführt werden, obwohl er tatsächlich durch Messreaktivität entstand. Dies könnte zu einer Fehleinschätzung der Wirksamkeit gesundheitsfördernder Interventionen oder Maßnahmen führen und in weiterer Folge dazu beitragen, dass Interventionen auf Basis verzerrter Evidenz entwickelt und implementiert werden (French & Sutton, 2010; Miles et al., 2020). Um Forschungsdaten zuverlässig interpretieren zu können, ist eine präzise Einschätzung möglicher Messreaktivitätseffekte daher zentral.

In der digitalen Gesundheitsforschung wurde Messreaktivität hauptsächlich im Bereich der Bewegungsforschung untersucht (König et al., 2022). Befunde aus einer Metaanalyse von König et al. (2022) zeigten hier einen kleinen bis mittleren Effekt (Cohen's $d = 0.27-0.30$). Konkret zeigte sich, dass Teilnehmende in den ersten ein bis zwei Tagen nach Beginn einer Messung mehr Schritte gingen als zuvor, obwohl keine zusätzliche Intervention stattfand (Arigo & König, 2024; French & Sutton, 2010; König et al., 2022). Dies weist darauf hin, dass allein das Teilnehmen an einer Studie zu gesundheitsförderlicherem Verhalten beitragen kann (French & Sutton, 2010; König et al., 2022). Die bisherige Studienlage ist insgesamt jedoch von hoher Heterogenität geprägt und Befunde aus anderen Gesundheitsbereichen liegen bisher kaum vor (French & Sutton, 2010; König et al., 2022).

So ist auch Messreaktivität in der digitalen Ernährungsforschung noch weitgehend unerforscht (König et al., 2022). Vor dem Hintergrund, dass digitales Ernährungstracking in der aktuellen Forschung und Praxis zunehmend eingesetzt wird und viele methodische Schwächen klassischer Selbstberichtsinstrumente adressieren kann (König et al., 2022; Rollo

et al., 2016; Sharp & Allman-Farinelli, 2014; Thompson & Subar, 2017), ergibt sich die Notwendigkeit, die Validität dieser Methodik auch unter dem Aspekt potenzieller Messreaktivität zu untersuchen.

Begrifflichkeiten und konzeptuelle Differenzierung von Messreaktivitätskontexten

Literatur und Studien zu Messreaktivität sind nicht nur von einer heterogenen Befundlage geprägt, sondern auch von einer undifferenzierten und inkonsistenten Verwendung von Begrifflichkeiten und Kontexten in denen Messreaktivität betrachtet wird (French & Sutton, 2010; McCambridge et al., 2014; Miles et al., 2020). Dies erschwert zuverlässige Schlussfolgerungen über das Ausmaß und Kontexte in denen Messreaktivität auftritt (McCambridge et al., 2014).

Hawthorne Effekt. Historisch betrachtet findet das Konzept der Messreaktivität seinen Ursprung in den sogenannten Hawthorne-Studien. In diesen zeigte eine Untergruppe von Arbeitenden, die im Rahmen eines Forschungsprogramms intensiv durch Vorgesetzte betreut wurden, eine deutliche Steigerung ihrer Produktivität (McCambridge et al., 2014). Auch wenn der Hawthorne Effekt ursprünglich aus der industriellen Forschung stammt, wurde der Begriff inzwischen auch auf andere Forschungsbereiche ausgeweitet (McCambridge et al., 2014). Typischerweise bezeichnet der Hawthorne Effekt eine Verhaltensänderung, die durch das Bewusstsein entsteht, Teil einer wissenschaftlichen Untersuchung zu sein. Der Fokus der Definition liegt also vor allem auf dem Bewusstsein erforscht zu werden. Häufig wird er allerdings auch als Synonym für jede Form von Verhaltensänderung im Rahmen einer Untersuchung gebraucht, ohne klare Abgrenzung und einheitlicher Verwendung (McCambridge et al., 2014). Eine Metaanalyse von McCambridge et al. (2014), mit dem Ziel einen Überblick über den aktuellen Forschungsstand des Hawthorne Effekts zu schaffen, zeigte, dass die bloße Teilnahme an einer Studie tatsächlich Verhaltensveränderungen hervorrufen kann. Gleichzeitig wird jedoch betont, dass verlässliche Aussagen zur Größe und zu den spezifischen Bedingungen dieses Effekts kaum möglich sind, da der Begriff uneinheitlich definiert und operationalisiert wird (McCambridge et al., 2014).

Während es beim Hawthorne Effekt typischerweise um Verhaltensveränderung durch das Bewusstsein erforscht zu werden geht, schließt der Begriff der Messreaktivität zusätzliche Aspekte ein (French & Sutton, 2010; McCambridge et al., 2014). So beschreibt Messreaktivität einerseits Veränderungen nicht nur im Verhalten, sondern auch in Kognitionen und Emotionen. Weiters werden Veränderungen miteinbezogen, die durch alle Aspekte der Teilnahme an einer Studie ausgelöst werden. Dies schließt somit auch andere Aspekte der Studienteilnahme ein, die neben dem Bewusstsein erforscht zu werden aufkommen können, wie beispielsweise Befragungen (French & Sutton, 2010; Miles et al., 2020).

Frage-Verhalten-Effekt. Somit ergibt sich auch ein Unterkonzept der Messreaktivität: der sogenannte Frage-Verhalten-Effekt (Question-Behavior Effect, QBE) (Miles et al., 2020). Dieser beschreibt Verhaltensänderungen, die durch den Befragungsaspekt einer Studie ausgelöst werden, also allein durch das Stellen von Fragen zu einem bestimmten Verhalten, wie es oft anhand von Fragebögen stattfindet (Miles et al., 2020; Sprott et al., 2006). Zwei theoretische Strömungen, unter dem Begriff des Mere-Measurement-Effekts und des Self-Prophecy-Effekts, fließen in diesen Begriff ein, wobei erstere stärker im Bereich des Konsumverhaltens, letztere im sozialen Kontext verortet ist. Beiden Ansätzen ist die Annahme gemeinsam, dass das Auseinandersetzen mit Fragen zu einem Verhalten dieses Verhalten selbst beeinflussen kann und werden somit als sogenannter Frage-Verhalten-Effekt zusammengefasst (Dholakia & Malhotra, 2010; Sprott et al., 2006).

Eine Vielzahl an Befunden zeigten kleine, aber signifikante Frage-Verhalten-Effekte in Bezug auf unterschiedliche selbstberichtete Verhaltensweisen (McCambridge & Kypri, 2011; Miles et al., 2020; Wilding et al., 2016; Wood, 2016). Eine speziell auf den Gesundheitsbereich fokussierte Metaanalyse von Miles et al. (2020) lieferte Evidenz dafür, dass diese Effekte auch bei gesundheitsbezogenen Selbstberichten auftreten. Die Ergebnisse zeigten jedoch auch hier große Heterogenität hinsichtlich der Effektgröße und der Kontexte, in denen der Frage-Verhalten-Effekt auftritt (Miles et al., 2020).

Messkontext. Neben den schon genannten Aspekten der Studienteilnahme, der Bewusstheit erforscht zu werden und der Befragung, kann auch das Gemessen werden im Rahmen einer Studie eine Rolle in der Entstehung der Messreaktivität spielen (Barta et al., 2012; König et al., 2022). Dieser Effekt ist aus der Interventionsforschung bekannt, wo das regelmäßige Aufzeichnen von Verhalten im Sinne des Selbstmonitorings gezielt zur Verhaltensänderung eingesetzt wird, unter anderem in Bezug auf Ernährungsverhalten und Gewichtsabnahme (Burke et al., 2011; Teasdale et al., 2018). Dadurch liegt nahe, dass dieser Effekt auch unerwünschterweise durch Forschungsmethoden ausgelöst werden kann, die das Aufzeichnen von Verhalten als Erhebungsmethode verwenden, wie es bei Tagebuchstudien oder Ecological Momentary Assessment (EMA) der Fall ist (Barta et al., 2012; König et al., 2022; Maher et al., 2024).

Aktiv eingesetzt als Intervention ist dieser Effekt zielführend (Burke et al., 2011; Teasdale et al., 2018). Wenn diese Verhaltensänderung jedoch durch die Messung als Mittel zur Verhaltenserhebung entsteht, ist dieser Effekt problematisch, da er das Erfassen des tatsächlichen Verhaltens verzerrt (French & Sutton, 2010; König et al., 2022).

Relevanz einer Differenzierung zwischen Studienaspekten

Insgesamt zeigt sich, dass verschiedene Aspekte der Studienteilnahme im Hinblick auf Messreaktivität relevant sind. Angelehnt an König et al. (2025) werden in dieser Arbeit zwei Kontexte für die Entstehung von Messreaktivität voneinander abgegrenzt: der Forschungskontext und der Messkontext. Der Forschungskontext umfasst die Aspekte, die

durch die Erforschung entstehen, dazu gehört unter anderem das Bewusstsein, erforscht zu werden, sowie mögliche Befragungen im Rahmen der Studie. Der Messaspekt hingegen umfasst das Gemessen werden, also beispielsweise das regelmäßige Aufzeichnen von Verhalten, durch ein neues Gerät wie einen Schrittzähler oder im Fall des digitalen Ernährungstracking durch die App (König et al., 2025).

In der Forschung wird jedoch bislang kaum klar zwischen diesen Aspekten unterschieden (König et al., 2025). Gerade in Studien mit regelmäßiger Verhaltensaufzeichnung als Erhebungsmethode, wie es im digitalen Ernährungstracking der Fall ist, fällt es jedoch schwer, den Effekt dieser Messung von dem des Forscherkontexts zu trennen, da beide gleichzeitig mit der Studienteilnahme eingeführt werden (König et al., 2025). Eine explizite Differenzierung findet bisher lediglich in der Studie von König et al. (2025) statt. Wissen über den Messreaktivitätseffekt in den einzelnen Kontexten, kann jedoch Aufschluss darüber geben wie und warum Messreaktivität entsteht und dementsprechend bei der Entwicklung gezielter Maßnahmen und Empfehlungen zur Minimierung von Messreaktivität verhelfen (König et al., 2025). Vor diesem Hintergrund liegt der Fokus dieser Arbeit bewusst auf dem Effekt der Messreaktivität im Forschungskontext, klar abgegrenzt vom Messkontext. Dies wird konkret umgesetzt, indem die Stichprobe nur aus Personen besteht, die schon vor der Studienteilnahme ihre Ernährung digital aufzeichneten. Dadurch ist es nicht erforderlich, ein neues Messinstrument im Zuge der Studienteilnahme einzuführen. Auf diese Weise entsteht kein zusätzlicher Messkontext und der Forschungskontext kann gezielt isoliert untersucht werden. Ziel ist es, somit zu einem differenzierteren Verständnis der Messreaktivität in der digitalen Ernährungsforschung beizutragen.

Mechanismen der Messreaktivität

Zu einem differenzierteren Verständnis der Messreaktivität trägt auch das Untersuchen zugrundeliegender Mechanismen der Messreaktivität bei. Das Verständnis dieser beruht zu einem Großteil auf theoretischen Vermutungen (Barta et al., 2012; Wilding et al., 2016; Wood, 2016). Das Untersuchen von Moderatoren im Kontext des Messreaktivitätseffekts kann empirischen Aufschluss über mögliche Wirkmechanismen geben und stellt somit einen zentralen Fokus dieser Arbeit dar.

Kognitive Dissonanz

In der Literatur wird das Phänomen der Messreaktivität oft auf den Mechanismus der Kognitiven Dissonanz zurückgeführt (Conner et al., 2017; Dholakia & Malhotra, 2010). Es wird angenommen, dass die Teilnahme an einer Studie die Wahrnehmung des eigenen Verhaltens bezüglich des untersuchten Themas beeinflussen kann und somit Selbstregulationsprozesse fördert (Barta et al., 2012; Conner et al., 2017; MacNeill et al., 2016; Moos, 2008). Konkreter kann die Studienteilnahme zu einer verstärkten Wahrnehmung möglicher Diskrepanzen zwischen tatsächlichem Verhalten und Vorstellungen über

Idealverhalten führen (Barta et al., 2012; Conner et al., 2017; Dholakia & Malhotra, 2010; Moos, 2008). Durch diese Diskrepanz wird ein Spannungszustand erzeugt, welcher als unangenehm empfunden wird und daher verringert werden soll (Barta et al., 2012; Conner et al., 2017; Dholakia & Malhotra, 2010; Freijy & Kothe, 2013; Moos, 2008). Dieser Spannungszustand kann auch als kognitive Dissonanz, definiert als Spannungszustand zwischen widersprüchlichen Kognitionen und/oder Verhaltensweisen, bezeichnet werden (Conner et al., 2017; Dholakia & Malhotra, 2010; Wilding et al., 2016; Wood, 2016). Die kognitive Dissonanz kann durch eine Änderung des Verhaltens in Richtung Zielverhalten reguliert werden, also durch einen selbstregulatorischen Prozess (Barta et al., 2012; Carver & Scheier, 1982; Conner et al., 2017; Freijy & Kothe, 2013; Moos, 2008; Stice et al., 2008). Die Wirksamkeit dieser Selbstregulation für Verhaltensveränderung lässt sich durch Befunde aus der Interventionsforschung belegen, in der das Erzeugen von Dissonanzen aktiv zur Anwendung kommt (Freijy & Kothe, 2013; Stice et al., 2008). So zeigte beispielsweise eine Metastudie, dass das gezielte Hervorrufen kognitiver Dissonanz durch eine Konfrontation mit Diskrepanzen zwischen dem eigenen Verhalten und gesundheitsbezogenem Zielverhalten zu positiven Veränderungen im Gesundheitsverhalten führen kann (Freijy & Kothe, 2013).

In Anbetracht dessen liegt nahe, dass dieser Mechanismus der kognitiven Dissonanz auch unerwünscht im Studienkontext auftreten kann (Barta et al., 2012; Conner et al., 2017; Wilding et al., 2016). Dieser Erklärungsansatz wird weiters gestützt durch Studien, die zeigen, dass Personen, die eine starke Notwendigkeit verspüren ihr Gesundheitsverhalten zu verbessern, also deren tatsächliches Verhalten von Wissen über ihr ideales Verhalten abweicht, verstärkt Messreaktivität zeigen. Dies konnte in Studien an Personen mit problematischem Alkoholkonsum und Personen mit erhöhtem Herz-Kreislauf-Risiko gezeigt werden (Arigo & König, 2024; Poulton et al., 2019). Diese Ergebnisse stützen die Annahme, dass das Erleben einer Diskrepanz zwischen tatsächlichem Verhalten und Zielverhalten, sowie die daraus resultierende kognitive Dissonanz und Selbstregulationsprozesse, eine zentrale Rolle in der Entstehung von Messreaktivität spielen könnten.

Allerdings werden diese Selbstregulationsprozesse vor allem im Kontext von Studien diskutiert, in denen Verhalten gemessen wird, indem es aufgezeichnet wird, wie Tagebuchstudien oder EMA (Barta et al., 2012; Maher et al., 2024; Poulton et al., 2019), also im Rahmen des Messkontexts einer Studie. Hier wird davon ausgegangen, dass durch das Aufzeichnen bzw. den Zugang zu den aufgezeichneten Daten, die Wahrnehmung auf das eigene Verhalten und mögliche Diskrepanzen erfolgt (Barta et al., 2012; König et al., 2022; Moos, 2008). In der Ernährungsforschung zeigt sich das insofern, dass Aufzeichnen von Essen selbst, Ernährungsentscheidungen beeinflussen kann (Thompson & Subar, 2017).

Ob vergleichbare Mechanismen jedoch auch isoliert im Forschungskontext eine Rolle spielen, abseits des Messungsaspekts, ist jedoch unklar (Wood, 2016). In der Literatur über

den Frage-Verhalten-Effekt, wird oft diskutiert, dass bereits die Auseinandersetzung mit einer Frage über das untersuchte Verhaltensthema zu einer Diskrepanzwahrnehmung und somit zu kognitiver Dissonanz führen kann, welche wiederum Verhaltensänderungen auslösen kann (Conner et al., 2017; Dholakia & Malhotra, 2010; Spangenberg et al., 2003; Wilding et al., 2016; Wood, 2016). Dies stützt eine Studie, in der Personen, die durch eine Befragung kognitive Dissonanz erlebten weniger Verhaltensveränderung zeigten, wenn diese ihre Dissonanz durch Abwärtsvergleich auflösen konnte und somit nicht ihr Verhalten ändern mussten, um die Dissonanz aufzulösen (Spangenberg et al., 2003). Allerdings ist die empirische Evidenz hierfür heterogen und die Möglichkeit eindeutige Schlüsse über Wirkmechanismen zu ziehen begrenzt (Wood, 2016). Eine Metaanalyse von Wood (2016) kommt zu dem Schluss, dass die Befundlage nicht ausreicht, um Dissonanz- oder Selbstregulationsmechanismen allein durch Befragung als sicher anzunehmen. Daher ergibt sich die Frage, ob der Forschungskontext einer Studie, ohne Messaspekt, tatsächlich messreaktive Effekte über vergleichbare Mechanismen auslösen kann.

Falls der Forschungskontext einer Studienteilnahme durch eine verstärkte Wahrnehmung von Diskrepanzen zwischen tatsächlichem Verhalten und Vorstellungen über Idealverhalten zu Verhaltensveränderung führt (Conner et al., 2017; Spangenberg et al., 2003; Wilding et al., 2016; Wood, 2016), könnte dies im Kontext der Ernährungsforschung folgendes bedeuten: Bei Personen, die angeben, ihre Kalorienzufuhr stark verändern zu wollen, und somit weit weg von ihrer Vorstellung des Idealverhalten sind, könnte die Studienteilnahme zu einem verstärkten Fokus auf diese Diskrepanz führen. Im Gegensatz dazu würden Personen, die ihre Kalorienzufuhr nicht verändern wollen, also sich gemäß ihrer Vorstellung des Ideals verhalten, keine Diskrepanz verspüren. Dementsprechend würden bei Personen mit hohem Ernährungsziel verstärkt selbstregulatorische Prozesse aktiviert werden, um die wahrgenommene Diskrepanz zu regulieren. Dadurch könnten Personen mit stark ausgeprägtem Kalorienziel ihr Verhalten eher an ihr Zielverhalten anpassen als Personen, die kein Veränderungsziel äußern. Die Erforschung der Rolle persönlicher Ernährungsziele in der Entstehung von Messreaktivität bietet somit die Möglichkeit, Hinweise bezüglich möglicher Diskrepanz- und Selbstregulationsprozesse in der Entstehung der Messreaktivität durch den Studienkontext zu gewinnen. Dementsprechend untersucht die vorliegende Studie, ob persönliche Ernährungsziele die Entstehung von Messreaktivität moderieren.

Soziale Erwünschtheit

Neben dem eben dargestellten Mechanismus der kognitiven Dissonanz wird in der Literatur auch der sozialen Erwünschtheit durch die Bewusstheit der Forscherbeobachtung eine Rolle in der Entstehung von Messreaktivität zugeschrieben (Barta et al., 2012; McCambridge et al., 2014; Wilde & Garvin, 2007; Wood, 2016). Soziale Erwünschtheit (Social Desirability) beschreibt eine Veränderung im Verhalten oder Antworten in Richtung

sozialer Normen. Dabei lassen sich zwei zentrale Formen unterscheiden: Selbsttäuschung (Self-Deception), als die unbewusste Selbsttäuschung im Sinne eines verzerrt positiven Selbstbildes, und Selbstdarstellung (Impression Management), als die bewusste Kontrolle des Verhaltens mit dem Ziel, von anderen positiv wahrgenommen zu werden (Börger, 2012). Selbsttäuschung tritt verstärkt bei retrospektiven Selbstberichtsinstrumenten wie dem 24-Stunden-Ernährungsprotokoll oder Lebensmittel-Häufigkeits-Fragebögen auf, bei denen Teilnehmende sich an vergangenes Verhalten erinnern und dabei möglicherweise ein geschöntes Bild ihrer Ernährung abrufen (Barta et al., 2012).

Da im digitalen Ernährungstracking nicht retrospektiv berichtet wird (König et al., 2022), liegt der Fokus dieser Arbeit insbesondere auf der Selbstdarstellung (Barta et al., 2012). Diese Form der sozialen Erwünschtheit spielt insbesondere in sozialen Situationen eine Rolle, also dort, wo Personen sich beobachtet und beurteilt fühlen bzw. Erwartungen verspüren (Barta et al., 2012). Bei einer Studienteilnahme geschieht dies durch die Forscherbeobachtung, also dadurch, dass das untersuchte Verhalten und Daten von den Forschenden beobachtet, dokumentiert und ausgewertet wird (Barta et al., 2012; McCambridge et al., 2014). Selbstdarstellung im Forschungskontext bedeutet also, dass sich Teilnehmende so präsentieren, dass sie die angenommenen Erwartungen und sozialen Normen der Forschenden erfüllen (Barta et al., 2012; McCambridge et al., 2014). Im Gesundheitsbereich äußern sich Verhaltensveränderungen durch Messreaktivität besonders in die Richtung eines gesundheitsförderlichen Verhaltens, wie mehr Bewegung oder geringerer Alkoholkonsum (French & Sutton, 2010; König et al., 2022). Die Annahme, dass soziale Erwünschtheit eine Rolle in der Entstehung von Messreaktivität spielt, wird durch Ergebnisse einer Metaanalyse unterstützt, welche zeigte, dass Befragung in einem Studienkontext die größte Verhaltensänderung bewirkte, wenn das Verhalten sozial erwünscht war (Wood, 2016).

Wenn soziale Erwünschtheit auch in der digitalen Gesundheitsforschung eine Rolle in der Messreaktivität spielt, sollte das Ausmaß der wahrgenommenen Forscherbeobachtung, durch die das Bedürfnis sich erwartungs- und normkonform zu verhalten ausgelöst wird, ebenfalls eine Rolle spielen (Barta et al., 2012; McCambridge et al., 2014). Dies konnte jedoch in der digitalen Bewegungsforschung über den Zusammenhang zwischen wahrgenommener Forscherbeobachtung und Messreaktivität nicht direkt beobachtet werden. So zeigten Befunde einer Studie, in der die Salienz der Forscherbeobachtung variiert wurde, keine Unterschiede im Effekt der Messreaktivität. Jedoch wurde nicht geprüft, ob die Salienzmanipulation erfolgreich war. Daher könnte der fehlende Effekt auch darauf zurückzuführen sein, dass die Manipulation nicht wirksam war (König et al., 2025).

Befunde aus der Ernährungsforschung im Laborsetting, geben wiederum Anlass zur Annahme, dass Forscherbeobachtung durch soziale Erwünschtheit Messreaktivität auslösen

kann, konkreter zu kalorienärmere Ernährung führen kann (Robinson et al., 2014, 2015). So gaben Personen in einer qualitativen Studie an, in einem Laborsetting aufgrund der Forscherbeobachtung, aktiv weniger zu essen (Robinson et al., 2014). Dies wird weiter durch Ergebnisse einer Metaanalyse gestützt, welche zeigte, dass sich Personen unter ausgeprägter Beobachtung kalorienärmer ernähren (Robinson et al., 2015).

Allerdings untersuchte diese Metaanalyse ausschließlich Essverhalten in einem Laborsetting (Robinson et al., 2015). Es wird angenommen, dass selbst-administrierte Methoden wie digitales Ernährungstracking weniger Verzerrung durch soziale Erwünschtheit mit sich bringen, verglichen zu Kontexten, in denen man direkten Kontakt mit den Forschenden hat, wie in Laborsettings (Barta et al., 2012). Eine direkte Übertragung der Ergebnisse aus dem Laborsetting auf digitales Ernährungstracking ist daher nicht möglich, dennoch wird dadurch die Rolle von sozialer Erwünschtheit in der Ernährungsforschung verdeutlicht.

Verstärkt wird diese Annahme durch Befunde im Kontext von Ernährungs-Selbstberichten. Hier lässt sich beobachten, dass insbesondere Kalorien stärker unterberichtet wurden als Protein (Schatzkin et al., 2003). Ebenso zeigte sich zwar ein systematisches Unterberichten in allen Nahrungsmittelgruppen, jedoch verstärkt in ungesünderen (Subar et al., 2003). Beide Befunde interpretieren Naska et al. (2017) als ein bewusstes Unterberichten von ungesundem Essverhalten, welches als eine Art Selbstdarstellung zu sehen ist. Dies legt nahe, dass Selbstdarstellung auch in weniger kontrollierten, selbstadministrierten Kontexten, wie beim digitalen Ernährungstracking, relevant sein könnte (Naska et al., 2017).

Vor dem Hintergrund dieser Befunde und der fehlenden Forschung im Bereich des digitalen Ernährungstracking, ergibt sich der Bedarf die Rolle sozialer Erwünschtheit in diesem Kontext zu erforschen. Die Theorie, dass Forscherbeobachtung durch den Mechanismus der sozialen Erwünschtheit zu Messreaktivität führt (Barta et al., 2012; McCambridge et al., 2014), legt nahe, dass Messreaktivität abhängig von der wahrgenommenen Forscherbeobachtung sein könnte. Somit kann die Untersuchung einer möglichen moderierenden Rolle der Stärke der wahrgenommenen Forscherbeobachtung in der Entstehung von Messreaktivität Hinweise auf soziale Erwünschtheit als möglichen Wirkmechanismus geben. Diese Erkenntnisse bringen nicht nur ein besseres Verständnis über mögliche Entstehungsmechanismen der Messreaktivität, sondern ermöglichen es auch die Forscherbeobachtung so zu gestalten, dass möglichst wenig Messreaktivität ausgelöst wird.

Bewusstheit über Verhaltensveränderung

Weiters wurde in der digitalen Bewegungsforschung die Frage aufgeworfen, ob Verhaltensveränderungen aufgrund von Messreaktivität bewusst oder unbewusst erfolgen (Arigo & König, 2024). Abgeleitet von dem möglichen Mechanismus der Selbstdarstellung als

Hintergrund der Messreaktivität, welche per Definition eine bewusste Verhaltensanpassung ist, würde eine Verhaltensveränderung bewusst geschehen (Börger, 2012). In der digitalen Gesundheitsforschung, und somit auch im digitalen Ernährungstracking, findet sich jedoch keine Forschung zu dieser Fragestellung.

Wenn Studienteilnehmende ihr Verhalten bewusst verändern, würde ein anderer Ansatz zur Reduktion der Messreaktivität Sinn machen. So könnte eventuell alleine die Aufforderungen, das Verhalten nicht zu verändern schon Messreaktivität reduzieren. Bei unbewussten Veränderungen würde dies jedoch kaum Sinn machen (Fitzsimons & Williams, 2000).

Neben der Untersuchung der Rolle von Ernährungszielen und der wahrgenommenen Forscherbeobachtung, kann damit auch das Erforschen, davon, ob Verhaltensveränderungen bewusst auftreten, zu einem besseren Verständnis der Messreaktivität und dadurch zu sinnvollen Reduktionsmaßnahmen verhelfen. In dieser Studie wird daher das Maß an Bewusstheit über Verhaltensveränderungen erhoben und folglich untersucht, ob diese die Entstehung von Messreaktivität moderiert.

Ziel der Studie

Die Studie setzt sich zum Ziel Messreaktivität im digitalen Ernährungstracking zu erforschen. Dadurch soll eine umfassendere Beurteilung der Validität des digitalen Ernährungstracking als vielversprechende und zunehmend eingesetzte Erhebungsmethode in der Ernährungsforschung ermöglicht werden.

Ein weiteres Ziel ist es, die Messreaktivität im Forschungskontext, isoliert vom Messkontext zu betrachten. Dies kann zu einer differenzierteren Einschätzung der Studienkontexte, die zu Messreaktivität führen, verhelfen. Um diese Trennung des Forschungskontexts und des Messkontexts durchzuführen, werden nur Personen rekrutiert, die schon vor der Studienteilnahme anhand einer App ihre Ernährung trackten. Dadurch wird nur der Forschungskontext neu eingeführt, unabhängig von der Ernährungsaufzeichnung als Messung. Durch diesen Ansatz bestehen zu dem auch bereits Daten über das Ernährungsverhalten der Teilnehmenden vor der Studie. Das ermöglicht einen direkten Vergleich der Ernährungsdaten vor und nach der Studienteilnahme und eine Veränderung dieser wäre auf einen Messreaktivitätseffekt zurückzuführen. Aus diesen Studienzielen ergibt sich folgende Fragestellung und Hypothese:

F1: Verändert sich das selbstberichtete Ernährungsverhalten im Alltag durch die Studienteilnahme?

H1: Die berichtete Kalorieneinnahme ist in den ersten zwei Tagen nach der Befragung niedriger als in der Woche vor der Befragung.

Weiters sollen drei Variablen als Moderatoren der Messreaktivität untersucht werden: Das selbstgesetzte Ernährungsziel, die wahrgenommene Forscherbeobachtung und die Bewusstheit über eine Verhaltensveränderung. Die Erforschung eines möglichen Einflusses dieser Variablen kann Aufschluss über Mechanismen geben, die der Entstehung der Messreaktivität zugrunde liegen. Erkenntnisse darüber können zur Entwicklung gezielter Maßnahmen und Empfehlungen zur Minimierung von Messreaktivität beitragen.

Aus dem Ziel mögliche Selbstregulationsmechanismen in der Messreaktivität zu untersuchen, soll der moderierende Effekt selbstgesetzter Ernährungsziele (Intention wird im weiteren als Synonym verwendet), als Indikator für bestehende Diskrepanz zwischen tatsächlichem und Zielverhalten, auf die Messreaktivität untersucht werden. Daraus ergibt sich die zweite Fragestellung und Hypothese:

F2: Besteht ein Zusammenhang zwischen dem angegebenen Ernährungsziele und der tatsächlichen Veränderung des Ernährungsverhaltens?

H2: Es besteht ein Zusammenhang zwischen der Intention, die Kalorieneinnahme zu verändern und der tatsächlichen Veränderung der Kalorieneinnahme in Richtung der Intention.

Aus dem Ziel Erkenntnisse einer möglichen Rolle der sozialen Erwünschtheit in der Entstehung von Messreaktivität zu gewinnen, soll der moderierende Effekt der wahrgenommene Forscherbeobachtung auf die Messreaktivität untersucht werden. Anstatt Forscherbeobachtung in verschiedenen Ausmaßen zu erzeugen, werden Personen am Ende der Studie gefragt, wie sehr sie sich dieser bewusst waren. Dadurch soll verhindert werden, dass ein Effekt aufgrund einer eventuell nicht wirksamen Manipulation nicht erkannt werden kann. Daraus ergibt sich die zweite Fragestellung und Hypothese:

Fragestellung 3: Besteht ein Zusammenhang zwischen der wahrgenommenen Forscherbeobachtung und der tatsächlichen Veränderung des Ernährungsverhaltens?

H3: Je ausgeprägter die Forscherbeobachtung wahrgenommen wird, desto stärker ist die tatsächliche Veränderung des Ernährungsverhaltens.

Aufgrund der kaum vorhandenen Forschungsgrundlage für die Bewusstheit einer Verhaltensveränderung im Kontext der Messreaktivität, ergibt sich folgende explorative Fragestellung.

E1: Besteht ein Zusammenhang zwischen der Bewusstheit von Ernährungsveränderung durch die Studienteilnahme und der tatsächlichen Veränderung des Ernährungsverhaltens?

Methodik

Die Arbeit wurde in der arbeitsbereichsinternen Studiendatenbank des Arbeitsbereichs Gesundheitspsychologie vor Durchführung der Datenauswertung präregistriert. Materialien, Daten und Auswertungsskripte dieser Arbeit sind in der arbeitsbereichsinternen Studiendatenbank des Arbeitsbereichs Gesundheitspsychologie hinterlegt. Zudem ist diese Arbeit Teil eines übergeordneten Projektes (König & Ruf, 2024), auf dessen Daten weitere Masterarbeiten basieren (Lackner, 2025; Lichtenberger, 2025).

Stichprobe

Da bislang keine Studien zur Messreaktivität im Ernährungsbereich existieren, die einen aussagekräftigen Aufschluss über die Effektstärke erlauben, konnte keine a-priori-Stichprobenumfangsberechnung durchgeführt werden. Das übergeordnete Projekt strebte bis zum 30. November 2025 eine Stichprobengröße von $N = 200$ an. Für die Masterarbeit wurde die Stichprobe verwendet, die bis zum 30. Juni 2025 vorlag. Somit ergab sich für diesen Teil des Projektes eine Gelegenheitsstichprobe von $N = 99$.

Die Rekrutierung erfolgte über die universitäre Plattform LABS, soziale Medien, Flyer und Mundpropaganda. Teilnehmende, die über LABS rekrutiert wurden, erhielten nach Abschluss beider Befragungen eine Entlohnung von vier LABS-Credits. Unter allen anderen Teilnehmenden wurden drei 100-Euro-Wunschgutscheine verlost.

Teilnahmevoraussetzungen waren ein Mindestalter von 18 Jahren sowie die Nutzung einer Ernährungstracking-App seit mindestens einer Woche. Während der Studienteilnahme wurden Personen ausgeschlossen, die beide in den ersten Fragebogen eingebauten Attention Checks nicht erfüllten. Ebenso wurden nur Personen inkludiert, die den zweiten Teil der Befragung innerhalb von 24 Stunden am siebten Tag nach der ersten Befragung abgeschlossen haben. Um sicherzustellen, dass nur Teilnehmende mit aussagekräftiger Datengrundlage in die Analysen einbezogen werden, wurden nachträglich Personen ausgeschlossen, die in der ersten Woche weniger als zwei tägliche Kalorienangaben und/oder in der zweiten Woche weniger als einen Eintrag am ersten oder zweiten Tag hatten. Von den 99 Personen, die die Teilnahmevoraussetzungen erfüllten, war dies bei $n = 13$ Personen der Fall. Somit verringerte sich die Stichprobengröße auf $N = 86$.

Materialien und Maße

Die Befragung erfolgte anhand von zwei Online-Fragebögen, daher war für die Teilnahme ein digitales Endgerät (z. B. Smartphone, Laptop, Tablet) erforderlich. Im Rahmen des Gesamtprojekts wurden Variablen erhoben, die für die vorliegende Arbeit nicht von Bedeutung waren. Im Folgenden werden ausschließlich die für die Forschungsfrage relevanten Variablen dargestellt.

Stichprobenbeschreibende Variablen

Im ersten Fragebogen wurden zunächst soziodemografische Angaben zu Alter, Geschlecht, Bildungsstand, Einkommen und Beschäftigungsstatus erhoben. Zusätzlich

machten die Teilnehmenden Angaben zu ihrer Körpergröße (in cm) und ihrem Gewicht (in kg). Bezüglich der Nutzung der Ernährungsapp wurden die Nutzungsdauer (in Monaten), die Nutzungskonsistenz auf einer fünfstufigen Likert-Skala (*1 = überhaupt nicht, 5 = sehr*) sowie der Name der verwendeten App erfasst. Weiters hatten Teilnehmende in einem offenen Textfeld die Möglichkeit, einen Vorfall in den letzten 7 Tagen zu berichten, der ihr Ernährungsverhalten beeinflussen könnte.

Tägliche Kalorieneinnahme kodiert nach Woche

Studienteilnehmende berichteten zu beiden Befragungszeitpunkten in einem Textfeld ihren täglichen Kalorienkonsum der letzten 7 Tage anhand der von ihnen in ihrer Ernährungsapp dokumentierten Daten (König, 2025). Die täglichen Kalorien der ersten Befragung wurden der Woche 1 zugeordnet, die der zweiten Befragung der Woche 2. Durch diese Kodierung ergibt sich die Variable „Woche“.

Kalorienziel

Beim ersten Erhebungszeitpunkt gaben die Teilnehmenden auf einer siebenstufigen Likert-Skala (*1 = Ich möchte viel weniger Kalorien essen bis 7 = Ich möchte viel mehr Kalorien essen*) an, inwieweit sie ihre Kalorienzufuhr verändern möchten.

Forscherbeobachtung

Gegen Ende des zweiten Fragebogens bewerteten die Teilnehmenden auf einer fünfstufigen Likert-Skala (*1 = überhaupt nicht bis 5 = sehr*), in welchem Ausmaß sie sich darüber bewusst waren, dass ihre Ernährungsdaten einem Team von WissenschaftlerInnen zur Verfügung gestellt werden würden.

Bewusste Ernährungsveränderung aufgrund der Studienteilnahme

Bei der zweiten Befragung sollte auf einer 5-stufigen Likert-Skala berichtet werden, inwieweit das Ernährungsverhalten bewusst aufgrund der Studienteilnahme verändert wurde. Die Skala ging von 1 (*nicht zutreffend*) bis 5 (*zutreffend*).

Design und Ablauf

Das Studiendesign orientierte sich an der Untersuchung von König et al. (2025) zu Messreaktivität in der digitalen Bewegungsforschung. Hier wurde bewusst der Forschungskontext vom Messkontext isoliert, indem Personen rekrutiert wurden, die schon vor der Studienteilnahme ihre Schrittzahl aufzeichneten. Analog wurden in dieser Studie nur Personen rekrutiert, die schon vor der Studienteilnahme ihr Ernährung trackten. Die vorliegende Studie umfasste zwei Erhebungszeitpunkte mittels Online-Fragebogen, die im Abstand von sieben Tagen durchgeführt wurden. Die Studie stellte somit eine beobachtende Längsschnittuntersuchung mit einem Within-Person-Design dar.

Die Datenerhebung erfolgte über die Plattformen *Unipark* und *SoSci Survey*. Interessierte Personen konnten über einen bereitgestellten Link an der Studie teilnehmen. Zu Beginn wurden die Teilnehmenden anhand eines Informed Consents darüber informiert, dass die Studie das Verhalten ihres Ernährungstracking untersucht. Es wurde jedoch nicht

erwähnt, dass auch mögliche Veränderungen des Ernährungsverhaltens im Zusammenhang mit der Studienteilnahme Gegenstand der Untersuchung sind, um eine mögliche Verzerrung der Antworten zu vermeiden. Nachdem die Teilnehmenden den Informed Consent gelesen und diesem zugestimmt hatten, erhielten sie den Link zum ersten Fragebogen. Personen, die die Einschlussvoraussetzungen nicht erfüllten, wurden vom weiteren Studienverlauf ausgeschlossen. Am Ende der ersten Befragung wurden die Teilnehmenden gebeten, ihre E-Mail-Adresse anzugeben, um sieben Tage später den Link zur zweiten Befragung zu erhalten. Dieser konnte innerhalb eines 24-Stunden-Zeitfensters am siebten Tag ausgefüllt werden. Nach Abschluss der zweiten Befragung erfolgte ein Debriefing. Im Anschluss hatten die Teilnehmenden die Möglichkeit, ihre E-Mail-Adresse zu hinterlassen, um an der Verlosung von drei 100-Euro-Wunschgutscheinen teilzunehmen. Teilnehmende, die über *LABS* rekrutiert wurden, wurden stattdessen auf die Plattform weitergeleitet und erhielten dort vier Credits als Vergütung.

Die Datenerhebung wurde jeweils für eine Woche vor und nach universitären Ferien (z. B. Weihnachten, Ostern) sowie eine Woche vor Semesterende bis zwei Wochen nach Semesterbeginn pausiert, um potenzielle Verzerrungen durch veränderte Routinen oder traditionsbedingte Ernährungsgewohnheiten zu vermeiden.

Datenaufbereitung

Zur Vorbereitung der Datenauswertung wurden zunächst in *Excel* alle Variablen entfernt, die für die vorliegende Teilstudie nicht relevant waren. Anschließend wurde überprüft, ob die Angaben zu Körpergröße und -gewicht in der korrekten Einheit (cm bzw. kg) vorlagen und gegebenenfalls angepasst. Auf Basis dieser Angaben wurde in *Excel* der Body-Mass-Index (BMI) berechnet. Weiters wurden die Namen identischer Ernährungsapps manuell vereinheitlicht, um eine sinnvolle deskriptive Analyse zu ermöglichen.

Die weitere Datenaufbereitung und Analyse erfolgte in R (Version 4.5.0; R Core Team, 2025) unter Verwendung von RStudio (Version 2025.05.0; Posit Software, 2025). Zunächst wurde die Zuordnung der Bildungskategorien überprüft. Bei $n = 4$ Personen traten Unstimmigkeiten zwischen dem Inhalt des offenen Textfelds und der zugeordneten Kategorie auf. Diese wurden den offenen Angaben entsprechend der passenden Kategorie zugeordnet. Anschließend wurde die Kennzeichnung von fehlenden Werten einheitlich als *NA* kodiert. Eine mögliche Systematik der fehlenden Werte wurde mithilfe des Little-MCAR-Tests überprüft. Der Test fiel nicht signifikant aus ($p = 0.233$), sodass von einer zufälligen Verteilung fehlender Werte ausgegangen werden konnte.

Als Vorbereitung auf die Multilevel Analyse wurde der Datensatz in ein Long-Format transformiert, wodurch jede tägliche Kalorienangabe einer Person als separate Beobachtung abgebildet ist. Die entsprechenden Angaben zur Vollständigkeit der Kalorienangabe wurden den jeweiligen Tageswerten zugeordnet. Die einzelnen Kalorienangaben wurden jeweils durch die Variable *Woche* (T1 oder T2), den *Tag* (1-7) und *Woche_Tag* (T1_Kcal_1,

T1_Kcal2,...T2_Kcal_7) eindeutig benannt. Die Kodierung der Woche weicht somit von der Präregistrierung ab (Woche 1 als 0, Woche 2 als 1), um zu vermeiden, dass die Kodierung 1 mit Woche 1 verwechselt wird. Anschließend wurden die Variablen *Person*, *Woche* als Faktoren definiert, sodass sie von *R* als kategoriale Variablen behandelt wurden. *Tag* blieb als numerische Variable definiert, um den zeitlichen Verlauf innerhalb der Woche abbilden zu können.

Die Tage 3 bis 7 der zweiten Woche wurden aus dem Datensatz ausgeschlossen, da sie für die Hypothesenprüfung nicht relevant waren. Die Variablen *Kalorienziel*, *Forscherbeobachtung* und *bewusste Ernährungsveränderung* wurden grand-mean zentriert, um Cross-Level-Analysen zu ermöglichen. Personen mit weniger als zwei gültigen Kalorienangaben in Woche 1 oder weniger als einer gültigen Angabe in Woche 2 wurden ausgeschlossen ($n = 13$), somit ergab sich eine Stichprobe von $N = 86$. Die Vollständigkeit der Angaben zu *Kalorienziel*, *Kalorieneinnahme* und *wahrgenommener Forscherbeobachtung* wurde überprüft und zeigte sich als vollständig.

Deskriptive Statistik

Für die deskriptive Auswertung wurden je nach Skalenniveau unterschiedliche Kennwerte berechnet. Für nominalskalierte Variablen wurden der Modus sowie absolute und relative Häufigkeiten bestimmt, für ordinalskalierte Variablen der Median, sowie die absolute und relative Häufigkeit, und für metrische Variablen der Mittelwert, die Standardabweichung sowie Minimum und Maximum. Die genannten Vorfälle der letzten 7 Tage wurden grob kategorisiert und beschrieben. Somit wurden mehr Kennwerte berechnet als in der Präregistrierung vorgesehen. Dadurch soll ein möglichst umfangreicher Überblick über die Stichprobe geschaffen werden.

Inferenzstatistische Analysen

Aufgrund der verschachtelten Struktur wurden zur Prüfung der Hypothesen lineare gemischte Modelle (LMMs) mit dem *R*-Paket *lmerTest* (Kuznetsova et al., 2017) berechnet und verglichen. Die Modelle berücksichtigten Messwiederholungen auf zwei Ebenen. Auf der ersten Ebene befanden sich die täglichen Kalorienangaben, auf der zweiten Ebene die Personen. Für jegliche Signifikanzprüfungen wurde ein Alpha-Niveau von $p < .05$ festgelegt.

Modelle für H1: Veränderung der Kalorienaufnahme zwischen den Wochen. Zur Prüfung der Hypothese 1 erfolgte eine schrittweise Modellierung von drei Modellen, deren Passung miteinander verglichen wurde, um das mit dem besten Modellfit zu bestimmen. Zunächst wurde als Ausgangspunkt ein Nullmodell mit zufälligem Interzept geschätzt (M10). Anschließend wurde dies zu einem random Intercept/fixed Slope Modell (M11) erweitert, mit der Woche als fester Effekt. Im dritten Schritt wurde für Woche zusätzlich ein zufälliger Slope auf Personenebene spezifiziert (M12), um interindividuelle Unterschiede in der Veränderung zwischen den Wochen zu erlauben.

Zur Beurteilung und zum Vergleich der Modellgüte wurden das Akaike-Informationskriterium (AIC) und das Bayesianische Informationskriterium (BIC) herangezogen, sowie die Signifikanz des Likelihood-Ratio-Tests (LRT) mittels Maximum Likelihood (ML) geprüft. Um die Hypothese zu prüfen, wird die Signifikanz des festen Effekts, also der Woche, beurteilt. Dafür wird das Modell mit dem besten Fit herangezogen. Wenn der feste Effekt signifikant ist, wird H1 angenommen, andernfalls verworfen. Als Maß der Effektgröße wurde das marginale und konditionale R^2 herangezogen. Zusätzlich zu diesem präregistrierten Vorgehen wurde auch R^2 -within und R^2 -between bestimmt, um die Varianzanteile innerhalb und zwischen Personen differenzierter einordnen zu können.

Cross-Level Modelle: Moderatoren der Messreaktivität. Zur Prüfung der weiteren Hypothesen bzw. für die Exploration wurde das Grundmodell mit dem besten Fit um den Cross-Level-Prädiktor *Intention* (M2) für Hypothese 2, *wahrgenommene Forscherbeobachtung* (M3) für Hypothese 3 sowie *Bewusstheit über Ernährungsveränderung* (M4) für die Exploration erweitert, um zu prüfen, ob diese Variablen die Veränderung der Kalorienaufnahme zwischen Woche 1 und Woche 2 moderieren.

Zur Beurteilung, inwieweit die Einbeziehung eines Cross-Level-Prädiktors die Modellgüte verbessert, wurden die jeweiligen Cross-Level-Modelle mit dem zuvor identifizierten bestpassenden Grundmodell verglichen. Der Vergleich erfolgte, analog zur Vorgehensweise der ersten Hypothese, mittels AIC/BIC und LRT. Es wird jeweils der feste Effekt der Interaktion zwischen Woche und Cross-Level Prädiktor geprüft. Fällt dieser Effekt signifikant aus, wird die jeweilige Alternativhypothese (H2/H3) angenommen. Auch für die Exploration wird der feste Effekt der Interaktion herangezogen. Zusätzlich werden auch die festen Effekte der Cross-Level Prädiktoren geprüft. Die erklärte Varianz wurde zusätzlich über das marginale und konditionale R^2 sowie R^2 -within und R^2 -between bestimmt.

Modellannahmen. Nach der Spezifizierung der einzelnen Modelle wurden die zentralen Annahmen linearer gemischter Modelle sowohl formell mithilfe des DHARMA-Pakets (Hartig, 2022) als auch visuell für jedes Modell einzeln geprüft. Da sich die Ergebnisse der Annahmeprüfung zwischen den Modellen allerdings kaum unterscheiden, werden sie zusammengefasst dargestellt.

Die Normalverteilung der Residuen wurde anhand von Histogrammen, QQ-Plots sowie dem DHARMA-Uniformity-Test beurteilt. Das Histogramm der Level-1-Residuen zeigte eine annähernd glockenförmige Verteilung. Die QQ-Plots sind zwar leichte Abweichungen in den Extrembereichen ersichtlich, allerdings fiel der Uniformity-Test nicht signifikant ($p > .05$) aus. Daher wird die Annahme der Normalverteilung als erfüllt betrachtet. Auch der QQ-Plot der Random Effects zeigte Abweichungen in den Extrembereichen. Diese wurde aufgrund der ausreichenden Stichprobengröße von $N = 86$ jedoch als nicht weiter gravierend bewertet (Pek et al., 2018). Die Überprüfung der Homoskedastizität erfolgte über den DHARMA-

Dispersion-Test, der keine Hinweise auf Heteroskedastizität ergab ($p > .05$). Zudem zeigen die Residuen-Fitted Plots keine Systematik in der Streuung. Dies stützt weiter die Voraussetzungserfüllung der Homoskedastizität und weist weiters auf eine Erfüllung der Linearität hin.

Zusätzlich zu den präregistrierten Annahmeprüfungen wurden die Prädiktoren der Cross-Level-Modelle auf Multikollinearität überprüft, um sicherzustellen, dass die Schätzungen der festen Effekte nicht durch korrelierte Prädiktoren verzerrt werden. Die Variance Inflation Factors (VIF = 1.00–1.02) lagen im unproblematischen Bereich und zeigten somit keine Hinweise auf Multikollinearität.

Insgesamt zeigten sich keine bedeutsamen Verletzungen der zentralen Modellannahmen. Damit können alle Modelle als angemessen spezifiziert und robust bewertet werden.

Ausreißer- und Einflussdiagnostik. Im Zuge der Ausreißerdiagnostik wurden potenzielle Ausreißer sowohl auf Level 1 als auch auf Level 2 geprüft. Auf Level 1 erfolgte die Identifikation ungewöhnlicher Residuen anhand des DHARMA-Outlier-Tests, der keine Auffälligkeiten ergab ($p > .05$). Auf Level 2 wurden potenziell einflussreiche Personen mittels Cook's D bestimmt. Keine Person übertrat den Schwellenwert von 1, somit ergaben sich keine Hinweise darauf, dass einzelne Personen die Modellergebnisse übermäßig beeinflussten. Daher wurden keine Sensitivitätsanalysen durchgeführt.

Ergebnisse

Deskriptive Statistik der Soziodemographischen Daten

Unter Ausschluss der Personen mit zu wenigen Daten ergab sich eine Stichprobe von $N = 86$ Personen im Alter von 19 bis 65 Jahren ($M = 27.23$, $SD = 7.94$). Davon identifizieren sich 64.0% ($n = 55$) als weiblich, 36.0% ($n = 31$) als männlich. Der durchschnittliche BMI beträgt 24.3 ($SD = 4.1$), mit einer Range von 19.1 bis 36.8.

In der ersten Woche gaben 24 Personen einen Vorfall an, der die Ernährung beeinflusst, in der zweiten Woche 27 Personen. Die genannten Vorfälle lassen sich grob fünf Bereichen *körperlicher Zustand/Eingriff* (z. B. Schmerzen, Krankheit, Heißhunger, Zyklus, OP, Gastroskopie), *Stress* (z. B. Prüfungsstress, erhöhtes Arbeitspensum), *soziale Ereignisse* (z. B. Geburtstage, Essen mit befreundeten Personen, Familienfeiern, Beerdigung, geschäftliches Abendessen), *Veränderungen in der Routine* (z. B. Reisen, Besuch, neues Projekt) sowie *Veränderungen im Lebensstil- und Gesundheitsverhalten* (z. B. mehr/weniger Training, Keto-Diät, Muskelaufbau, bewusst weniger Tracking) gliedern.

Die Verteilung und der Median des höchsten Schulabschlusses, des höchsten Ausbildungsabschlusses, des derzeitigen Beschäftigungsstandes sowie des monatlichen Nettoeinkommens sind jeweils in Tabelle 1 dargestellt. Eine Person gab in der offenen Texteingabe an über einen Fachhochschulabschluss zu verfügen, eine weitere Person gab

einen Werkmeister in Elektrotechnik an. Dies sind keine Angabe über den höchsten Schulabschluss, somit ergeben sich zwei fehlende Angabe in der Erfassung des höchsten Schulabschlusses. Beide Angaben werden im höchsten Ausbildungsabschluss berücksichtigt. Bezüglich der Beschäftigung wurden Personen, die Angaben zu ihrem Studienfach machten, der Gruppe in Ausbildung/Umschulung/Studium zugeordnet.

Tabelle 1

Soziodemographische Verteilung

Beschreibung	n	%	Median
Höchster Schulabschluss	84		6 = Hochschulreife
(Noch) kein Schulabschluss	0	0.0	
Volksschul-/Hauptschulabschluss	0	0.0	
Mittlere Reife / Realschulabschluss	1	1.2	
Abschluss der Polytechnischen Oberschule	0	0.0	
Abschluss einer (Berufs-)Fachschule	3	3.5	
Hochschulreife (Abitur/Matura)	80	94.1	
Anderer Abschluss	0	0.0	
Höchster Ausbildungsabschluss	86		5 = Hochschule
(Noch) keine abgeschlossene Berufsausbildung	29	33.7	
Lehre	3	3.5	
Fachoberschule / Berufsoberschule	7	8.1	
Fachhochschule	6	7.0	
Hochschule	38	44.2	
Promotion	2	2.3	
Anderer Abschluss	1	1.2	
Monatliches Nettoeinkommen	86		6 = 1.500–2.000 €
< 150 €	4	4.7	
150–300 €	1	1.2	
300–500 €	4	4.7	
500–1.000 €	14	16.3	
1.000–1.500 €	13	15.1	
1.500–2.000 €	8	9.3	
2.000–2.500 €	9	10.5	
2.500–3.000 €	9	10.5	
3.000–5.000 €	10	11.6	
5.000–10.000 €	12	14.0	
> 10.000 €	2	2.3	

Beschreibung	n	%	Median
Beschäftigungsstatus	86		
Berufstätig (Vollzeit)	21	24.4	
Berufstätig (Teilzeit)	9	10.5	
Arbeitssuchend	3	3.5	
In Ausbildung / Umschulung / Studium	38	44.2	
Hausfrau / Hausmann	0	0.0	
Rentner*in	0	0.0	
Studierende (Studienfach angegeben)	15	17.4	

Deskriptive Statistik der App-Nutzung

Die Verteilung der verwendeten Apps ist der Tabelle 2 zu entnehmen. Als durchschnittliche Nutzungsdauer der Apps wurde 15.11 Monate ($SD = 26.17$), mit einem Minimum von 2 Wochen und einem Maximum von 120 Monaten, angegeben. Die Nutzungskonsistenz zeigte sich als überwiegend regelmäßig ($Mdn = 4$). 30.2% ($n = 26$) bewerteten ihre Nutzungskonsistenz mit einer 4 und 33.7% ($n = 29$) mit einer 5. 17 Personen (19.8%) gaben eine mittlere Nutzungskonsistenz von 3 an, während 10 Personen (11.6%) eine niedrige Nutzungskonsistenz von 2 berichteten. Vier Personen (4.7%) gaben eine *überhaupt nicht konsistente* Nutzung von 1 an.

Tabelle 2

Verteilung verwendeter Apps

App	n	%
Yazio	33	38.4
MyFitnessPal	21	24.4
Lifesum	8	9.3
Fddb	6	7.0
FatSecret	3	3.5
Samsung Health	3	3.5
Fitbit	1	1.2
Noom	2	2.3
Foodvisor	1	1.2
Happy Cat	1	1.2
Macrofactor	1	1.2
mci	1	1.2
ChatGPT	1	1.2

App	n	%
Mehrere Apps (Yazio, Fastic Gesund, Lifesum, Apple Health)	2	2.4

Deskriptive Statistik der hypothesenspezifischen Daten

In Bezug auf die Kalorienaufnahme berichteten die Teilnehmenden eine durchschnittliche Energieaufnahme von 1833.22 Kcal ($SE = 517.53$) in Woche 1 mit einem Minimum von 543.4 Kcal bis zu einem Maximum von 3421.57 Kcal. Für die ersten zwei Tage der zweiten Woche lag die durchschnittliche Aufnahme bei 1877.80 Kcal ($SE = 579.16$), mit einer Spanne von 790 bis 3324.50 Kcal.

Bezüglich des Kalorienziels zeigte sich mit $n = 58$ (64.4%) eine Tendenz zu einer niedrigeren Kalorienaufnahme. 18 Personen (22.1%) gaben an gleich viele Kalorien zu sich nehmen zu wollen, während 9 Personen (10.5%) ein höheres Kalorienziel berichteten. Keine Person gab an sehr viel mehr Kalorien zu sich nehmen zu wollen.

Bezüglich der Bewusstheit über die wahrgenommene Forscherbeobachtung gab 48.8% ($n = 42$) mit einem Wert von 5 ($n = 29$; 33.7%) oder 4 ($n = 13$; 15.1%) an, sich der Forscherbeobachtung bewusst zu sein. 20.9% ($n = 18$) gaben einen Wert von 3 an. 26 Personen (30.2%) gaben mit einem Wert von 1 ($n = 11$; 12.8%) oder 2 ($n = 15$; 17.4%) an sich der Forscherbeobachtung weniger bewusst gewesen zu sein. Der Median lag bei 3.5.

Zur bewussten Veränderung des Ernährungsverhaltens gab der Großteil der Personen an, dass dies nicht zutreffe ($n = 59$; 68.6%) bzw. eher nicht zutreffe ($n = 22$; 25.6%). Je zwei Personen (2.3%) gaben 3 (*teils-teils*) und 4 (*eher zutreffend*) an. Eine Person (1.2%) gab 5 (*zutreffend*) an. Der Median lag bei 1 (*nicht zutreffend*).

Hypothese 1: Veränderung der Kalorienaufnahme zwischen den Wochen

Zur Überprüfung der ersten Hypothese wurde untersucht, ob die berichtete Kalorienaufnahme in den ersten zwei Tagen nach der Befragung niedriger ist als in der Woche vor der Befragung. Der Vergleich der drei Modelle (M10–M12) zeigte, dass das Random-Intercept/Random-Slope-Modell (M12) insgesamt die beste Passung an die Daten aufwies. Im AIC-Vergleich erzielte M12 die niedrigsten Werte und weist damit auf eine verbesserte Modellgüte hin. Der BIC fiel hingegen am höchsten aus, was darauf zurückzuführen sein könnte, dass der BIC komplexere Modellstrukturen stärker bestraft. Im Likelihood-Ratio-Test zeigte M12 eine signifikant bessere Passung im Vergleich zu M11 ($p = .047$). Zudem erklärte M12 den größten Anteil an Varianz (konditionales $R^2 = .749$). Insgesamt weist M12 somit die beste Modellpassung auf und wird als Grundmodell für die Hypothesenprüfungen und die explorative Analyse herangezogen. Die vollständigen Modellkennwerte sind in Tabelle 3 dargestellt.

Der feste Effekt der Woche zeigte in M12 eine leichte Zunahme der Kalorienaufnahme ($b = 45.46$), erreichte jedoch keine statistische Signifikanz, $t(87.09) = 1.46$, $p = .147$. Hypothese 1 wird somit nicht bestätigt. Der fehlende Effekt spiegelt sich auch in der Varianzaufteilung wider: Der feste Effekt der Woche erklärte lediglich einen minimalen Anteil der Gesamtvarianz ($R^2\text{-marginal} = .001$) und trug damit kaum zur Veränderung der Kalorienaufnahme bei. Die hohe erklärte Gesamtvarianz des Modells (konditionales $R^2 = .749$) ist vor allem auf Unterschiede zwischen den Personen zurückzuführen. Der innerhalb der Personen erklärte Varianzanteil ($R^2\text{-within} = .058$) weist zwar auf etwas individuelle Schwankungen hin, diese wurden jedoch nicht durch die Woche erklärt.

Tabelle 3

Vergleich der Grundmodelle (M10/M11/M12)

Modell	AIC	BIC	R^2 (marg.)	R^2 (cond.)	R^2 (within)	R^2 (between)	LRT	p
M10	11,031.8	11,045.6	0.000	0.734	0.000	0.012		
M11	11,030.8	11,049.3	0.001	0.735	0.004	0.013	2.95	0.086
M12	11,028.7	11,056.5	0.001	0.749	0.058	0.024	6.12	0.047*

Anmerkung: Nullmodell mit zufälligem Interzept (M10); Modell mit zufälligem Interzept/festem Slope Modell (M11); Modell mit zufälligem Interzept/zufälligem Slope (M12); LRT = Likelihood-Ratio-Test mit jeweilig vorhergehendem Modell; $p < .05^*$, $p < .01^{**}$, $p < .001^{***}$.

Hypothese 2: Interaktion von Woche und Intention

Zur Überprüfung der zweiten Hypothese wurde untersucht, ob die Intention zur Kalorienreduktion den Zusammenhang zwischen Woche und berichteter Kalorienaufnahme moderiert. Im Modellvergleich zu M12 wies M2 niedrigere AIC und BIC-Werte auf. Ebenso zeigte der Likelihood-Ratio-Test eine signifikante Verbesserung der Modellpassung ($p < .001$). Insgesamt zeigte der Modellvergleich, dass die Einbeziehung der Intention die Gesamtmodellgüte verbessert.

Auf Ebene der festen Effekte zeigte sich ein signifikanter Haupteffekt der Intention ($b = 198.78$, $p < .001$). Das marginale R^2 stieg im Vergleich zu M12 deutlich an ($R^2\text{-marginal} = .208$) ebenso wie der zwischenpersonenbezogene Anteil ($R^2\text{-between} = .292$). Die Intention erklärt somit Unterschiede zwischen den Personen. Der Interaktionseffekt zwischen Woche und Intention war jedoch nicht signifikant ($t(86.24) = 0.63$, $p = .532$). Somit stellt die Intention keinen signifikanten Moderator einer Veränderung der berichteten Kalorien zwischen Woche 1 und 2 dar. Die Erklärung der Varianz ist hauptsächlich auf den Haupteffekt der Intention zurückzuführen. Hypothese 2 wird damit nicht bestätigt. Bezüglich des konditionalen R^2 und $R^2\text{-within}$ zeigten sich keine Unterschiede zu M12.

Hypothese 3: Interaktion von Woche und wahrgenommener Forscherbeobachtung

Für die dritte Hypothese wurde geprüft, ob die wahrgenommene Forscherbeobachtung die Veränderung der Kalorienaufnahme zwischen den Wochen moderiert. Das Cross-Level-Modell M3 wurde hierzu mit dem Grundmodell M12 verglichen (siehe Tabelle 4). Im AIC- oder BIC-Vergleich zeigte sich eine Reduktion der Modellgüte gegenüber M12, und auch der Likelihood-Ratio-Test ergab keine signifikante Verbesserung (LRT = 1.20, $p = .550$). Damit lieferte der Modellvergleich keinen Hinweis auf eine gesteigerte Modellgüte durch die Einbeziehung der wahrgenommenen Beobachtung. Auch die festen Effekte zeigten keinen Hinweis auf einen bedeutsamen Einfluss der wahrgenommenen Beobachtung. Weder der Haupteffekt ($b = 38.80$, $p = .319$) noch die Interaktion mit der Woche ($t(87.06) = 0.39$, $p = .696$) waren signifikant. Hypothese 3 wird daher nicht bestätigt.

Bezüglich der Varianzerklärung zeigten sich im Vergleich zu M12 kaum Unterschiede. Die wahrgenommene Forscherbeobachtung führte lediglich zu einer leicht erhöhten erklärten Varianz durch die festen Effekte ($R^2\text{-marginale} = .011$) sowie der Unterschiede zwischen Personen ($R^2\text{-between} = .036$).

Explorative Analyse: Interaktion von Woche und Bewusstheit über Ernährungsveränderungen

Im explorativen Modell (M4) wurde untersucht, ob die Bewusstheit über Ernährungsveränderungen eine Veränderung der Kalorienaufnahme zwischen der ersten und der zweiten Woche moderiert. Der AIC- und BIC-Vergleich ergab eine geringere Modellgüte von M4 gegenüber M12. Weiters zeigte der Likelihood-Ratio-Test keine signifikante Verbesserung der Modellgüte (LRT = 0.41, $p = .815$). Die Hinzunahme der Bewusstheit führte somit nicht zu einer verbesserten Modellpassung.

Auf Ebene der festen Effekte zeigte sich weder ein signifikanter Haupteffekt der Bewusstheit ($b = 35.30$, $p = .633$) noch ein Interaktionseffekt mit der Woche ($t(89.66) = 0.39$, $p = .695$). Weiters erklärte die Hinzunahme der Bewusstheit über Ernährungsveränderung kaum zusätzliche Varianz. Das marginale R^2 stieg minimal auf .003 an und der zwischenpersonenbezogene Anteil auf .027. Damit ergaben sich keine Hinweise auf einen moderierenden Einfluss der Bewusstheit auf eine Veränderung der Kalorienaufnahme zwischen den Wochen.

Tabelle 4

Vergleich der Cross-Level Modelle (M2/M3/M4) mit dem Grundmodell M12

Modell	AIC	BIC	R^2 (marg.)	R^2 (cond.)	R^2 (within)	R^2 (between)	LRT	p
M12	11,028.7	11,056.5	0.001	0.749	0.058	0.024		

Modell	AIC	BIC	R ² (marg.)	R ² (cond.)	R ² (within)	R ² (between)	LRT	p
M2	11,005.6	11,042.6	0.208	0.749	0.058	0.292	27.12	< .001***
M3	11,031.5	11,068.5	0.011	0.749	0.058	0.036	1.20	0.550
M4	11,032.3	11,069.3	0.003	0.749	0.058	0.027	0.41	0.815

Anmerkung: Modell mit zufälligem Interzept/zufälligem Slope (M12); Cross-Level Modell mit Ernährungsziel (M2), wahrgenommener Forscherbeobachtung (M3) und Bewusstheit über Verhaltensveränderung (M4) als Level 2 Prädiktor; LRT = Likelihood-Ratio-Test je mit M12; p < .05*, p < .01**, p < .001***.

Diskussion

Ziel der Studie war es, Messreaktivität im digitalen Ernährungstracking zu untersuchen. Dadurch sollte eine differenziertere Beurteilung des digitalen Ernährungstracking als vielversprechende Erhebungsmethode ermöglicht werden. Ein besonderer Fokus lag dabei auf dem Effekt des Forschungskontexts, isoliert von möglichen Verhaltensveränderungen, die allein durch die Einführung eines neuen Messinstruments entstehen könnten. Konkret wurde geprüft, ob sich die Kalorieneinnahmen von Personen, die bereits vor der Studienteilnahme ihre Ernährung mithilfe einer App aufzeichneten, zwischen der Woche vor Studienbeginn und den ersten beiden Tagen der Studienteilnahme verändert. Darüber hinaus wurde die moderierende Rolle von Ernährungszielen, der wahrgenommenen Forscherbeobachtung sowie der Bewusstheit über Ernährungsveränderungen im Kontext der Messreaktivität untersucht, um zu einem vertieften Verständnis der Entstehungsbedingungen von Messreaktivität beizutragen.

Entgegen den aufgestellten Hypothesen zeigten die Daten weder eine signifikante Veränderung der Kalorieneinnahme infolge der Studienteilnahme noch signifikante Interaktionseffekte der untersuchten Moderatoren. Entsprechend fanden sich in dieser Studie weder bedeutsame Hinweise auf Messreaktivität im digitalen Ernährungstracking noch auf moderierende Effekte von Ernährungszielen, wahrgenommener Forscherbeobachtung oder der Bewusstheit über Ernährungsveränderungen.

Hypothese 1

In der ersten Hypothese wurde angenommen, dass die berichtete Kalorienaufnahme in den ersten beiden Tagen nach der Befragung niedriger ausfällt als in der Woche davor. Insgesamt zeigt sich jedoch keine systematische Veränderung durch den Forschungskontext. Stattdessen sprechen die Befunde dafür, dass Veränderungen im Ernährungsverhalten vor allem durch individuelle Unterschiede zwischen den Teilnehmenden bedingt sind und nicht systematisch durch die Studienteilnahme. Zudem zeigte sich, wenn auch statistisch nicht bedeutsam, numerisch eine erhöhte Kalorienangabe während der Studienteilnahme und somit eine Veränderung entgegen der postulierten Richtung. Entsprechend unterstützen die Ergebnisse der Studie die erste Hypothese nicht.

Die Befunde dieser Studie ordnen sich somit zunächst nicht direkt in den Forschungsstand der Messreaktivität im Bereich der digitalen Gesundheitsforschung ein, in dem überwiegend kleine bis mittleren Effekten berichtet werden (König et al., 2022). Gleichzeitig liegen jedoch auch Studien vor, in denen keine Messreaktivität nachgewiesen wurde, sodass sich die Befundlage insgesamt als heterogen darstellt (für einen Überblick siehe König et al., 2022). Insbesondere vor dem Hintergrund, dass sich die vorliegende Studie in mehreren zentralen Aspekten von einem Großteil der bisherigen Untersuchungen unterscheidet (für einen Überblick siehe König et al., 2022), lassen sich die Ergebnisse eher als differenzierende Ergänzung der bestehenden Befundlage und nicht als Widerspruch

interpretieren. Für die Einordnung der Befunde der ersten Hypothese sind dabei insbesondere zwei Unterschiede zentral: die untersuchte Verhaltensweise und die gezielte Weiterverwendung eines bestehenden Messinstruments zur Isolierung des Forschungskontexts.

Untersuchtes Verhalten

Ein möglicher Erklärungsansatz für die Befunde betrifft zunächst die Art des untersuchten Verhaltens. Die bisherige Forschung zu Messreaktivität in der digitalen Gesundheitsforschung konzentriert sich überwiegend auf Bewegungsverhalten (König, 2022). In diesem Bereich wurden wiederholt kleine bis mittlere Effekte nachgewiesen. Demgegenüber ist die Befundlage zu Messreaktivität in anderen gesundheitsbezogenen Verhaltensweisen bislang deutlich begrenzter und zeigt sich als heterogen. Hier liegen lediglich vereinzelte Studien zu weiteren Verhaltensbereichen vor, etwa zur Medikamentenadhärenz (Cook et al., 2012; Sutton et al., 2014) oder zum Konsum alkoholischer (Labhart et al., 2020; Poulton et al., 2019; Yang et al., 2015) und nicht-alkoholischer Getränke (Labhart et al., 2020). Während in einigen dieser Studien Hinweise auf Messreaktivität gefunden wurden (Cook et al., 2012; Labhart et al., 2020, Poulton et al., 2019), blieben entsprechende Effekte in anderen Untersuchungen aus (Labhart et al., 2020; Sutton et al., 2014; Yang et al., 2015). Insgesamt lässt sich somit festhalten, dass Messreaktivität außerhalb des Bewegungsverhaltens bislang nicht konsistent belegt ist (König et al., 2022). Dies legt nahe, dass sich messreaktive Effekte, die im Bewegungsverhalten beobachtet werden, nicht ohne Weiteres auf das Ernährungsverhalten übertragen lassen könnten.

Eine mögliche Ursache für eine eingeschränkte Übertragbarkeit könnte in der unterschiedlichen Veränderbarkeit verschiedener Verhaltensweisen liegen. So legen Befunde aus der Bewegungsforschung nahe, dass leichter veränderbare Verhaltensweisen, wie die tägliche Schrittzahl, besonders anfällig für messreaktive Effekte sind, verglichen mit schwerer veränderbarem Verhalten, wie moderate und intensive körperliche Aktivität (König et al., 2022). Ernährungsverhalten stellt ein stark habituiertes und schwer veränderbares Verhalten dar (Pierce-Messick & Corbit, 2021). Entsprechend könnte eine schwere Veränderbarkeit des Ernährungsverhaltens dazu beitragen, dass sich messreaktive Effekte in diesem Bereich weniger deutlich oder gar nicht zeigen.

Verwendung des eigenen Messinstruments

Weitere Erkläransätze für die vom Forschungsstand abweichenden Befunde, ergeben sich daraus, dass ein Kernaspekt der vorliegenden Studie darin besteht, dass bewusst keine Verwendung eines neuen Messinstruments eingeführt wird, sondern die schon bekannte Ernährungsapp weiterverwendet wird. Dies unterscheidet die vorliegende Untersuchung deutlich von den meisten Studien im Bereich digitaler Gesundheitsforschung, in denen Messreaktivität gezeigt wurde. In diesen wird typischerweise zu Beginn der Datenerhebung

ein neues Gerät, wie ein Schrittzähler oder ein Aktivitätstracker, bereitgestellt (König et al., 2022). Somit wird in diesen Designs mit Beginn der Studie nicht nur der Forschungskontext, sondern auch ein Messkontext, durch die Verwendung eines neuen Messinstruments, eingeführt. Die vorliegende Studie zeichnet sich somit dadurch aus, dass der Forschungskontext isoliert untersucht wird. Zudem könnte das gewählte Studiendesign unbeabsichtigt dazu geführt haben, dass Ernährungsverhalten nicht als eindeutiger Forschungsfokus wahrgenommen wurde. Beide diese Aspekte stellen die Grundlage für weitere Erkläransätze dar.

Isolierung des Forschungskontexts. Die vorliegende Studie unterscheidet sich von einem Großteil anderer Studien, in dem Aspekt, dass der Forschungskontext isoliert vom Messkontext untersucht wird. Die Abwesenheit von Messreaktivität in dieser Studie könnte somit darin begründet sein, dass zur Entstehung dieser weniger der Forschungskontext, sondern eher der Messkontext, beiträgt. Diese Annahme wird unterstützt durch die Befunde einer Studie von König et al. (2025). Ziel der Studie war eine isolierte Untersuchung von Messreaktivität im Forschungskontext, indem Personen, die schon vor Studienbeginn ein Gerät zur Bewegungsaufzeichnung nutzten, dies während der Studie weiterverwendeten. Ebenso wie in der vorliegenden Studie konnte hier keine Messreaktivität festgestellt werden (König et al., 2025).

Weiters stützen Erkenntnisse aus der Forschung im Bereich des Self-Monitorings die Annahme, dass in der Messreaktivitätsforschung beobachtete Verhaltensveränderungen unter anderem auf die Verwendung neuer Messinstrumente zurückgeführt werden könnten. Self-Monitoring beschreibt das systematische Aufzeichnen von Daten, unter anderem Verhalten, durch die Teilnehmenden selbst in ihrem Alltagssetting (Korotitsch & Nelson-Gray, 1999). Dabei kommen unter anderem digitale Anwendungen zum Einsatz, darunter auch Schrittzähler oder Apps zum Ernährungstracking (Berry et al., 2021; Butryn et al., 2020; Page et al., 2020; Sanders et al., 2016).

In der Forschung zur Verhaltensänderung zeigt sich Self-Monitoring als eine der wirksamsten Techniken zur Förderung von Verhaltensveränderungen, unter anderem im Bereich des Bewegungs- und Ernährungsverhaltens (Michie et al., 2009). Entsprechend zeigen Studien und systematische Reviews, dass der Einsatz digitaler Self-Monitoring-Tools mit Veränderungen im Bewegungsverhalten, der Ernährung sowie folglich mit Gewichtsreduktion einhergeht (Berry et al., 2021; Bravata et al., 2007; Burke et al., 2011)

Vor diesem Hintergrund erscheint es plausibel, dass die Einführung neuer Messinstrumente im Rahmen von Studien auch ungewollt verstärkte Self-Monitoring-Prozesse aktiviert und dadurch messreaktive Effekte begünstigen kann.

Zusammenfassend legen die Befunde auf der Self-Monitoring Forschung sowie die Studie von König et al. (2025) nahe, dass Messreaktivität in digitalen Gesundheitsstudien

möglicherweise weniger durch den Forschungskontext, sondern eher durch den Messkontext, konkreter durch die Verwendung von Messgeräten, die Self-Monitoring fördern, hervorgerufen wird. Da keine neuen Messgeräte in dieser Studie eingesetzt wurden, könnte so auch das Ausbleiben messreaktiver Effekte erklärt werden.

Salienz der untersuchten Verhaltensweisen. Da in der vorliegenden Studie kein neues Messgerät eingeführt wurde, sondern bestehendes Ernährungstracking genutzt wurde, nahm das Trackingverhalten unbeabsichtigt einen zentralen Stellenwert im Studiendesign ein. Entsprechend war Kalorientracking eine Teilnahmevoraussetzung und es wurden zu Beginn der Studie viele Aspekte des Trackings abgefragt. Dadurch enthielt der erste Fragebogen mehr Items zum Trackingverhalten als zum Ernährungsverhalten, sodass dem Trackingverhalten im Studienkontext eine mindestens vergleichbare Bedeutung zugekommen sein könnte. Vor diesem Hintergrund ergeben sich zwei Erklärungsansätze.

Salienz des Ernährungsverhalten. Einerseits besteht die Möglichkeit, dass Messreaktivität durch diesen Fokus auf mehrere Verhaltensweisen ausblieb. Hinweise aus der Tagebuchforschung deuten darauf hin, dass messreaktive Effekte bei einer gleichzeitigen Selbstbeobachtung mehrerer Verhaltensweisen ausbleiben (Barta et al., 2012). Bei einer auf ein einzelnes Verhalten gerichteten Selbstbeobachtung, konnte Messreaktivität jedoch beobachtet werden (Barta et al., 2012; Korotitsch & Nelson-Gray, 1999). Dies wird theoretisch auf eine erhöhte Salienz und ein damit einhergehendes gesteigertes Bewusstsein für das Verhalten zurückgeführt (Barta et al., 2012). Vor diesem Hintergrund erscheint es plausibel, dass das Ernährungsverhalten in der vorliegenden Studie durch das zusätzliche Thematisieren des Trackingverhaltens keine hinreichende Salienz erlangte, um messreaktive Effekte auszulösen.

Salienz des Trackingverhalten. Neben der Möglichkeit das Ernährungsverhalten nicht ausreichend salient war, um messreaktive Effekte auszulösen, besteht auch die Möglichkeit das durch den Fokus auf Trackingverhalten in der Studie, dieses besonders salient wurde. Dadurch könnten sich messreaktive Effekte im Trackingverhalten und nicht, wie ursprünglich angenommen, im Ernährungsverhalten gezeigt haben. Eine verstärkte Fokussierung auf das Tracking könnte zu sozial oder persönlich erwünschterem Trackingverhalten führen. Das könnte sich in einer vollständigeren, und somit erhöhten, Dokumentation der Kalorieneinnahme äußern, ohne zwingend eine tatsächliche Veränderung des Ernährungsverhaltens vorauszusetzen.

Vor diesem Hintergrund lässt sich auch die Richtung der beobachteten Ergebnisse einordnen. Entgegen der ursprünglichen Erwartung wurde während der Studienwoche keine Reduktion, sondern eine numerisch erhöhte Kalorienaufnahme berichtet. Auch wenn der Effekt nicht signifikant ausfiel, sollte dies vor dem Hintergrund einer potenziell begrenzten Teststärke nicht unberücksichtigt bleiben.

Die numerische Erhöhung könnte dahingehend interpretiert werden, dass in der Woche der Studienteilnahme nicht mehr Kalorien zu sich genommen wurden, sondern mehr Kalorien berichtet wurden. Eine tatsächliche Erhöhung der Kalorienzufuhr erscheint dahingegen vor dem Hintergrund, dass insgesamt eine Tendenz geäußert wurde, weniger oder gleichviel essen zu wollen, weniger plausibel. Die höheren Kalorienangaben könnten somit auf eine Veränderung im Trackingverhalten, und nicht wie erwartet im Ernährungsverhalten, hinweisen. Insgesamt bedeutet dies, dass das Ausbleiben der postulierten niedriger berichteten Kalorieneinnahme nicht unbedingt als ein Ausbleiben von Messreaktivität zu interpretieren ist. Stattdessen könnten sich, angesichts des Studienfokus auf Trackingverhalten und des numerischen erhöhten Kalorienberichts, messreaktive Effekte im Trackingverhalten gezeigt haben.

Moderatoren der Messreaktivität

Auch wenn sich im Mittel keine bedeutsame Veränderung der Kalorienangaben zwischen den Erhebungszeiträumen zeigt, ist es theoretisch möglich, dass sich messreaktive Effekte zwischen den Teilnehmenden systematisch unterscheiden und sich dadurch insgesamt nicht als einheitlicher Effekt abbilden lassen. In diesem Fall wären vorhandene messreaktive Effekte nur unter Berücksichtigung moderierender Variablen zu sehen.

Dies ist vereinbar mit der Varianzstruktur der vorliegenden Daten. In der vorliegenden Studie wurde nur ein sehr geringer Anteil der Veränderung der Kalorienangaben durch die Studienwoche erklärt, während der überwiegende Teil der erklärten Varianz auf Unterschiede zwischen den Teilnehmenden zurückzuführen ist. Dies deutet darauf hin, dass messreaktive Effekte, sofern vorhanden, zwischen Personen unterschiedlich ausgeprägt gewesen sein könnten. Allein aus der Varianzstruktur lässt sich jedoch nicht darauf schließen, ob diese Unterschiede systematisch bedingt sind. Aufschluss darüber können Moderationsanalysen liefern. In diesen zeigten sich jedoch keine signifikanten Interaktionseffekte zwischen den untersuchten Moderatorvariablen und der Veränderung der Kalorienangaben. Im Hinblick auf eine Veränderung der berichteten Kalorien durch die Studienteilnahme, liefern die Befunde somit keine empirische Evidenz für systematische Unterschiede zwischen den Teilnehmenden.

Somit ergeben sich zwei mögliche Interpretationen: Einerseits ist es möglich, dass im Ernährungs- oder Trackingverhalten keine messreaktiven Effekte auftraten und somit kein Effekt vorlag, der hätte moderiert werden können. Andererseits ist denkbar, dass messreaktive Effekte grundsätzlich vorhanden waren, systematische Unterschiede in der vorliegenden Studie jedoch durch die ausgewählten Moderatorvariablen nicht angemessen erfasst wurden. Erkläransätze zur Einordnung der Befunde ergeben sich aus einer möglichen Operationalisierungsschwächen, Besonderheiten der Stichprobe sowie dem onlinebasierten Forscherkontakt.

Hypothese 2

In der zweiten Hypothese wurde ein Zusammenhang zwischen der Intention die Kalorieneinnahme zu verändern und der tatsächlichen Veränderung der Kalorieneinnahme in Richtung der Intention angenommen. Allerdings zeigte sich kein Interaktionseffekt der Intention. Entsprechend fanden sich keine Hinweise darauf, dass die Intention die Veränderung der Kalorienangaben vor und während der Studienteilnahme moderierte. Somit stützen die Befunde der vorliegenden Studie die zweite Hypothese nicht.

Allerdings zeigte sich ein signifikanter Haupteffekt der Intention. Die positive Richtung dieses Effekts weist darauf hin, dass das geäußerte Kalorienziel systematisch mit dem durchschnittlich berichteten Kalorienniveau zusammenhing. Personen, deren Intention stärker in Richtung einer Erhöhung der Kalorienaufnahme ausgerichtet war, berichteten im Mittel höhere Kalorienwerte, während eine stärkere Ausrichtung in Richtung einer Reduktion mit niedrigeren Kalorienangaben einherging. Dieser Zusammenhang lässt die Annahme zu, dass sich die Teilnehmenden zumindest teilweise bereits entsprechend ihres Kalorienziels ernähren. Daraus ergeben sich zwei Implikationen.

Einerseits wirft dieser Befund die Frage auf, inwieweit das Ernährungsziel in der vorliegenden Studie tatsächlich ein geeignetes Maß zur Erfassung der theoretisch postulierten Diskrepanz zwischen tatsächlicher und persönlich angestrebter Kalorieneinnahme darstellt. Die Operationalisierung basiert auf der Annahme, dass Personen ausgehend von ihrem aktuellen Kalorienniveau angeben, wie stark und in welche Richtung sie ihre Kalorienaufnahme verändern möchten. Wenn sich Personen jedoch bereits vor Studienteilnahme in Richtung ihres Ziels ernähren, und das möglicherweise auch in unterschiedlichem Ausmaß, bildet die Intentionsangabe keine verlässliche Grundlage mehr, um Unterschiede im Ausmaß der Diskrepanz zwischen Personen abzubilden. Die fehlende Moderation der Intention könnte demnach darauf zurückzuführen sein, dass sie die Diskrepanz nicht passend abbildet und den postulierten Dissonanzmechanismus somit nicht erfasst.

Andererseits könnte der Zusammenhang zwischen Intention und durchschnittlichem Kalorienniveau mit einem begrenzten Veränderungsspielraum einhergehen, der sich im Sinne eines Deckeneffekts äußert und somit eine Veränderung der Kalorieneinnahme während der Studienteilnahme einschränken könnte. Falls sich Personen bereits nahe an ihrem individuellen Zielbereich befinden, ist der Spielraum für weitere Anpassungen geringer als bei Personen, deren aktuelles Verhalten noch deutlich vom Zielverhalten abweicht.

Vergleichbare Befunde sind in einer Studie von König et al. (2025) zu sehen, in welcher Personen mit niedrigerem Ausgangsniveau der Schrittzahl stärkere Zunahmen im Studienverlauf zeigten als Personen mit bereits hohem Ausgangsniveau. Dies wurde dahingehend interpretiert, dass bei Personen mit niedrigem Ausgangsniveau eine größere Diskrepanz zum Zielverhalten, und damit ein größerer Veränderungsspielraum, bestand.

Vor diesem Hintergrund ist zudem zu berücksichtigen, dass in der vorliegenden Studie ausschließlich Personen rekrutiert wurden, die bereits vor Studienbeginn ihre Ernährung aufzeichneten. Da das Aufzeichnen von Ernährung an sich schon zu einer Verhaltensveränderung führen kann (Berry et al., 2021; Burke et al., 2011) besteht die Möglichkeit, dass die Teilnehmenden das Potenzial ihrer Ernährungsveränderung bereits zu einem gewissen Maß ausgeschöpft haben. Analog zu den Befunden bei Personen mit hohem Ausgangsniveau könnte der Veränderungsspielraum in der vorliegenden Stichprobe somit insgesamt begrenzt gewesen sein. In diesem Fall wäre es plausibel, dass messreaktive Effekte im Ernährungsverhalten nicht oder nur in geringem Ausmaß auftraten.

Weiters ist zu beachten, dass sich, falls sich messreaktive Effekte primär im Trackingverhalten und nicht im Ernährungsverhalten gezeigt haben, auch der theoretisch relevante Moderationsmechanismus verschiebt. In diesem Fall wäre nicht die Diskrepanz zwischen tatsächlichem und idealem Ernährungsverhalten der zentrale Moderator, sondern vielmehr die Diskrepanz zwischen tatsächlichem und idealem Trackingverhalten. Somit stellt auch dies eine mögliche Erklärung für die fehlende Interaktion mit der Ernährungsintention dar.

Hypothese 3

In der dritten Hypothese wurde angenommen, dass eine stärker wahrgenommene Forscherbeobachtung mit stärkeren Veränderungen der berichteten Kalorienaufnahme einhergeht. Insgesamt zeigten die Ergebnisse jedoch keinen Zusammenhang zwischen der wahrgenommenen Forscherbeobachtung und einer Veränderung der berichteten Kalorienwerte. Die dritte Hypothese wird somit durch die Befunde nicht gestützt. Gleichzeitig berichteten die Teilnehmenden insgesamt eine Tendenz, sich der Forscherbeobachtung bewusst gewesen zu sein. Somit, scheint naheliegend, dass sich die Teilnehmenden zwar der Forscherbeobachtung bewusst waren, dies jedoch keine Veränderung im Verhalten zur Folge hatte.

Dies könnte unter anderem daran liegen, dass das Format der Studie ausschließlich online durchgeführt wurde, eine hohe Anonymität bestand und kein direkter Forscherkontakt im Studiendesign vorgesehen war. So legen Befunde aus der Tagebuchforschung nahe, dass Effekte sozialer Erwünschtheit insbesondere dann stärker ausfallen, wenn Erhebungssituationen weniger anonym sind oder in persönlichem Kontakt stattfinden (Barta et al., 2012). Weiters zeigt eine Meta-Analyse zum Frage-Verhaltenseffekt, dass Messreaktivität verstärkt in Laborsettings verglichen mit Onlinesettings auftritt (Wilding et al., 2016). Diese Annahme wird weiters durch die Studie von König et al. (2025) gestützt, die ebenfalls weder Messreaktivität noch eine Moderation durch Forscherbeobachtung zeigten. Sowohl in der vorliegenden Studie als auch in der Studie von König et al. (2025) erfolgte die Erhebung ausschließlich online und ohne persönlichen Forscherkontakt. Im Gegensatz dazu,

konnte in Studien mit explizit erwähntem Forscherkontakt Messreaktivität beobachtet werden (Arigo & König, 2024; Baumann et al., 2018).

Dies legt nahe, dass die Art des Forscherkontakts möglicherweise eine Rolle in der Entstehung von Messreaktivität spielt. Die onlinebasierte Forscherbeobachtung in der vorliegenden Studie reichte somit möglicherweise nicht aus, um soziale Erwünschtheit in einem Ausmaß zu aktivieren, das zu messbaren Verhaltensveränderungen führte.

Eine weitere Erklärung lässt sich aus der Zusammensetzung der Stichprobe ableiten. Da diese ausschließlich aus Personen besteht, die Ernährung tracken, besteht die Möglichkeit, dass es sich hier um eine recht ernährungsbewusste Stichprobe handelt, welche sich möglicherweise schon weitgehend sozial erwünscht ernährt. Somit könnte die fehlende Moderation darauf zurückzuführen sein, dass die Forscherbeobachtung zwar wahrgenommen wurde, aber nicht zu einer Verhaltensveränderung führte, da nicht die Notwendigkeit bestand, das Verhalten zu verändern, um positiv wahrgenommen zu werden.

Exploration

Weiters wurde exploriert, ob ein Zusammenhang zwischen der Bewusstheit einer durch die Studienteilnahme ausgelösten Ernährungsveränderung und der tatsächlichen Veränderung des Ernährungsverhaltens besteht. In der vorliegenden Studie zeigte sich kein signifikanter Interaktionseffekt bezüglich der Bewusstheit über eine Verhaltensveränderung. Somit weisen die Befunde nicht auf eine moderierende Rolle dieser auf Messreaktivität hin.

Dies lässt sich vor dem Hintergrund des Befundes, dass auch die wahrgenommene Forscherbeobachtung keine Messreaktivität moderierte, plausibel einordnen. Die Grundlage für die Exploration stellt die Annahme dar, dass Impression Management, als Form sozialer Erwünschtheit, Messreaktivität auslöst (Barta et al., 2012). Die Bewusstheit über eine Verhaltensveränderung wurde in der vorliegenden Studie als Hinweis auf Impression Management theoretisiert. Demnach ist kein moderierender Effekt der Bewusstheit über eine Verhaltensveränderung zu erwarten, wenn keine soziale Erwünschtheit aktiviert wurde.

Da in der vorherigen Hypothese keine Moderation der wahrgenommenen Forscherbeobachtung vorlag, liegen keine Hinweise auf eine Aktivierung von sozialer Erwünschtheit in dieser Studie vor. Dementsprechend könnte die fehlende Moderation der Bewusstheit über eine Verhaltensveränderung auf Messreaktivität, dahingehend zurückgeführt werden, dass möglicherweise keine soziale Erwünschtheit in der vorliegenden Studie aktiviert wurde.

Zusammenfassende Einordnung der Befunde

Die Befunde zeigen, dass unter den gegebenen Bedingungen keine messreaktiven Effekte beobachtet wurden und auch die untersuchten Moderatoren keine systematischen Unterschiede in einer Kalorienveränderung zwischen den Personen erklären konnten. Insgesamt sind in der vorliegenden Studie daher keine Hinweise zu finden, dass die Validität digitalen Ernährungstrackings durch Messreaktivität eingeschränkt ist.

Die vorliegenden Befunde lassen sich im Vergleich zu Studien, in denen Messreaktivität beobachtet wurde, dahingehend einordnen, dass das Auftreten messreaktiver Effekte möglicherweise abhängig von den jeweiligen Studienbedingungen ist. Während in bestimmten Studien Messreaktivität berichtet wird, blieb sie unter den spezifischen Bedingungen der vorliegenden Untersuchung aus. Konkret unterschied sich die vorliegende Studie in der untersuchten Verhaltensweise und der Isolierung des Forschungskontext. Im Zuge dessen ergab sich ein potenziell uneindeutiger Forschungsfokus sowie eine spezifische Stichprobenzusammensetzung. Zusätzlich erfolgte der Forscherkontakt in der vorliegenden Studie rein onlinebasiert. All diese spezifischen Kontexte könnten dementsprechend ausschlaggebend für das Ausbleiben von Messreaktivität in dieser Studie sein. Insgesamt legt dies nahe, dass die Entstehung von Messreaktivität als kontextabhängig zu betrachten ist.

Stärken

Eine zentrale Stärke der vorliegenden Studie besteht darin, dass Messreaktivität in Ernährungsverhalten untersucht und somit eine Forschungslücke in der digitalen Gesundheitsforschung aufgegriffen wird, welche sich bisher größtenteils auf Bewegungsverhalten konzentriert (König et al., 2022). Dadurch trägt die Studie zum Verständnis der Validität digitalen Ernährungstrackings als Erhebungsinstrument bei.

Darüber hinaus zeichnet sich die Studie durch den Einbezug verschiedener Moderatoren aus. Dadurch können systematische Unterschiede in möglichen messreaktiven Effekten zwischen den Personen berücksichtigen werden und so Aufschluss über mögliche theoretisierte Wirkmechanismen geben.

Eine weitere Stärke liegt in der Weiterverwendung der eigenen Ernährungsapp. Dies hebt die vorliegende Studie von den meisten anderen Studien zu Messreaktivität in der digitalen Gesundheitsforschung ab (für einen Überblick siehe König et al., 2022). Einerseits wird dadurch ein Vergleich der Kalorienangaben vor und während der Studienteilnahme innerhalb derselben Personen ermöglicht. Veränderungen konnten somit auf individueller Ebene abgebildet und zugleich Unterschiede zwischen den Teilnehmenden berücksichtigt werden. Dieser Vorteil wurde methodisch durch den Einsatz von Multilevel-Modellen ergänzt, die es ermöglichten, unterschiedliche Ausgangsniveaus sowie individuelle Veränderungen im Ernährungsverhalten zu berücksichtigen.

Andererseits ermöglichte die Weiterverwendung der bereits genutzten Ernährungsapp, messreaktive Effekte zu untersuchen, die allein auf den Forschungskontext zurückzuführen sind und damit vom Messkontext getrennt betrachtet werden konnten (König et al., 2025). Dies erlaubt eine erweiterte Einordnung der Bedingungen, unter denen Messreaktivität entsteht, und ergänzt damit den bisherigen Forschungsstand, in dem Mess- und Forschungskontext meist gemeinsam betrachtet werden.

Limitationen

Die zuvor dargestellten Stärken der Studie gehen zugleich mit mehreren Einschränkungen einher, die bei der Interpretation der Ergebnisse berücksichtigt werden sollten. Eine zentrale Einschränkung der vorliegenden Studie besteht darin, dass die beobachteten Befunde nicht eindeutig einem einzelnen Aspekt des Studiendesigns zugeordnet werden können. Verglichen mit Studien in der digitalen Gesundheitsforschung, in denen Messreaktivität nachgewiesen wurde (für einen Überblick siehe König et al., 2022), unterscheidet sich die vorliegende Untersuchung in mehreren Punkten. Aufgrund dieser gleichzeitigen Abweichungen bleibt unklar, auf welchen dieser Aspekte die Abwesenheit messreaktiver Effekte im Vergleich zu Studien mit nachgewiesener Messreaktivität zurückzuführen ist. Diese Limitation erschwert somit eine eindeutige Einordnung der Befunde im Vergleich zu bestehender Messreaktivitätsforschung.

Darüber hinaus ist die Generalisierbarkeit der Ergebnisse durch die Zusammensetzung der Stichprobe eingeschränkt. Ergebnisse zum Ernährungsverhalten von Personen, die regelmäßig ihre Ernährung tracken, sind möglicherweise spezifisch für diese Personengruppe. Darüber hinaus setzte sich die Stichprobe überwiegend aus Frauen mit relativ hohem sozioökonomischem Status zusammen. Diese Stichprobenmerkmale schränken die Übertragbarkeit der Ergebnisse auf die Allgemeinbevölkerung ein.

Unabhängig von der inhaltlichen Zusammensetzung der Stichprobe stellt die Größe der Stichprobe eine weitere Limitation dar. So könnte die Anzahl der Teilnehmenden zu gering sein, um kleine, aber dennoch relevante Effekte zuverlässig zu entdecken. Entsprechend kann nicht ausgeschlossen werden, dass vorhandene Effekte aufgrund begrenzter statistischer Power unentdeckt geblieben sind.

Neben diesen Aspekten ergibt sich eine weitere Einschränkung aus dem potenziell uneindeutigen Forschungsfokus der Studie. Neben dem Ernährungsverhalten wurde auch das Trackingverhalten explizit thematisiert. Der Studienfokus auf zwei Verhaltensweisen könnte die Entstehung messreaktiver Effekte beeinflusst haben (Barta et al., 2012; Korotitsch & Nelson-Gray, 1999). Entsprechend sind die Befunde dieser Studie möglicherweise nur eingeschränkt auf Kontexte übertragbar, die eindeutig Ernährungsverhalten als Fokus haben.

Eine weitere Limitation betrifft die Verlässlichkeit der erhobenen Ernährungsdaten. Durch den fehlenden Zugriff auf die Rohdaten der verwendeten Ernährungsapps, musste in der vorliegenden Studie angenommen werden, dass die Teilnehmenden ihre Kalorienangaben zeitnah erfasst und korrekt übertragen haben. Es lässt sich jedoch nicht überprüfen, in welchem Ausmaß das Tracking tatsächlich unmittelbar erfolgte. So ist beispielsweise denkbar, dass fehlende Einträge nachträglich für den Fragebogen geschätzt wurden. Folglich besteht die Möglichkeit, dass Daten durch Recall-Bias verzerrt wurden. Zudem können Übertragungsfehler zwischen App und Befragung nicht ausgeschlossen werden.

Schließlich besteht eine weitere Einschränkung in der Operationalisierung der Diskrepanz zwischen tatsächlichem und idealem Ernährungsverhalten. In der vorliegenden Studie wurde diese Diskrepanz über die Intention erfasst, es besteht jedoch die Möglichkeit, dass diese kein passendes Maß dafür darstellt. Dadurch kann nicht eindeutig zwischen einer unpassenden Operationalisierung und dem tatsächlichen Ausbleiben des postulierten Dissonanzmechanismus unterschieden werden. Dies stellt eine Einschränkung für die Interpretation der Befunde dar.

Implikationen

Unter den Bedingungen der vorliegenden Studie finden sich keine Hinweise auf eine relevante Beeinträchtigung der Validität digitalen Ernährungstrackings durch Messreaktivität. Vor diesem Hintergrund kann digitales Ernährungstracking im Rahmen vergleichbarer Studiensettings im Hinblick auf Messreaktivität als zuverlässiges Erhebungsinstrument betrachtet werden. Gleichzeitig ist die bisherige Forschungslage zu Messreaktivität im digitalen Ernährungstracking begrenzt, sodass weiterführende Untersuchungen erforderlich sind, um diese Einschätzung empirisch abzusichern.

Da Messreaktivität in anderen Studien beobachtet wurde, unter den spezifischen Bedingungen der vorliegenden Untersuchung jedoch ausblieb, spricht dies dafür, dass messreaktive Effekte kontextabhängig auftreten. Daraus ergibt sich die Notwendigkeit, Messreaktivität in der digitalen Gesundheitsforschung in Zukunft gezielt und systematisch unter variierenden Bedingungen zu untersuchen. Erkenntnisse über die spezifischen Bedingungen, unter denen Messreaktivität entsteht, können dazu beitragen, das Risiko messreaktiver Verzerrungen einzuschätzen oder durch eine dementsprechende Gestaltung des Studiendesigns zu reduzieren (French & Sutton, 2010).

Ein zentraler Ansatzpunkt hierfür stellt die Untersuchung von Messreaktivität in unterschiedlichen Verhaltensweisen dar. Die Befunde der vorliegenden Studie geben Hinweise darauf, dass Ernährungsverhalten möglicherweise weniger anfällig für Messreaktivität ist als Bewegungsverhalten. Zukünftige Forschung könnte Ernährungs- und Bewegungsverhalten bzw. andere Verhaltensweisen unter möglichst vergleichbaren Rahmenbedingungen untersuchen, um verhaltensspezifische Unterschiede messreaktiver Effekte besser einordnen zu können. Innerhalb des Ernährungsverhaltens wäre zudem eine Differenzierung nach der Veränderbarkeit einzelner ernährungsspezifischer Verhaltensweisen sinnvoll, da sich messreaktive Effekte möglicherweise verstärkt bei modifizierbaren Aspekten zeigen (König et al., 2022).

Darüber hinaus stellt eine differenzierte Betrachtung des Forschungs- und Messkontexts ein relevanten Forschungsaspekt dar (König et al., 2025). So könnten zukünftige Studien diese beiden Kontexte systematisch variieren und sowohl isoliert als auch

in Kombination untersuchen. Ein solcher Vergleich würde es ermöglichen die Rolle des Messkontexts und des Forschungskontexts in der Entstehung der Messreaktivität besser zu verstehen.

Eine weitere Implikation betrifft die Eindeutigkeit des wahrgenommenen Forschungsfokus. Zukünftige Studien könnten die Salienz des untersuchten Verhaltens gezielt variieren, indem entweder ein einzelnes Verhalten klar in den Mittelpunkt gestellt oder mehrere Verhaltensweisen parallel thematisiert werden. Dadurch ließe sich prüfen, ob messreaktive Effekte eher dann auftreten, wenn sich der vermittelte Forschungsfokus eindeutig auf ein spezifisches Verhalten richtet. Ergänzend könnte erhoben werden, welches Verhalten aus Sicht der Teilnehmenden tatsächlich im Fokus der Studie stand.

Auch die Art des Forscherkontakts stellt einen relevanten Ansatzpunkt für zukünftige Forschung dar. Da in der vorliegenden Studie die onlinebasierte Forscherbeobachtung zwar wahrgenommen wurde, jedoch nicht mit Verhaltensänderungen einherging, erscheint es sinnvoll, unterschiedliche Formen des Forscherkontakts systematisch zu vergleichen. Der Vergleich rein onlinebasierter Studien mit Designs, die persönlichen Kontakt beinhalten, könnte dazu beitragen, eine potenzielle Rolle des spezifischen Forscherkontakts bei der Entstehung messreaktiver Effekte genauer zu untersuchen.

Aus den methodischen Einschränkungen der vorliegenden Studie ergeben sich weitere Implikationen für zukünftige Forschung. Einerseits sollte Messreaktivität im Kontext digitalen Ernährungstrackings an repräsentativeren Stichproben untersucht werden, um die Übertragbarkeit der Befunde besser einschätzen zu können. Zudem sollten zukünftige Studien über eine ausreichende Stichprobengröße verfügen, um auch kleine Effekte verlässlich nachweisen zu können. Andererseits sollte die Diskrepanz zwischen tatsächlichem und idealem Verhalten adäquat operationalisiert werden. Dies könnte beispielsweise über die Erhebung eines konkreten Kalorienzielwerts erfolgen, der anschließend mit dem tatsächlichen Kalorienkonsum verglichen wird. Mit einem zuverlässigen Maß für die Diskrepanz ließe sich der postulierte Dissonanzmechanismus differenzierter untersuchen. Darüber hinaus erscheint es sinnvoll, dass zukünftige Studien nach Möglichkeit mit Rohdaten der verwendeten Ernährungsapps arbeiten, um die Datenqualität und damit die Verlässlichkeit der Befunde besser absichern zu können.

Conclusio

Die vorliegende Studie liefert keine Hinweise darauf, dass die Validität digitalen Ernährungstrackings unter den studienspezifischen Bedingungen durch Messreaktivität beeinträchtigt wird. Zukünftige Untersuchungen sind jedoch erforderlich, um diese Einschätzung empirisch abzusichern und die Entstehungsbedingungen und Wirkmechanismen von Messreaktivität differenzierter einschätzen zu können. Insbesondere das systematische Untersuchen verschiedener Kontexte, die das Risiko von Messreaktivität

reduzieren oder begünstigen, stellen hierfür einen zentralen Ansatz dar. So kann zu einer verlässlicheren Forschung in Ernährungsbereich beigetragen werden.

Literaturverzeichnis

- Arigo, D., & König, L. M. (2024). Examining reactivity to the measurement of physical activity and sedentary behavior among women in midlife with elevated risk for cardiovascular disease. *Psychology & Health*, 39(3), 319–335.
<https://doi.org/10.1080/08870446.2022.2055024>
- Barta, W. D., Tennen, H., & Litt, M. D. (2012). *Measurement Reactivity in Diary Research*. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 108–123). The Guilford Press.
- Baumann, S., Groß, S., Voigt, L., Ullrich, A., Weymar, F., Schwaneberg, T., Dörr, M., Meyer, C., John, U., & Ulbricht, S. (2018). Pitfalls in accelerometer-based measurement of physical activity: The presence of reactivity in an adult population. *Scandinavian Journal of Medicine & Science in Sports*, 28(3), 1056–1063.
<https://doi.org/10.1111/sms.12977>
- Berry, R., Kassavou, A., & Sutton, S. (2021). Does self-monitoring diet and physical activity behaviors using digital technology support adults with obesity or overweight to lose weight? A systematic literature review with meta-analysis. *Obesity Reviews*, 22(10), e13306. <https://doi.org/10.1111/obr.13306>
- Börger, T. (2012). *Social Desirability and Environmental Valuation*. Peter Lang.
<https://doi.org/10.3726/978-3-653-01583-6>
- Bravata, D. M., Smith-Spangler, C., Sundaram, V., Gienger, A. L., Lin, N., Lewis, R., Stave, C. D., Olkin, I., & Sirard, J. R. (2007). Using pedometers to increase physical activity and improve health: a systematic review. *JAMA*, 298(19), 2296–2304.
<https://doi.org/10.1001/jama.298.19.2296>
- Burke, L. E., Wang, J., & Sevvick, M. A. (2011). Self-monitoring in weight Loss: A systematic review of the literature. *Journal of the American Dietetic Association*, 111(1), 92–102.
<https://doi.org/10.1016/j.jada.2010.10.008>
- Burrows, T. L., Ho, Y. Y., Rollo, M. E., & Collins, C. E. (2019). Validity of dietary assessment methods when compared to the method of doubly labeled water: A systematic review in adults. *Frontiers in Endocrinology*, 10. <https://doi.org/10.3389/fendo.2019.00850>
- Butryn, M. L., Godfrey, K. M., Martinelli, M. K., Roberts, S. R., Forman, E. M., & Zhang, F. (2020). Digital self-monitoring: Does adherence or association with outcomes differ by self-monitoring target? *Obesity Science & Practice*, 6(2), 126–133.
<https://doi.org/10.1002/osp4.391>
- Carver, C. S., & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality–social, clinical, and health psychology. *Psychological Bulletin*, 92(1), 111–135. <https://doi.org/10.1037/0033-2909.92.1.111>

- Cena, H., & Calder, P. C. (2020). Defining a healthy diet: Evidence for the role of contemporary dietary patterns in health and disease. *Nutrients*, 12(2), 334. <https://doi.org/10.3390/nu12020334>
- Conner, M., Sandberg, T., Nekitsing, C., Hutter, R., Wood, C., Jackson, C., Godin, G., & Sheeran, P. (2017). Varying cognitive targets and response rates to enhance the question-behaviour effect: An 8-arm randomized controlled trial on influenza vaccination uptake. *Social Science & Medicine*, 180, 135–142. <https://doi.org/10.1016/j.socscimed.2017.03.037>
- Cook, P., Schmiede, S., McClean, M., Aagaard, L., & Kahook, M. (2012). Practical and analytic issues in the electronic assessment of adherence. *Western Journal of Nursing Research*, 34(5), 598–620. <https://doi.org/10.1177/0193945911427153>
- Dholakia, U. M., & Malhotra, N. K. (2010). A critical review of question–behavior effect research. In *Review of Marketing Research* (Vol. 7, pp. 145–197). Emerald Group Publishing Limited. [https://doi.org/10.1108/S1548-6435\(2010\)0000007009](https://doi.org/10.1108/S1548-6435(2010)0000007009)
- Fitzsimons, G. J., & Williams, P. (2000). Asking questions can change choice behavior: Does it do so automatically or effortfully? *Journal of Experimental Psychology: Applied*, 6(3), 195–206. <https://doi.org/10.1037/1076-898X.6.3.195>
- Freijy, T., & Kothe, E. J. (2013). Dissonance-based interventions for health behaviour change: A systematic review. *British Journal of Health Psychology*, 18(2), 310–337. <https://doi.org/10.1111/bjhp.12035>
- French, D. P., & Sutton, S. (2010). Reactivity of measurement in health psychology: How much of a problem is it? What can be done about it? *British Journal of Health Psychology*, 15(3), 453–468. <https://doi.org/10.1348/135910710X492341>
- Hartig, F. (2022). *DHARMA: Residual diagnostics for hierarchical (multi-level / mixed) regression models* (R package version 0.7.4). <https://cran.r-project.org/package=DHARMA>
- Illner, A.-K., Freisling, H., Boeing, H., Huybrechts, I., Crispim, S., & Slimani, N. (2012). Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. *International Journal of Epidemiology*, 41(4), 1187–1203. <https://doi.org/10.1093/ije/dys105>
- Koene, R. J., Prizment, A. E., Blaes, A., & Konety, S. H. (2016). Shared risk factors in cardiovascular disease and cancer. *Circulation*, 133(11), 1104–1114. <https://doi.org/10.1161/CIRCULATIONAHA.115.020406>
- Korotitsch, W. J., & Nelson-Gray, R. O. (1999). An overview of self-monitoring research in assessment and treatment. *Psychological Assessment*, 11(4), 415–425. (1999-01917-002). <https://doi.org/10.1037/1040-3590.11.4.415>

- König, L. M., Allmeta, A., Christlein, N., Van Emmenis, M., & Sutton, S. (2022). A systematic review and meta-analysis of studies of reactivity to digital in-the-moment measurement of health behaviour. *Health Psychology Review*, 16(4), 551–575.
- König, L. M., Pasko, K., Baga, K., Harsora, R., & Arigo, D. (2025). Isolating the role of researcher observation on reactivity to the measurement of physical activity. *Applied Psychology: Health and Well-Being*, 17(1), e12630.
<https://doi.org/10.1111/aphw.12630>
- König, L. M., & Ruf, A. (2024). Does researcher observation influence food intake in daily life? A data donation study. Abgerufen von
https://osf.io/b6cyz/?view_only=c7c7060793e04eee9c5e982352741961
- König, L. M., Van Emmenis, M., Nurmi, J., Kassavou, A., & Sutton, S. (2022). Characteristics of smartphone-based dietary assessment tools: A systematic review. *Health Psychology Review*, 16(4), 526–550. <https://doi.org/10.1080/17437199.2021.2016066>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017).
lmerTest package: Tests in linear mixed effects models.
Journal of Statistical Software, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Labhart, F., Tarsetti, F., Bornet, O., Santani, D., Truong, J., Landolt, S., Gatica-Perez, D., & Kuntsche, E. (2020). Capturing drinking and nightlife behaviours and their social and physical context with a smartphone application – Investigation of users' experience and reactivity. *Addiction Research & Theory*, 28(1), 62–75.
<https://doi.org/10.1080/16066359.2019.1584292>
- Lackner, S. (2025). *Messreaktivität bei der digitalen Ernährungserfassung: verändert Forscherbeobachtung den alltäglichen Proteinkonsum?* (Masterarbeit). Universität Wien.
- Lichtenberger, M. (2025). Messreaktivität beim Fettkonsum: Der Einfluss von Forscherbeobachtung und die moderierende Rolle von Geschlecht und Gewissenhaftigkeit. (Unveröffentlichte Masterarbeit). Universität Wien.
- MacNeill, V., Foley, M., Quirk, A., & McCambridge, J. (2016). Shedding light on research participation effects in behaviour change trials: A qualitative study examining research participant experiences. *BMC Public Health*, 16(1), 91.
<https://doi.org/10.1186/s12889-016-2741-6>
- Maher, J. P., Arigo, D., Baga, K., Salvatore, G. M., Pasko, K., Hudgins, B. L., & König, L. M. (2024). Measurement reactivity in ecological momentary assessment studies of movement-related behaviors. *Journal for the Measurement of Physical Behaviour*, 7(1), jmpb.2023-0035. <https://doi.org/10.1123/jmpb.2023-0035>

- McCambridge, J., & Kypri, K. (2011). Can simply answering research questions change behaviour? Systematic review and meta analyses of brief alcohol intervention trials. *PLoS One*, 6(10), e23748. <https://doi.org/10.1371/journal.pone.0023748>
- McCambridge, J., Witton, J., & Elbourne, D. R. (2014). Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. *Journal of Clinical Epidemiology*, 67(3), 267–277. <https://doi.org/10.1016/j.jclinepi.2013.08.015>
- Michie, S., Abraham, C., Whittington, C., McAteer, J., & Gupta, S. (2009). Effective techniques in healthy eating and physical activity interventions: A meta-regression. *Health Psychology*, 28(6), 690–701. (2009-20990-006). <https://doi.org/10.1037/a0016136>
- Miles, L. M., Rodrigues, A. M., Sniehotta, F. F., & French, D. P. (2020). Asking questions changes health-related behavior: An updated systematic review and meta-analysis. *Journal of Clinical Epidemiology*, 123, 59–68. <https://doi.org/10.1016/j.jclinepi.2020.03.014>
- Moos, R. H. (2008). Context and mechanism of reactivity to assessment and treatment. *Addiction*, 103(2), 249–250. <https://doi.org/10.1111/j.1360-0443.2007.02123.x>
- Naska, A., Lagiou, A., & Lagiou, P. (2017). *Dietary assessment methods in epidemiological research: Current state of the art and future prospects* (No. 6:926). F1000Research. <https://doi.org/10.12688/f1000research.10703.1>
- Page, E. J., Massey, A. S., Prado-Romero, P. N., & Albadawi, S. (2020). The Use of self-monitoring and technology to increase physical activity: A review of the literature. *Perspectives on Behavior Science*, 43(3), 501–514. <https://doi.org/10.1007/s40614-020-00260-0>
- Pek, J., Wong, O., & Wong, A. C. M. (2018). How to address non-normality: A taxonomy of approaches, reviewed, and illustrated. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.02104>
- Pierce-Messick, Z., & Corbit, L. H. (2021). Problematic eating as an issue of habitual control. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 110, 110294. <https://doi.org/10.1016/j.pnpbp.2021.110294>
- Posit Software, PBC. (2025). *RStudio* (Version 2025.05.0). <https://posit.co/>
- Poulton, A., Pan, J., Bruns Jr, L. R., Sinnott, R. O., & Hester, R. (2019). A smartphone app to assess alcohol consumption behavior: Development, compliance, and reactivity. *JMIR mHealth and uHealth*, 7(3), e11157. <https://doi.org/10.2196/11157>
- R Core Team. (2025). *R: A language and environment for statistical computing* (Version 4.5.0). R Foundation for Statistical Computing. <https://www.R-project.org/>

- Rebro, S. M., Patterson, R. E., Kristal, A. R., & Cheney, C. L. (1998). The effect of keeping food records on eating patterns. *Journal of the American Dietetic Association*, 98(10), 1163–1165.
- Reedy, J., Krebs-Smith, S. M., Miller, P. E., Liese, A. D., Kahle, L. L., Park, Y., & Subar, A. F. (2014). Higher diet quality is associated with decreased risk of all-cause, cardiovascular disease, and cancer mortality among older adults. *The Journal of Nutrition*, 144(6), 881–889. <https://doi.org/10.3945/jn.113.189407>
- Robinson, E., Hardman, C. A., Halford, J. C., & Jones, A. (2015). Eating under observation: A systematic review and meta-analysis of the effect that heightened awareness of observation has on laboratory measured energy intake. *The American Journal of Clinical Nutrition*, 102(2), 324–337. <https://doi.org/10.3945/ajcn.115.111195>
- Robinson, E., Kersbergen, I., Brunstrom, J. M., & Field, M. (2014). I'm watching you. Awareness that food consumption is being monitored is a demand characteristic in eating-behaviour experiments. *Appetite*, 83, 19–25. <https://doi.org/10.1016/j.appet.2014.07.029>
- Rollo, M. E., Williams, R. L., Burrows, T., Kirkpatrick, S. I., Bucher, T., & Collins, C. E. (2016). What are they really eating? A review on new approaches to dietary intake assessment and validation. *Current Nutrition Reports*, 5(4), 307–314. <https://doi.org/10.1007/s13668-016-0182-6>
- Sanderman, R., & Morgan, K. (Eds.). (2025). *The Routledge International Handbook of Health Psychology: Global and Contemporary Issues*. Routledge. <https://doi.org/10.4324/9781003318026>
- Sanders, J. P., Loveday, A., Pearson, N., Edwardson, C., Yates, T., Biddle, S. J., & Esliger, D. W. (2016). Devices for self-monitoring sedentary time or physical activity: A scoping review. *Journal of Medical Internet Research*, 18(5), e90. <https://doi.org/10.2196/jmir.5373>
- Schatzkin, A., Kipnis, V., Carroll, R. J., Midthune, D., Subar, A. F., Bingham, S., Schoeller, D. A., Troiano, R. P., Freedman, L. S., & Freedman, L. S. (2003). A comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological cohort study: Results from the biomarker-based Observing Protein and Energy Nutrition (OPEN) study. *International Journal of Epidemiology*, 32(6), 1054–1062. <https://doi.org/10.1093/ije/dyg256>
- Sharp, D. B., & Allman-Farinelli, M. (2014). Feasibility and validity of mobile phones to assess dietary intake. *Nutrition*, 30(11–12), 1257–1266. <https://doi.org/10.1016/j.nut.2014.02.020>
- Spangenberg, E. R., Sprott, D. E., Grohmann, B., & Smith, R. J. (2003). Mass-communicated prediction requests: Practical application and a cognitive dissonance explanation for

- self-prophecy. *Journal of Marketing*, 67(3), 47–62.
<https://doi.org/10.1509/jmkg.67.3.47.18659>
- Sprott, D. E., Spangenberg, E. R., Block, L. G., Fitzsimons, G. J., Morwitz, V. G., & Williams, P. (2006). *The QBE. What we know and where we go from here*.
- Stice, E., Shaw, H., Becker, C. B., & Rohde, P. (2008). Dissonance-based interventions for the prevention of eating disorders: Using persuasion principles to promote health. *Prevention Science*, 9(2), 114–128. <https://doi.org/10.1007/s11121-008-0093-x>
- Subar, A. F., Kipnis, V., Troiano, R. P., Midthune, D., Schoeller, D. A., Bingham, S., Sharbaugh, C. O., Trabulsi, J., Runswick, S., Ballard-Barbash, R., Sunshine, J., & Schatzkin, A. (2003). Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: The OPEN study. *American Journal of Epidemiology*, 158(1), 1–13. <https://doi.org/10.1093/aje/kwg092>
- Sutton, S., Kinmonth, A.-L., Hardeman, W., Hughes, D., Boase, S., Prevost, A. T., Kellar, I., Graffy, J., Griffin, S., & Farmer, A. (2014). Does electronic monitoring influence adherence to medication? Randomized controlled trial of measurement reactivity. *Annals of Behavioral Medicine*, 48(3), 293–299. <https://doi.org/10.1007/s12160-014-9595-x>
- Teasdale, N., Elhussein, A., Butcher, F., Piernas, C., Cowburn, G., Hartmann-Boyce, J., Saksena, R., & Scarborough, P. (2018). Systematic review and meta-analysis of remotely delivered interventions using self-monitoring or tailored feedback to change dietary behavior. *The American Journal of Clinical Nutrition*, 107(2), 247–256.
<https://doi.org/10.1093/ajcn/nqx048>
- Thompson, F. E., & Subar, A. F. (2017). Dietary Assessment Methodology. In *Nutrition in the Prevention and Treatment of Disease* (S. 5–48). Elsevier.
<https://doi.org/10.1016/B978-0-12-802928-2.00001-1>
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.
- Turner-McGrievy, G. M., Dunn, C. G., Wilcox, S., Boutté, A. K., Hutto, B., Hoover, A., & Muth, E. (2019). Defining adherence to mobile dietary self-monitoring and assessing tracking over time: Tracking at least two eating occasions per day is best marker of adherence within two different mobile health randomized weight loss interventions. *Journal of the Academy of Nutrition and Dietetics*, 119(9), 1516–1524.
<https://doi.org/10.1016/j.jand.2019.03.012>
- Wilde, M. H., & Garvin, S. (2007). A concept analysis of self-monitoring. *Journal of Advanced Nursing*, 57(3), 339–350. <https://doi.org/10.1111/j.1365-2648.2006.04089.x>
- Wilding, S., Conner, M., Sandberg, T., Prestwich, A., Lawton, R., Wood, C., Miles, E., Godin, G., & Sheeran, P. (2016). The question-behaviour effect: A theoretical and

- methodological review and meta-analysis. *European Review of Social Psychology*, 27(1), 196–230. <https://doi.org/10.1080/10463283.2016.1245940>
- Wirtz, M. A. (Hrsg.). (2023). Kognitive Dissonanz. In *Dorsch – Lexikon der Psychologie*. Hogrefe. <https://dorsch.hogrefe.com/stichwort/kognitive-dissonanz>
- Wood, C. (2016). *The impact of asking intention or self-prediction questions on subsequent behavior*. <https://doi.org/10.1177/1088868315592334>
- Yang, C., Linas, B., Kirk, G., Bollinger, R., Chang, L., Chander, G., Siconolfi, D., Braxton, S., Rudolph, A., & Latkin, C. (2015). Feasibility and acceptability of smartphone-based ecological momentary assessment of alcohol use among African American men who have sex with men in Baltimore. *JMIR mHealth and uHealth*, 3(2), e67. <https://doi.org/10.2196/mhealth.4344>

Erklärung zur Verwendung von Künstlicher Intelligenz (KI)

Hiermit erkläre ich, dass ich die folgenden KI-Hilfsmittel zur Erstellung meiner Masterarbeit mit dem Titel *Messreaktivität im digitalen Ernährungstracking: Die Rolle von Ernährungszielen und wahrgenommener Forscherbeobachtung* verwendet habe. Ich versichere, dass die Kennzeichnung der verwendeten KI-Hilfsmittel vollständig ist, inkl. Verwendungszweck, Produktname, als auch ggfs. Prompts und Outputs des KI-Hilfsmittels. Ich verantworte die Auswahl und die Übernahme sämtlicher Ergebnisse der von mir verwendeten KI-Hilfsmittel vollumfänglich selbst.

Verwendungszweck	Produktname des KI-Hilfsmittels
Hilfe bei Erstellung der R-Syntax (wie Behebung von Fehlermeldungen etc.)	ChatGPT
Unterstützung bei sprachlicher Überarbeitung (wie kleine Umformulierungen, Rechtschreib- und Grammatikfehler etc.)	ChatGPT

Alternativ:

☐ Ich bestätige, dass ich zur Erstellung meiner Masterarbeit keine KI-Hilfsmittel verwendet habe.