



MASTERARBEIT | MASTER'S THESIS

Titel | Title

Messreaktivität bei der digitalen Ernährungserfassung:
Verändert Forscherbeobachtung den alltäglichen
Proteinkonsum?

verfasst von | submitted by
Stefan Lackner BSc

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

Wien | Vienna, 2025

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.

Zusammenfassung

Messreaktivität beschreibt die Veränderung von Kognition, Emotion und Verhalten einer Person allein durch die Tatsache, dass diese gemessen wird, wodurch die Interpretation beobachteter Effekte erschwert sein kann. Das Auftreten potenzieller reaktiver Effekte in der digitalen Ernährungserfassung wurde bislang allerdings noch nicht untersucht, trotz der steigenden Beliebtheit von Ernährungstracking-Apps. Ziel der vorliegenden Arbeit ist es daher, anhand des täglichen Proteinkonsums zu prüfen, ob sich das Ernährungsverhalten unter Forscherbeobachtung verändert und ob das Geschlecht beziehungsweise die bisherige Trackingdauer diese Effekte moderieren. Grundlage bildeten bereits verfügbare, über zwei Online-Fragebögen erhobene Daten der Ernährungstracking-Apps der Teilnehmenden ($n = 75$), anhand derer die Proteinwerte in der Woche vor (Woche 1) und nach Induktion der Forscherbeobachtung (Woche 2) miteinander verglichen wurden. Zur Hypothesenprüfung wurden lineare Multilevel-Modelle mit einem Random Intercept auf Personenebene berechnet. Insgesamt konnten keine Hinweise auf relevante Reaktivitätseffekte gefunden werden, es zeigte sich zwar ein leichter, jedoch nicht signifikanter Anstieg des Proteinkonsums von Woche 1 zu Woche 2 ($p = .399$; $\text{Total-}R^2 = .825$; $\text{Between-}R^2 = .000$; $\text{Within-}R^2 = .005$). Auch die getesteten Interaktionseffekte des Geschlechts ($p = .419$) und der Trackingdauer ($p = .395$) ergaben keine signifikanten Effekte. Die Ergebnisse deuten darauf hin, dass bei Weiterverwendung der gewohnten Ernährungstracking-App im Rahmen einer Studie kaum reaktive Effekte auftreten, zumindest im Hinblick auf den Proteinkonsum. Zukünftige Forschung sollte in diverseren Stichproben auch andere Ernährungsparameter wie den Kalorienkonsum untersuchen sowie Teilnehmende ohne Trackingerfahrung berücksichtigen.

Abstract

Measurement reactivity refers to changes in a person's cognition, emotion, and behavior that occur merely as a consequence of being measured, which can complicate the interpretation of observed effects. However, potential reactivity effects in digital dietary assessment tools have not yet been systematically investigated, despite the growing popularity of nutrition apps. The present study therefore examined whether dietary behavior, operationalized via daily protein intake, changes under researcher observation and whether gender or prior tracking duration moderate such effects. The study was based on already existing data from the app-based dietary records of the participants ($n = 75$) collected through two online questionnaires, which were used to compare protein values in the week before (Week 1) and after the induction of researcher observation (Week 2). Linear multilevel models with a random intercept at the person level were estimated. No evidence of relevant reactivity effects was found, although protein intake slightly increased from Week 1 to Week 2, this change was not significant ($p = .399$; Total- $R^2 = .825$; Between- $R^2 = .000$; Within- $R^2 = .005$). Likewise, the tested interaction effects of gender ($p = .419$) and tracking duration ($p = .395$) were non-significant. The findings suggest that continued use of one's habitual nutrition-tracking app while taking part in a research project produces minimal reactivity effects, at least with respect to protein intake. Future research should examine additional dietary parameters such as caloric intake and include participants without prior tracking experience.

INHALTSVERZEICHNIS

EINLEITUNG	1
MESSREAKTIVITÄT	1
Konzeptuelle Grundlagen der Messreaktivität	1
Erklärungsmodelle von Messreaktivität	2
Methodische Implikationen von Messreaktivität	4
DIGITALE ERNÄHRUNGSERFASSUNG	6
Gesundheitliche Relevanz von proteinreicher Ernährung	6
Der Stellenwert von Protein in der Ernährung	7
Konzept der digitalen Ernährungserfassung	7
Potenzielle Reaktivitätseffekte bei digitaler Ernährungserfassung	9
Potenzielle Moderatoren von Messreaktivität im Kontext digitaler Ernährungserfassung	12
ZIELE DER VORLIEGENDEN ARBEIT	13
FORSCHUNGSFRAGEN UND HYPOTHESEN	14
METHODE	16
STICHPROBE	16
MAßE UND MATERIALIEN	19
DESIGN UND ABLAUF	20
DATENAUSWERTUNG	20
Datenbereinigung	20
Deskriptivstatistische Auswertung	21
Voraussetzungstests	21
Modellvergleiche	22
Inferenzstatistische Auswertung	23
ERGEBNISSE	25
DESKRIPTIVE STATISTIKEN	25

ERGEBNISSE HYPOTHESE 1: EFFEKT DER FORSCHERBEOBACHTUNG AUF DEN TÄGLICHEN PROTEINKONSUM	25
ERGEBNISSE HYPOTHESE 2: INTERAKTION VON FORSCHERBEOBACHTUNG UND GESCHLECHT	26
ERGEBNISSE HYPOTHESE 3: INTERAKTION VON FORSCHERBEOBACHTUNG UND TRACKINGDAUER	26
DISKUSSION	28
EINORDNUNG IN DEN FORSCHUNGSSTAND UND INTERPRETATION	28
Generelle Erklärungsansätze	29
Untersuchte Moderatoren	32
STÄRKEN UND SCHWÄCHEN DER VORLIEGENDEN STUDIE	35
IMPLIKATIONEN UND ZUKÜNFTIGE FORSCHUNG	37
KONKLUSION	38
LITERATURVERZEICHNIS	40

Einleitung

Die vorliegende Arbeit befasst sich mit dem Phänomen der Messreaktivität im Zusammenhang mit digitaler Ernährungserfassung. Dabei wird konkret untersucht, ob und in welchem Ausmaß das Aufzeichnen der Ernährung im Alltag durch die Teilnahme an einer Studie verändert wird. Als Untersuchungsvariable dient dabei der tägliche Proteinkonsum, dazu werden das Geschlecht sowie die Vorerfahrung mit Ernährungstracking als potenzielle Moderatoren miteinbezogen.

Messreaktivität

Messreaktivität bezeichnet die Möglichkeit, dass allein die Teilnahme an einer wissenschaftlichen Studie oder das Wissen, gemessen zu werden, zu kognitiven, emotionalen oder verhaltensbezogenen Veränderungen führt. Messreaktivität stellt ein methodisches Problem für die empirische Forschung dar, da das Erheben von Daten das zu untersuchende Verhalten selbst beeinflussen kann. Da die Messreaktivität den grundlegenden Rahmen der vorliegenden Arbeit bildet, wird das Konzept im Folgenden theoretisch eingeführt, Erklärungsmodelle diskutiert sowie methodische Implikationen beschrieben, bevor mit der digitalen Ernährungserfassung der Anwendungsbereich dieser Untersuchung präzisiert wird.

Konzeptuelle Grundlagen der Messreaktivität

Das Phänomen der Messreaktivität (engl. *measurement reactivity*) lässt sich inhaltlich in die Nähe des weiter gefassten Konzepts des *Hawthorne-Effekts* rücken. Dieser beschreibt, dass sich Personen allein aufgrund des Wissens beobachtet zu werden, in ihrem Verhalten verändern (Roethlisberger & Dickson, 1939). Der Begriff wird seit den 1950er-Jahren in wissenschaftlichen Publikationen verwendet, wurde bislang jedoch nur begrenzt empirisch untersucht (McCambridge et al., 2014). In der Literatur fungiert der Hawthorne-Effekt zumeist als Oberbegriff für verschiedene Reaktivitätsphänomene, darunter eben auch Messreaktivität (French et al., 2021b).

Messreaktivität beschreibt die Veränderung von Kognition, Emotion und Verhalten einer Person allein durch die Tatsache, dass diese gemessen wird (French & Sutton, 2010). Bereits die bloße Teilnahme an einer wissenschaftlichen Studie kann somit reaktive Effekte auslösen, die ausschließlich im Zusammenhang mit der Erhebung entstehen (French et al., 2021b; McCambridge et al., 2014). Dieses Phänomen stellt eine Herausforderung für die verhaltenswissenschaftliche Forschung dar, da es sowohl die interne als auch die externe Validität empirischer Studien beeinflussen und somit die Interpretation von Effekten erschweren kann (z.B. French et al., 2021a; König et al., 2022a; Miles et al., 2018; Ram et al., 2017).

Grundsätzlich kann jede Messung, die an Menschen vorgenommen wird, reaktive Effekte auslösen. Die dadurch resultierenden Verhaltensänderungen können sowohl in die intendierte als auch in die entgegengesetzte Richtung wirken (Barta et al., 2012). Die meiste Evidenz für das Auftreten von Messreaktivität findet sich bislang in zwei Bereichen: Einerseits, wenn bereits das Stellen von Fragen zu einem bestimmten Verhalten Veränderungen eben dieses Verhaltens hervorruft, andererseits bei der Erfassung körperlicher Aktivität mittels Schrittzählern, insbesondere dann, wenn Studienteilnehmende ihre Schrittzahl selbst einsehen können, was zu einer Steigerung physischer Aktivität führen kann (French et al., 2021a). Die aktuelle Studienlage deutet darauf hin, dass Messreaktivität tendenziell ein allgemeines Phänomen darstellt, welches sich relativ konsistent über verschiedene Personen hinweg zeigt (Arigo & König, 2024). Das konkrete Design des Messinstruments kann allerdings entscheidend beeinflussen, wie ausgeprägt Reaktivität auftritt beziehungsweise in welchem Maße die Validität der Erhebung dadurch beeinträchtigt wird (Barta et al., 2012; Ram et al., 2017). Die Möglichkeiten, durch die Messreaktivität zu *Bias* (engl. für systematische Verzerrung von Schlussfolgerungen) führen kann, sind zudem äußerst vielfältig, genauso wie die verschiedenen Arten und Zeitpunkte während des Studienverlaufs, durch die sich Messreaktivität manifestieren kann (French & Sutton, 2010).

Erklärungsmodelle von Messreaktivität

In der wissenschaftlichen Literatur existieren mehrere theoretische Ansätze zur Erklärung der Entstehung von Messreaktivität. Ein häufig diskutierter Erklärungsansatz ist die erhöhte Aufmerksamkeit und Selbstwahrnehmung, die durch den Messprozess ausgelöst wird (French & Sutton, 2010). Diese erhöhte Aufmerksamkeit führt eventuell dazu, dass Personen ihr Verhalten bewusster wahrnehmen und reflektieren, was wiederum Anpassungen des Verhaltens begünstigen kann, unabhängig von etwaigen Interventionen (Barta et al., 2012). Durch das Erfassen eines spezifischen Verhaltens können Studienteilnehmende zum Beispiel angeregt werden, über Konsequenzen ihres Handelns nachzudenken oder Unterschiede zwischen ihrem aktuellen und einem erwünschten Verhalten festzustellen und entsprechend mit Selbstregulation zu reagieren (König et al., 2022a). Individuen setzen sich möglicherweise erstmals aktiv mit bestimmten Themen auseinander, was zu einer Verhaltensänderung führen kann (McCambridge et al., 2014).

Die erhöhte Selbstaufmerksamkeit und Bewertungsangst, die durch das Wissen entsteht, Teil einer Studie zu sein, kann ebenfalls zu reaktiven Verzerrungen führen. Dieser Effekt kann je nach Schwierigkeit der Aufgabe die Leistung hemmen oder fördern und ist besonders ausgeprägt, wenn die Beobachtungssituation salient ist. So gelten beispielsweise schriftliche Selbstberichte als weniger reaktiv als persönliche Interviews, und Befragungen in Gruppensituationen reaktiver als in privaten

Kontexten. Eine höhere wahrgenommene Anonymität wie etwa bei digitalen Tagebuchstudien könnte somit protektiv gegenüber diesen Prozessen wirken. Dennoch bleibt ein potenzieller Verzerrungsfaktor bestehen, wenn Teilnehmende befürchten, aufgrund ihrer Angaben später Rückmeldung zu erhalten oder negative Konsequenzen zu erleben (Barta et al., 2012).

Ein weiteres häufig genanntes Erklärungsmodell ist die sozialen Erwünschtheit. Teilnehmende verändern ihr Verhalten oder ihre Antworten, um den Erwartungen der Forschenden gerecht zu werden oder sich selbst in einem möglichst positiven Licht darzustellen. Auch wenn Teilnehmende nicht explizit über die Ziele beziehungsweise Hypothesen der Studie informiert sind, reichen oft schon Hinweise auf den Kontext oder die Art der Fragen aus, um konforme Reaktionen zu erzeugen. Bereits das Bewusstsein, beobachtet zu werden oder Gegenstand einer Verhaltensmessung zu sein, aktiviert unweigerlich Annahmen über die Erwartungen der Forschenden, wodurch es in der Folge zu einer Anpassung des Verhaltens an diese angenommenen Erwartungen kommen kann. Dieses Konformitätsverhalten gegenüber wahrgenommenen Normen gilt als ein grundlegender Wirkmechanismus von Reaktivitätseffekten (McCambridge et al., 2014; Thompson & Subar, 2017). Bei jüngeren Kindern konnte etwa eine geringere Ausprägung von Messreaktivität beobachtet werden, was vermutlich auf ihr noch nicht vollständig entwickeltes Bewusstsein für soziale Normen zurückzuführen ist (König et al., 2022a).

Barta et al. (2012) differenzieren zusätzlich zwischen zwei Formen der sozialen Erwünschtheit: *Impression management* als bewusster Versuch sich vorteilhaft darzustellen, und *deceptive self-enhancement* als unbewusster Prozess, ein positives Selbstbild aufrechtzuerhalten. Während impression management stärker durch unmittelbare soziale Präsenz beeinflusst wird, stellt deceptive self-enhancement eine bedeutendere Herausforderung dar: Insbesondere bei retrospektiven Erhebungen, bei denen länger zurückliegende Ereignisse erfragt werden, neigen Individuen in der Regel dazu, die Erinnerungslücken durch schemageleitete und selbstaufwertende Rekonstruktion aufzufüllen. Methoden mit kürzeren Erinnerungszeiträumen wie einem Tag oder weniger scheinen weniger anfällig für diese Form der Verzerrung zu sein.

Falls Teilnehmende wiederholt zu sensiblen Themen befragt werden, kann abgesehen davon ein Gewöhnungseffekt eintreten, welcher die Wahrscheinlichkeit erhöht, sozial unerwünschte Verhaltensweisen im weiteren Studienverlauf offenzulegen. Wenn etwa eine Person einmal angibt, ein bedenkliches Verhalten gezeigt zu haben und daraufhin keine negativen Konsequenzen erfährt, steigt die Wahrscheinlichkeit, dass dieses Verhalten bei späteren Befragungen weiterhin erfolgt (Catania et al., 1990).

Ein weiterer etwaiger Wirkmechanismus stellt die Tendenz dar, bei der Beantwortung von Fragen nur ein begrenztes Maß an Anstrengung zu investieren. Insbesondere Zeitdruck, Ermüdung,

oder eine hohe Belastung – etwa durch eine große Anzahl oder komplexe Formulierung von Fragen – können zu oberflächlichem Antwortverhalten führen. Auch die von Teilnehmenden verwendete Erinnerungsstrategie spielt eine mögliche Rolle bei Reaktivitätseffekten: Während „*bottom-up*“-Strategien auf konkreten Einzelerinnerungen basieren, greifen „*top-down*“-Prozesse auf allgemeine Eindrücke zurück und sind anfälliger für Verzerrungseffekte. Speziell Tagebuchstudien können solchen Effekten begegnen, indem sie kurze, klar formulierte Fragen verwenden, subjektive Einschätzungen unterlassen sowie den Schwerpunkt auf konkrete Verhaltensberichte legen (Barta et al., 2012).

Methodische Implikationen von Messreaktivität

In der Gesundheitspsychologie stellt Messreaktivität ein grundlegendes Problem dar, da die Teilnahme an Studien selbst Verhaltensänderungen hervorrufen kann, unabhängig von anderen Interventionen (French & Sutton, 2010). Wenn also durch Reaktivitätseffekte Veränderungen im Verhalten oder Erleben der Teilnehmenden induziert werden, erfolgt die Datenerhebung *de facto* in einer veränderten Realität und somit unter Bedingungen, die nicht mehr der natürlichen Lebenswelt der Personen entsprechen. Die ökologische Validität, also die Verallgemeinerbarkeit der Studienergebnisse auf das Alltagsgeschehen, ist somit beeinträchtigt (Ram et al., 2017).

French und Sutton (2010) identifizieren drei zentrale Probleme, die durch Messreaktivität in experimentellen Studien entstehen können: Erstens kann es zu einer Konfundierung von Mess- und Interventionseffekten kommen, wenn beide auf denselben psychologischen Mechanismus wirken, etwa bei der Verwendung von Schrittzählern in beiden Gruppen, was die Unterschiede zwischen Interventions- und Kontrollgruppe verringert. Zweitens können Messungen mit der Intervention interagieren, selbst ohne direkte Konfundierung, etwa wenn bereits durch die Erhebung gesundheitsförderliches Verhalten angestoßen wird und somit das Veränderungspotenzial der Intervention eingeschränkt ist. Letztlich kann Messreaktivität zu systematischen Veränderungen der erhobenen Variablen führen, was insbesondere bei hohen Ausgangswerten zu Deckeneffekten führen und dadurch die Nachweisbarkeit von Interventionswirkungen und die statistische Power einer Studie verringern kann.

Der Effekt von Messreaktivität wird in der bisherigen Literatur generell als klein bis mittelgroß eingeschätzt. Bezüglich der Frage, wie stark Messreaktivität die Ergebnisse von Studien verzerren kann, liegt bislang jedoch nur begrenzte empirische Evidenz vor. Bestehende Übersichtsarbeiten und Metaanalysen verdeutlichen zudem eine erhebliche methodische und inhaltliche Heterogenität der einbezogenen Studien, was die Generalisierbarkeit der Befunde zusätzlich limitiert (König et al., 2022a). In einer Übersichtsarbeit zu Messreaktivität fanden French et al. (2021b) konsistent kleine Effektstärken, welche im Bereich von $d = 0.09$ bis $d = 0.28$ lagen. Der reaktive Effekt scheint zudem

eher von kurzer Dauer zu sein (König et al., 2022a), in den vorhandenen Studien werden hauptsächlich in den ersten zwei Tagen Effekte gefunden (Arigo & König, 2024), mitunter bis zu etwa einer Woche (Clemens & Deans, 2012). Generell erscheint es wenig ratsam, den potenziellen Einfluss von Reaktivitätseffekten zu vernachlässigen, selbst wenn diese nur geringe Effektstärken aufweisen. Die zentrale Frage ist dabei weniger, ob diese Effekte existieren, sondern vielmehr, wer davon betroffen ist, wie stark diese Effekte ausfallen und inwiefern sie in verschiedenen Forschungskontexten variieren (Bendtsen & McCambridge, 2021).

Allgemein wird angenommen, dass Messreaktivität Verhalten in jene Richtung beeinflusst, die als erwünscht oder gesund angesehen wird, beispielsweise in Form von gesteigerter körperlicher Aktivität. Diese Reaktivitätseffekte treten in der Regel zu Beginn der Erhebungsphase auf und lassen nach ein bis zwei Tagen nach, woraufhin das Verhalten wieder zu seinem Ausgangsniveau zurückkehrt (Arigo & König, 2024). Barta et al. (2012) führen diesen Rückgang reaktiver Effekte primär auf eine nachlassende Aufmerksamkeit der Teilnehmenden gegenüber der Erhebungssituation sowie auf eine zunehmende Gewöhnung an das Messverfahren zurück. Die Teilnahme an Forschungen kann darüber hinaus Selbstpräsentations- und soziale Erwünschtheitseffekte aktivieren, wodurch Teilnehmende ihr Verhalten stärker an allgemein bekannten Gesundheitsrichtlinien ausrichten. Da in vielen Studien Verhaltensdaten jedoch über mehrere Tage gemittelt werden, um ein typisches Maß des Zielverhaltens zu bestimmen, können bereits geringe Veränderungen zu Beginn der Erhebung Mittelwerte verzerren und somit zu ungenauen oder sogar irreführenden Schlussfolgerungen führen (König et al., 2025). Dementsprechend wird Messreaktivität häufig über zeitliche Veränderungen untersucht, etwa durch den Vergleich des Zielverhaltens am ersten Erhebungstag mit späteren Tagen (Arigo & König, 2024).

Außerdem wird argumentiert, dass die Leichtigkeit der Verhaltensänderung selbst ebenfalls einen Einfluss auf das Ausmaß von Messreaktivität haben kann. Verhalten, welches kurzfristig und ohne großen Aufwand modifizierbar ist, ist wahrscheinlich anfälliger für Reaktivitätseffekte als anstrengende oder zeitraubende Veränderungen (König et al., 2022a).

Alles in allem wird davon ausgegangen, dass das Reaktivitätspotenzial stark variiert und in erheblichem Maße von der thematischen Ausrichtung der Studie, der Gestaltung des Erhebungsprotokolls, individuellen Merkmalen der Teilnehmenden sowie der zeitlichen Struktur der Datenerhebung abhängt. Während etwa eine stündliche Erhebung sozialer Interaktionen wenig Einfluss auf das Verhalten von Interaktionspartnern nehmen dürfte, würde tägliches Erfassen von bestimmten Nahrungsmitteln deutlich wahrscheinlicher Veränderungen im Ernährungsverhalten auslösen (Ram et al., 2017).

Exakt diese Prämisse bildet den Ausgangspunkt der vorliegenden Arbeit. Da Messreaktivität grundsätzlich in einer Vielzahl von Forschungsfeldern auftreten kann, erscheint es plausibel, dass auch die Ernährungserfassung im Alltag von entsprechenden Effekten betroffen ist. Im Folgenden wird daher die digitale Ernährungserfassung als Untersuchungsbereich dieser Arbeit eingeführt und theoretisch eingeordnet.

Digitale Ernährungserfassung

Die Ernährungserhebung mithilfe digitaler Mittel stellt den untersuchten Anwendungsbereich der vorliegenden Arbeit dar, spezifischer der potenzielle Zusammenhang zwischen Messreaktivität und dem erhobenen Proteinkonsum im Alltag. Vor dem Hintergrund der gesundheitlichen und gesellschaftlichen Bedeutsamkeit einer ausgewogenen und proteinreichen Ernährung werden daher im Folgenden die theoretischen und methodischen Grundlagen digitaler Ernährungserfassung dargestellt. Dabei wird zunächst auf die gesundheitliche und gesellschaftliche Relevanz gesunder Ernährung eingegangen, bevor das Konzept digitaler Erhebungsverfahren selbst, methodische Aspekte sowie zuletzt mögliche reaktive Effekte dabei näher behandelt werden.

Gesundheitliche Relevanz von gesunder Ernährung

Nicht übertragbare Krankheiten (*Non-communicable diseases* [NCDs]), zu denen insbesondere Herz-Kreislauf-Erkrankungen, Krebs, chronische Atemwegserkrankungen, Diabetes und Adipositas gehören, zählen zu den weltweit häufigsten Todesursachen und betreffen sowohl die Bevölkerung in Industrie- als auch in Entwicklungsländern (Cena & Calder, 2020; World Health Organization, 2024). Obwohl genetische und umweltbedingte Faktoren ebenfalls zu einem erhöhten NCD-Risiko beitragen, gelten vor allem veränderbare gesundheitsbezogene Verhaltensweisen wie die Ernährung, physische Aktivität oder Alkohol- und Tabakkonsum als zentrale Determinanten auf individueller Ebene (z.B. Koene et al., 2016; O'Neil et al., 2015).

Insbesondere ungünstige Ernährungsentscheidungen tragen zum Risiko der Entwicklung von Bluthochdruck, Übergewicht, Hypercholesterinämie und Entzündungen bei, welche wiederum mit der Ausbildung von NCDs einhergehen (Sun et al., 2014), weltweit können mehr als 18% aller durch NCDs verursachte Tode ernährungsbedingten Risikofaktoren zugeschrieben werden (Melaku et al., 2019). Der deutliche Anstieg chronischer NCDs steht dabei in engem Zusammenhang mit der Ausbreitung globaler, westlich geprägter Ernährungsmuster. Diese zeichnen sich durch einen hohen Anteil an fettem und verarbeitetem Fleisch, gesättigten Fetten, raffiniertem Getreide, Salz sowie Zucker und einem geringem Konsum von frischem Obst und Gemüse aus (Cena & Calder, 2020; Lane et al., 2021).

Eine gesunde, ausgewogene Ernährung ist somit essenziell für die Erhaltung der allgemeinen Gesundheit, die Förderung des körperlichen und geistigen Wohlbefindens sowie die Vorbeugung

zahlreicher chronischer Krankheiten (World Health Organization, 2020). Das regelmäßige Verzehren von nährstoffreichen Lebensmitteln unterstützt den menschlichen Körper bei Energiehaushalt, Wachstum, Immunfunktion sowie der Aufrechterhaltung physiologischer Prozesse (Arias et al., 2025).

Der Stellenwert von Protein in der Ernährung

Protein ist mit Kohlenhydraten und Fetten einer der drei essenziellen Makronährstoffe und damit ein grundlegender Bestandteil einer gesunden Ernährung. Die in Proteinen enthaltenen Aminosäuren übernehmen im menschlichen Organismus eine Vielzahl grundlegender Funktionen und sind unentbehrlich für grundlegende physiologische Prozesse wie Wachstum, Entwicklung, Reproduktion sowie Laktation, zudem spielen sie speziell bei Aufbau sowie Erhaltung von Muskelmasse eine maßgebliche Rolle (Arentson-Lantz et al., 2015; Wu, 2016). Die täglich empfohlene Menge für Erwachsene liegt dabei zwischen 0.8–1.0 g pro Kilogramm Körpergewicht, obwohl diese für sportlich aktive Menschen deutlich höher (1.2–2.0 g/kg) liegen kann (Deutsche Gesellschaft für Ernährung, 2017).

Protein spielt in vielen Ernährungsstrategien eine prominente Rolle, insbesondere bei der Gewichtsabnahme (Moon & Koh, 2020). Diese Beliebtheit lässt sich auf mehrere Faktoren zurückführen: Einerseits fördert Protein den Aufbau von Muskelmasse durch die Stimulierung der Muskelproteinsynthese und die Unterdrückung des Muskelproteinabbaus (Stokes et al., 2018), andererseits geht der Konsum proteinreicher Lebensmittel mit einer hohen Sättigungswirkung einher (Soenen & Westerterp-Plantenga, 2008). Eine proteinhaltige Ernährung wird darüber hinaus mit einem erhöhtem physischen Leistungs- und Ausdauerniveau in Verbindung gebracht (Kimura et al., 2022; Zhao et al., 2024) und gilt allgemein als gesund, weshalb proteinreiche Lebensmittel einen entsprechend positiven Ruf besitzen (Drewnowski et al., 2010). Eine Kombination aus regelmäßiger körperlicher Aktivität und einer erhöhten Zufuhr hochwertiger Proteinquellen stellt zudem eine wirksame Strategie zur Reduktion von Körperfett und zur Verbesserung des Stoffwechsels dar (Moon & Koh, 2020), unterstützt die Gewichtsstabilisierung (Bardus et al., 2016) und führt bei längerfristigem Krafttraining zu deutlich größeren Zuwächsen an Muskelmasse und -kraft als bei unzureichenden Mengen (Morton et al., 2017).

Konzept der digitalen Ernährungserfassung

Vor dem Hintergrund der zuvor beschriebenen Bedeutung ausgewogener und proteinreicher Ernährungsweisen stellt sich die Frage, wie sich Ernährungsverhalten möglichst valide und alltagsnah erfassen lässt. Zur Erfassung des Ernährungsverhaltens werden in der Forschung verschiedene Methoden eingesetzt, die sich hinsichtlich Genauigkeit, Erhebungsaufwand und Anfälligkeit für

Verzerrungen unterscheiden. Zu den geläufigsten Verfahren zählen die 24-Stunden-Ernährungserinnerung (engl. *24-hour dietary recall*), bei der Studienteilnehmende gebeten werden, rückblickend alle konsumierten Lebensmittel eines Tages zu berichten, die Lebensmittel-Frequenz-Fragebögen (engl. *food frequency questionnaires*), die typischerweise über einen längeren Zeitraum hinweg erfassen, wie häufig bestimmte Lebensmittel konsumiert werden, sowie Selbstaufzeichnungen in Form von Tagebüchern (engl. *self-reported dietary records*), bei denen Teilnehmende für eine bestimmte Dauer alle konsumierten Nahrungsmittel und Getränke aufzeichnen (Thompson & Subar, 2017).

Self-reported dietary records gelten dabei als Goldstandard, da sie es ermöglichen, sämtliche konsumierte Lebensmittel und Getränke inklusive Portionsgrößen direkt zum Zeitpunkt des Konsums zu erfassen. Im Vergleich zu retrospektiven Verfahren sind Selbstaufzeichnungen somit weniger anfällig für Erinnerungsfehler und liefern folglich detailliertere sowie validere Daten (Shiffman et al., 2008; Thompson & Subar, 2017). Gleichzeitig setzt diese Methode allerdings auch ein hohes Maß an Adhärenz voraus und kann für Studienteilnehmende mit erheblichem Aufwand verbunden sein, weswegen empfohlen wird, sie nur für kürzere Zeiträume von wenigen Tagen anzuwenden. Für die Wirksamkeit ist entscheidend, dass die Erfassung konsumierter Lebensmittel sowohl regelmäßig als auch möglichst häufig erfolgt (Raber et al., 2021; Turner-McGrievy et al., 2019). Auf diese Weise kann nicht nur das Bewusstsein für das eigene Ernährungsverhalten gesteigert, sondern durch zunehmende Routine im Umgang mit der Methode auch die Genauigkeit der Angaben verbessert werden (Thompson & Subar, 2017).

Self-reported dietary records weisen allerdings auch eine Reihe von Limitationen auf, darunter mögliche Schwierigkeiten beim späteren Erinnern von konsumierten Lebensmitteln, begrenzte Nährwertdatenbanken sowie teilweise Probleme bei der Schätzung von Portionsgrößen (Eldridge et al., 2019). Diese Einschränkungen wirken sich wiederum negativ auf die Reliabilität sowie die ökologische Validität der erhobenen Daten aus (Maugeri & Barchitta, 2019). Darüber hinaus erfordert sie nicht nur für Nutzende, sondern auch für Forschende erheblichen Aufwand beziehungsweise Zeitbedarf, da die erfassten Lebensmittel nachträglich kodiert und in Nährwertangaben umgerechnet werden müssen (Eldridge et al., 2019).

Der technologische Fortschritt hat zur Entwicklung zahlreicher digitalen Ernährungserfassungsmethoden geführt, welche einige der vorhandenen Einschränkungen der traditionellen Erhebungsarten adressieren und zusätzliche technische Vorzüge integrieren (Naska et al., 2017, Patel et al., 2021). Besonders Smartphones haben sich als geeignetes Instrument für solche Verfahren erwiesen, da sie weit verbreitet sind, von Nutzer*innen nahezu ständig mitgeführt werden und damit die Möglichkeit bieten, Mahlzeiten unmittelbar vor dem Verzehr zu dokumentieren. Zudem

erlauben sie die simple Installation von zahlreichen Applikationen und gewährleisten durch ihre Internetverbindung sowie Tools wie die integrierte Kamerafunktion weitere Funktionalitäten (König et al., 2022b, Wahl et al., 2020). Da technologische Geräte mittlerweile fest in den Alltag vieler Menschen integriert sind, bevorzugen viele Nutzer*innen eine digitale Erfassung ihrer Nahrungsaufnahme gegenüber traditionellen Methoden, besonders jüngere Menschen (Eldridge et al., 2019).

Smartphonebasierte Verfahren zur Ernährungserfassung weisen eine vergleichbare Validität wie traditionelle Erhebungsmethoden auf und reduzieren zeitgleich den Aufwand für Teilnehmende (Eldridge et al., 2019). Digitale Ernährungsapps können für zahlreiche Zwecke genutzt werden, beispielsweise zur Selbstbeobachtung, zur Verbesserung der Ernährungsqualität oder sogar zur gezielten Gewichtszunahme (König et al., 2021). Ein Großteil der verwendeten Ernährungstagebuch-Apps zielt auf eine vollständige Erfassung der täglichen Nahrungsaufnahme ab und erfordert von den Nutzenden die Angabe des Kalorien- beziehungsweise Nährstoffgehalts oder exakter Grammangaben (Pala et al., 2024). Gleichzeitig unterscheiden sich die Funktionalitäten entsprechender Programme teils erheblich, wobei angesichts der zunehmenden Verbreitung smartphonebasierter Erhebungsverfahren davon auszugehen ist, dass sowohl deren Anzahl als auch deren Vielfalt künftig weiter zunehmen wird (z.B. Bardus et al., 2016; König et al., 2022b; Pala et al., 2024).

Digitale Erhebungsverfahren, insbesondere jene auf Smartphone-Basis, bieten den zusätzlichen Vorteil, nicht nur Unterschiede zwischen Personen, sondern auch intraindividuelle Schwankungen sowie individuelle Unterschiede in den beobachteten Effekten analysieren zu können. Da die Erhebung im Alltag der Teilnehmenden stattfindet, eröffnen sich darüber hinaus Möglichkeiten, psychologische Prozesse des Essverhaltens wie etwa situative Einflüsse oder zeitliche Dynamiken in Echtzeit zu untersuchen. Dies ermöglicht ein tiefergehendes Verständnis des Verhaltens selbst sowie seiner Veränderung über die Zeit hinweg und kann somit langfristig zur Weiterentwicklung psychologischer Theorien und Interventionen beitragen (König et al., 2022b, Perski et al., 2022). Zudem zeichnen sich digitale Anwendungen durch ihre Interaktivität, die Möglichkeit zur individuellen Anpassung sowie ihr Potenzial aus, wirksame und alltagstaugliche Interventionen für unterschiedliche Zielgruppen bereitzustellen. Entsprechend hat sich der Einsatz von Smartphones zur Erhebung alltäglichen Ernährungsverhaltens in der verhaltenswissenschaftlichen Forschung zunehmend etabliert (Villinger et al., 2019).

Potenzielle Reaktivitätseffekte bei digitaler Ernährungserfassung

Empirische Befunde weisen darauf hin, dass Messreaktivität sowohl im Rahmen einmaliger gesundheitsbezogener Erhebungen per Fragebogen (z.B. McCambridge et al., 2014; Miles et al., 2020; Rodrigues et al., 2015; Wilding et al., 2016), als auch bei regelmäßiger Selbstbeobachtung

gesundheitsbezogener Verhaltensweisen auftreten kann (z.B. Barta et al., 2012; Eisele et al., 2023; Ram et al., 2017). In einer Metaanalyse von Miles et al. (2020) zeigte sich beim reinen Beantworten von gesundheitsrelevanten Fragen ein konsistenter, kleiner Effekt ($SMD = 0.06$, 95% CI [0.02, 0.09]) von messbaren Verhaltensänderungen. Als zugrunde liegende Mechanismen werden ein gesteigertes Bewusstsein für das eigene Verhalten, die kognitive Aktivierung gesundheitsbezogener Aspekte sowie die Reflexion möglicher Konsequenzen diskutiert.

Bislang ist allerdings noch relativ unklar, inwieweit auch digital erfasste Verhaltensdaten durch Messreaktivität in ihrer Validität beeinträchtigt werden (French & Sutton, 2010). Da Personen bei der Nutzung digitaler Messinstrumente in der Regel über den Studienkontext informiert sind, erscheint es plausibel, dass auch die digitale Erhebung von Verhalten anfällig für reaktive Effekte ist. In bisherigen Studien zur Messreaktivität in der digitalen Erfassung von Gesundheitsverhalten wurden in einer Metaanalyse kleine bis mittelgroße Effektstärken ($d = 0.27$, 95% CI [0.16, 0.39]) gefunden, allerdings weisen die Ergebnisse eine hohe Heterogenität auf (König et al., 2022a).

Abgesehen von einer einzelnen Studie, die den Konsum alkoholischer und nicht-alkoholischer Getränke in Ausgehsituationen untersuchte, existieren allerdings keine Untersuchungen zur Reaktivität bei der digitalen Ernährungserfassung, nahezu sämtliche Forschung bezüglich potenziell auftretender reaktiver Effekte im Rahmen digital erfasster Verhaltensdaten konzentriert sich bislang auf den Zusammenhang mit körperlicher Aktivität (König et al., 2022a). Trotz der Beliebtheit von digitalen Tracking-Apps (Maugeri & Barchitta, 2019) fehlen daher Studien über potenzielle Reaktivitätseffekte bei der täglichen Nahrungsaufnahme.

Messreaktivität sollte prinzipiell bei jeder Studie berücksichtigt werden, die auf regelmäßig wiederholte Dokumentationen angewiesen ist (Strath et al., 2013), wie es auch bei der digitalen Ernährungserfassung der Fall ist. Grundsätzlich wird davon ausgegangen, dass self-reported dietary records anfällig für Reaktivitätseffekte sind (Thompson & Subar, 2017), speziell aufgrund des vorausschauenden Charakters von Ernährungsdokumentationen (Lucassen et al., 2022). So könnten Teilnehmende ihr Verhalten bereits im Vorfeld der Aufzeichnung anpassen, indem sie beispielsweise den Entschluss fassen, an diesem Tag besonders gesund zu essen, da jede Mahlzeit dokumentiert werden muss. Ebenso kann der Akt des Aufzeichnens selbst zu Veränderungen führen, wenn Teilnehmende etwa im Verlauf der Dokumentation bemerken, dass sie einen hohen Anteil ungesunder Lebensmittel konsumieren, und daraufhin beginnen, ihren Junkfood-Konsum zu reduzieren. (Harnack & Pereira, 2017). Self-reported dietary records bringen somit den Nachteil mit sich, dass die gleichzeitige Dokumentation vor beziehungsweise während des Essens sowohl die Auswahl der Lebensmittel als auch deren Menge beeinflussen kann. Allein das Wissen, dass sämtliche Speisen und Portionen erfasst werden müssen, sowie der damit verbundene Aufwand können das

Ernährungsverhalten in eine gesundheitskonformere Richtung verschieben, ein allgemein positiver Effekt, jedoch problematisch für die Forschung (Thompson & Subar, 2017).

Auch Robinson et al. (2015) fanden in einer Metaanalyse Evidenz für potenzielle Reaktivitätseffekte bei Ernährungserhebungen. In Laborstudien konnte gezeigt werden, dass allein das Bewusstsein, beim Essen beobachtet oder gemessen zu werden, die Nahrungsaufnahme signifikant reduziert. Dieser Effekt wird auch hier vor allem durch Selbstpräsentationsprozesse erklärt: Teilnehmende passen ihr Verhalten an, um den sozialen Erwartungen zu entsprechen und ein positives Bild von sich zu vermitteln, da übermäßiges Essen häufig negativ bewertet wird. Auch wenn es sich hierbei um Laborstudien handelt, verdeutlichen die Ergebnisse, dass die bloße Wahrnehmung von Beobachtung in der Lage ist, das Essverhalten zu verändern. Auch digitale Verfahren zur Ernährungserhebung, bei denen den Studienteilnehmenden bewusst ist, dass ihr Verhalten dokumentiert wird, könnten somit anfällig für vergleichbare Reaktivitätseffekte sein.

Da leichter veränderbares Verhalten vermutlich stärker zu reaktiven Effekten führt als schwieriger modifizierbares Verhalten (König et al., 2022a), erscheint dies besonders relevant für das Ernährungsverhalten, was sich vergleichsweise flexibel in den Tagesablauf integrieren lässt. Damit zusammenhängend ergab eine Studie von Allmeta et al. (2022) zudem, dass Ernährungsverhalten im Vergleich zu körperlicher Aktivität als leichter veränderbar wahrgenommen wird, was zusätzlich für potenzielle Reaktivitätseffekte im Kontext digitaler Ernährungserfassung spricht.

Trotz der technischen Vorzüge wird auch davon ausgegangen, dass digitale Ernährungserhebungen anfällig für reaktive Effekte sind, vermutlich da sie vergleichbare Mechanismen der Selbstaufmerksamkeit, sozialen Erwünschtheit und Auseinandersetzung mit den dokumentierten beziehungsweise konsumierten Lebensmitteln aktivieren (Lucassen et al., 2022). Da Messreaktivität Verhalten vermutlich auf eine erwünschte beziehungsweise gesundheitliche Art beeinflusst, ist im Kontext digitaler Ernährungserfassung ebenfalls davon auszugehen, dass Reaktivitätseffekte in eine gesundheitsförderliche Richtung wirken. Teilnehmende könnten in diesem Sinne ihre Ernährungsgewohnheiten kurzfristig an gängige Gesundheitsnormen anpassen und proteinreichere Lebensmittel wählen beziehungsweise mehr Protein pro Tag konsumieren. Zusätzlich könnten die durch Forschungsteilnahme aktivierten Prozesse der Selbstpräsentation und sozialen Erwünschtheit dazu führen, dass Teilnehmende ihre Angaben stärker an gesellschaftlich akzeptierte Ernährungsempfehlungen ausrichten. Die aktuelle Studienlage legt nahe, dass solche Effekte besonders zu Beginn der Erhebungsphase auftreten und nach ein bis zwei Tagen wieder abflachen, sobald eine Gewöhnung an das Messverfahren eintritt. Auch Reaktionen auf wiederholte Messungen in sehr kurzen, etwa stündlichen Intervallen, klingen typischerweise rasch ab (König et al., 2025). Vor diesem Hintergrund erscheint es plausibel, dass selbst im Falle von reaktiven Effekten bei der digitalen

Erfassung des Ernährungsverhaltens höchstens anfänglich messbare Verhaltensänderungen bestehen, welche aufgrund der raschen Gewöhnung an die Messsituation vermutlich kaum über die ersten Tage hinausreichen. Diese anfänglichen Veränderungen können dennoch das alltägliche Ernährungsverhalten verzerrt darstellen, wenn die Daten mehrerer Tage gemittelt werden. In dieser Untersuchung werden aus diesem Grund nur die ersten beiden Tage unter Forscherbeobachtung berücksichtigt.

Potenzielle Moderatoren von Messreaktivität im Kontext digitaler Ernährungserfassung

Verschiedene individuelle Merkmale werden als mögliche Moderatorvariablen von Messreaktivität diskutiert. Abgesehen von den untersuchten Haupteffekten stellt das Geschlecht einen potenziellen Moderator in Bezug auf Messreaktivität dar, die Studienlage dazu ist allerdings momentan eher heterogen (König et al., 2025). König et al. (2022a) identifizierten in einer Übersichtsarbeit fünf Studien, die den Einfluss des Geschlechts auf Reaktivitätseffekte beim Messen von täglichen Schritten untersuchten, wo überall keine Unterschiede festgestellt wurden. Eine weitere Untersuchung zu Aktivitätsdaten, welche eng mit Schrittzahlen korrelieren, zeigte jedoch, dass Reaktivität bei Mädchen, nicht aber bei Jungen auftrat (Ho et al., 2013).

Die Befunde einer Laborstudie von Stubbs et al. (2014), bei denen Frauen ihre Nahrungsaufnahme unter Forscherbeobachtung stärker reduzierten als Männer, deuten ebenfalls darauf hin, dass Frauen von diesen Effekten in stärkerem Maße betroffen sein könnten als Männer, was auf eine höhere Sensibilität gegenüber sozialen Normen im Ernährungsverhalten hindeutet. Generell scheinen Frauen anfälliger für soziale Erwünschtheit zu sein, wenn sie über ihr Gesundheitsverhalten berichten und somit wahrscheinlicher als Männer ihr Verhalten im Rahmen von Studienteilnahmen anpassen (Bernardi & Guptill, 2008; Hebert et al., 1997).

Frauen zeigen im Durchschnitt ein gesünderes Ernährungsverhalten als Männer (Leblanc et al., 2015) und berichten häufiger, ihre Ernährung an gängigen Ernährungsempfehlungen auszurichten (Fagerli & Wandel, 1999). Weiters legen Studien nahe, dass Frauen ein stärkeres Interesse an Ernährung und Nährwertinformationen haben und sensibler auf gesundheitsbezogene Informationen reagieren als Männer (Kiefer et al., 2005). Für die vorliegende Arbeit könnte man daher davon ausgehen, dass geschlechtsspezifische Unterschiede auftreten werden, da die sozialen Erwartungen sowie die individuellen Ziele von Männern und Frauen variieren, weswegen Frauen eventuell stärker auf die durch Messreaktivität induzierten Effekte reagieren.

Ein weiterer denkbarer Moderator von Reaktivitätseffekten bei der digitalen Ernährungserfassung ist die Trackingdauer, also der Zeitraum, seit der die aktuelle Tracking-App verwendet wird. Studien zeigen, dass selbstbeobachtungsbezogene Aufmerksamkeit mit

zunehmender Routine reduziert wird und automatisierte Verhaltensmuster resistenter gegenüber externen Reizen werden (Wood & Neal, 2007). Darüber hinaus weisen Befunde aus der Forschung zu Routinen und Selbstmonitoring darauf hin, dass neue Verhaltensmuster in der Anfangsphase besonders anfällig für Feedback, Kontrolle und soziale Bewertung sind (Fritz et al., 2014; Verplanken & Orbell, 2022). Aus theoretischer Perspektive könnte man also vermuten, dass unerfahrene Tracker*innen stärker auf die Aufnahme einer Studie reagieren, weil sie weniger Routine im Umgang mit Ernährungstracking-Apps haben und daher sensibler reagieren, wenn sie sich beobachtet fühlen oder sich der Studienteilnahme bewusster werden. Im Gegensatz dazu könnten erfahrene Nutzer*innen bereits gewohnt sein, ihr Verhalten zu protokollieren, sodass zusätzliche Beobachtung kaum mehr Aufmerksamkeit erzeugt beziehungsweise zu Verhaltensanpassungen führt.

Auch bei der digitalen Erfassung von Gesundheitsverhalten wird angenommen, dass erfahrenere Tracker*innen eventuell weniger anfällig für Verhaltensänderungen durch Forscherbeobachtung sind, da sie bereits Routine im Umgang mit Tracking-Geräten entwickelt haben und dadurch weniger anfällig für Reaktivitätseffekte sind (König et al., 2022a). Die Dauer, seit der Personen ihre Ernährung mit einer App dokumentieren, beeinflusst also möglicherweise ebenfalls die etwaige auftretende Messreaktivität. Erfahrenere Nutzer*innen einer Ernährungstracking-App könnten aufgrund routinierterer Handhabung und geringerer Neuheitseffekte mit ihrer App weniger stark reagieren, während unerfahrenere Tracker*innen im Umkehrschluss anfälliger für Reaktivität sein könnten. Diese Annahme wurde in der Form allerdings noch nicht empirisch untersucht, zudem wird die bisherige Trackingerfahrung der Studienteilnehmenden meist gar nicht erst berichtet (König et al., 2025). Um diese Lücke anzuschneiden, wird die Trackingdauer in der vorliegenden Arbeit als Moderator für etwaige auftretende reaktive Effekte bei der digitalen Ernährungserfassung miteinbezogen.

Ziele der vorliegenden Arbeit

In Summe verdeutlichen diese Befunde die Bedeutsamkeit, potenzielle Reaktivitätseffekte bei digitalen Ernährungserhebungen systematisch zu berücksichtigen, da sie die ökologische Validität der erhobenen Daten beeinträchtigen können. Trotz der Popularität von digitalen Tracking-Apps existieren bislang allerdings kaum Studien über womöglich auftretende Messreaktivität beim alltäglichen Ernährungsverhalten.

Die vorliegende Arbeit soll diese Forschungslücke adressieren, um die ökologische Validität digitaler Selbsterhebungen von alltäglichem Ernährungsverhalten zukünftig besser einschätzen zu können. Da leichter veränderbares Verhalten tendenziell anfälliger für Messreaktivität ist, dürfte dies auch für das Ernährungsverhalten gelten, das flexibel in den Alltag integriert werden kann und als

verhältnismäßig leichter veränderbar angesehen wird. Dazu zählt auch der tägliche Proteinkonsum, weshalb dieser, zusammen mit den angeführten gesundheitlichen Hintergründen, die abhängige Variable der vorliegenden Arbeit darstellt. Aufgrund der Forscherbeobachtung und der damit vermutlich einhergehenden bewussteren Auseinandersetzung mit der eigenen Ernährung kann davon ausgegangen werden, dass Studienteilnehmende ihre Proteinaufnahme tendenziell eher steigern als senken werden.

Darüber hinaus wird geprüft, ob bestimmte personenspezifische Merkmale die Stärke potenzieller Reaktivitätseffekte moderieren. So stellt das Geschlecht eine mögliche Moderatorvariable dar: Aufgrund unterschiedlicher sozialer Erwartungen und individueller Zielsetzungen in Bezug auf Ernährung könnten sich geschlechtsspezifische Unterschiede in der Reaktion auf die Erhebungssituation zeigen. Ebenso könnte die Dauer der bisherigen Nutzung von Tracking-Apps eine Rolle spielen: Personen mit mehr Erfahrung im Ernährungstracking haben womöglich bereits stabile Routinen entwickelt und sind daher weniger anfällig für Verhaltensänderungen im Zuge von Forscherbeobachtung, während weniger erfahrene Nutzer*innen stärker auf die Messsituation reagieren könnten. Dabei werden jeweils ausschließlich die ersten beiden Tage von Woche 2 berücksichtigt, da Reaktivitätseffekte vermutlich von relativ kurzer Dauer sind und primär in den ersten zwei Tagen der Erhebungsphase auftreten.

Die vorliegende Arbeit verfolgt daher das Ziel, bestehende Forschungslücken im Bereich der Messreaktivität bei der digitalen Ernährungserfassung zu schließen. Dabei wird untersucht, ob und in welchem Ausmaß sich die tägliche Erfassung der Nahrungsaufnahme mit gängigen Ernährungstracking-Apps unter Forscherbeobachtung auf das Essverhalten auswirkt. Grundlage bilden dabei bereits erhobene Daten aus Ernährungstracking-Apps, welche im Alltag der Teilnehmenden gesammelt wurden und anschließend mit den unter Forscherbeobachtung erhobenen Daten verglichen werden. Mithilfe dieser innovativen Methode soll die Rolle der Forscherbeobachtung, inklusive der potenziellen Effektmodifikatoren Geschlecht und Trackingdauer isoliert werden und im Kontext des täglichen Ernährungsverhaltens untersucht werden.

Forschungsfragen und Hypothesen

Auf Basis der dargestellten theoretischen Überlegungen und empirischen Befunde entstanden daher drei Forschungsfragen, aus welchen anschließend die folgenden drei in der Arbeit untersuchten Hypothesen abgeleitet wurden. Abweichend von der Präregistrierung wurde die geplante explorative Hypothese E4 verworfen.

RQ 1: Verändert sich der tägliche Proteinkonsum unter Forscherbeobachtung im Vergleich zu einer Woche ohne Forscherbeobachtung?

H1: Der Proteinkonsum ist in den ersten beiden Tagen von Woche 2 höher als der Ausgangswert von Woche 1.

RQ 2: Moderiert das Geschlecht den Effekt der Forscherbeobachtung auf den Proteinkonsum?

H2: Die Steigerung des Proteinkonsums von Woche 1 zu den ersten beiden Tagen von Woche 2 fällt bei Frauen stärker aus als bei Männern.

RQ 3: Wie beeinflusst die Dauer des Ernährungstrackings den Zusammenhang zwischen Forscherbeobachtung und dem Proteinkonsum?

H3: Eine kürzere Trackingdauer ist mit einer stärkeren Steigerung des Proteinkonsums in den ersten zwei Tagen von Woche 2 im Vergleich zu Woche 1 assoziiert.

Methode

Aufbauend auf den ausgeführten theoretischen Überlegungen wird im Folgenden die methodische Vorgehensweise der vorliegenden Studie dargestellt. Die zugrunde liegenden Daten wurden im Rahmen eines übergeordneten Forschungsprojekts erhoben (König & Ruf, 2024), auf dessen Basis mehrere Masterarbeiten verfasst werden (Förg, 2025; Lichtenberger, 2025). Die Arbeit wurde in der arbeitsbereichsinternen Studiendatenbank des Arbeitsbereichs Gesundheitspsychologie vor Durchführung der Datenauswertung präregistriert, die finale Version der Präregistrierung wurde im Mai 2025 eingereicht. Materialien, Daten und Auswertungsskripte dieser Arbeit sind in der arbeitsbereichsinternen Studiendatenbank des Arbeitsbereichs Gesundheitspsychologie hinterlegt. Jegliche Abweichungen von der Präregistrierung werden in den jeweiligen Abschnitten beschrieben. Der Titel der Arbeit wurde nachträglich von „Messreaktivität bei der digitalen Ernährungserfassung: Der Effekt der Studienteilnahme auf den alltäglichen Proteinkonsum“ zu „Messreaktivität bei der digitalen Ernährungserfassung: Verändert Forscherbeobachtung den alltäglichen Proteinkonsum?“ geändert.

Stichprobe

Da diese Studie die erste Untersuchung von Messreaktivität auf die tägliche Nahrungsaufnahme mithilfe einer Ernährungstracking-App darstellt, existieren keine vergleichbaren Pilotdaten für die Stichprobenumfangsplanung. Daher wurde stattdessen das Rekrutieren von $N = 200$ Teilnehmenden als Ziel gesetzt, nachdem die Anzahl Ende Juni 2025 noch nicht erreicht wurde, wurden die Analysen in der vorliegenden Arbeit mit den bis dahin erhobenen Daten durchgeführt. Die Datensammlung für das übergeordnete Forschungsprojekt läuft dagegen noch bis November 2025 weiter.

Teilnahmeberechtigt waren Personen ab einem Mindestalter von 18 Jahren, die bereits vor Studienbeginn seit mindestens einer Woche eine mobile App zur Ernährungserfassung nutzten, mit der sich Kilokalorien erfassen lassen. Ausgeschlossen wurden Teilnehmende, welche beide Aufmerksamkeitstests im ersten Fragebogen nicht bestanden oder den zweiten Fragebogen nicht innerhalb von 24 Stunden abschlossen.

Die Rekrutierung erfolgte über das Online-Rekrutierungssystem (LABS) der Universität Wien, die Website surveycircle.com, soziale Medien, durch Flyer sowie Mundpropaganda. Die Rekrutierung wurde in einem Zeitraum von jeweils einer Woche vor bis einer Woche nach großen Feiertagen beziehungsweise Ferien (Weihnachten und Ostern) pausiert, um Verzerrungen des alltäglichen Essverhaltens durch abweichende Routinen in den Ferien zu vermeiden. Die Studie wurde eine Woche vor Semesterende geschlossen und zwei Wochen nach Semesterbeginn wieder geöffnet, um typische

Ernährungsgewohnheiten während des Semesters abzubilden. Teilnehmende erhielten wahlweise 4 LABS-Credits (für Psychologiestudierende der Universität Wien) oder nahmen an einer Verlosung von drei 100€ Wunschgutscheinen (wunschgutschein.at/.ch/.de) teil.

Im finalen Datensatz befanden sich schlussendlich insgesamt 75 Personen ($n = 75$). Die Stichprobe setzte sich aus 61.3 % weiblichen und 38.7 % männlichen Teilnehmenden zusammen, niemand identifizierte sich als divers. Das durchschnittliche Alter betrug 27.6 Jahre ($SD = 8.16$), der durchschnittliche Body-Mass-Index (BMI) lag bei 24.5 ($SD = 4.05$). Hinsichtlich des höchsten Bildungsabschlusses verfügten die Teilnehmenden überwiegend über eine allgemeine Hochschulreife (92.0 %), der mediane Ausbildungsabschluss war ein Hochschulabschluss (45.3 %). Unter den Teilnehmenden, die ein Studienfach angaben, studierte die Mehrheit (59,5 %; $n = 25$) Psychologie. Kognitions- und Medienwissenschaft sowie Wirtschaftswissenschaften wurden jeweils von 2 Personen (4,8 %) genannt; alle weiteren Studienfächer wurden nur einmal angegeben (2.4 %). Die am häufigsten genutzten Ernährungstracking-Apps waren *Yazio* (42.7 %), *MyFitnessPal* (22.7 %) sowie *FDDDB* (8.0 %) und *Lifesum* (8.0 %). Eine vollständige Übersicht der deskriptiven Statistiken der Stichprobe ist in Tabelle 1 zu finden. Abweichend von der Präregistrierung wurde das monatliche Nettoeinkommen nicht mit dem Median und Interquartilsabstand angeführt, sondern in absoluten Häufigkeiten und Prozentwerten. Zudem wurden alle im Fragebogen erfragten demographischen Angaben für die Stichprobenbeschreibung angegeben.

Tabelle 1

Deskriptive Stichprobenbeschreibung

Merkmale	n (%)	M (SD)
Alter (in Jahren)	75 (100 %)	27.6 (8.16)
BMI	-	24.5 (4.05)
Geschlecht	-	-
weiblich	46 (61.3 %)	-
männlich	29 (38.7 %)	-
divers	0 (0.0 %)	-
Höchster Schulabschluss	-	-
Volksschul- / Hauptschulabschluss	0 (0.0 %)	-
Mittlere Reife / Realschlussabschluss	1 (1.3 %)	-
Abschluss der Polytechnischen Oberschule	0 (0.0 %)	-
Abschluss einer (Berufs-)Fachschule	3 (4.0 %)	-
Hochschulreife (Abitur/ Matura)	69 (92.0 %)	-

Anderer Abschluss	2 (2.7 %)	-
Höchster Bildungsabschluss	-	-
(Noch) keine abgeschlossene Berufsausbildung	25 (33.3 %)	-
Lehre	3 (4.0 %)	-
Fachoberschule/Berufsoberschule	6 (8.0 %)	-
Fachhochschule	4 (5.3 %)	-
Hochschule	34 (45.3 %)	-
Promotion	2 (2.7 %)	-
Anderer Abschluss	1 (1.3 %)	-
Derzeitige Beschäftigung	-	-
berufstätig (in Vollzeit)	19 (25.3 %)	-
berufstätig (in Teilzeit)	7 (9.3 %)	-
arbeitssuchend	3 (4.0 %)	-
in Ausbildung/ Umschulung/ Studium	46 (61.3 %)	-
ausschließlich Hausfrau/mann	0 (0.0 %)	-
Rentner/ Rentnerin	0 (0.0 %)	-
Studienfach ¹	42 (56 %)	-
Psychologie	25 (59.5 %)	-
Kognitions- und Medienwissenschaft	2 (4.8 %)	-
Wirtschaftswissenschaften	2 (4.8 %)	-
Monatliches Nettoeinkommen	-	-
weniger als 150€	4 (5.3 %)	-
150€ bis 300€	1 (1.3 %)	-
300€ bis 500€	4 (5.3 %)	-
500€ bis 1.000€	13 (17.3 %)	-
1.000€ bis 1.500€	11 (14.7 %)	-
1.500€ bis 2.000€	7 (9.3 %)	-
2.000€ bis 2.500€	8 (10.7 %)	-
2.500€ bis 3.000€	8 (10.7 %)	-
3.000€ bis 5.000€	7 (9.3 %)	-
5.000€ bis 10.000€	10 (13.3 %)	-
über 10.000€	2 (2.7 %)	-

Genutzte Tracking-App	-	-
Yazio	32 (42.7 %)	-
MyFitnessPal	17 (22.7 %)	-
Fddb	6 (8.0 %)	-
Lifesum	6 (8.0 %)	-
FatSecret	3 (4.0 %)	-
Noom	2 (2.7 %)	-
Samsung Health	2 (2.7 %)	-
Barmer Fit2Go	1 (1.3 %)	-
Calories and nutrition tracker	1 (1.3 %)	-
Fitbit	1 (1.3 %)	-
Foodvisor	1 (1.3 %)	-
HappyCat Fit	1 (1.3 %)	-
MacroFactor	1 (1.3 %)	-

Anmerkung. ¹ Studienfächer mit nur einer Nennung werden aus Gründen der Übersichtlichkeit nicht dargestellt.

Maße und Materialien

Die Studie umfasst zwei digitale Fragebögen, eine Übersicht der Fragebogeninhalte kann der folgenden Tabelle 2 entnommen werden:

Tabelle 2

Übersicht Studienmaterialien

Erhobene Daten	1. Teil	2. Teil	Quelle
Soziodemographische Daten (Alter, Geschlecht, höchster Schulabschluss, höchster Bildungsabschluss, monatliches Nettoeinkommen, Beschäftigung, BMI)	x		
Erfahrung und Nutzen von Ernährungsapp (Verwendete App, Trackingdauer, angedachte Verwendungszwecke der App)	x		König et al. (2018)
Tägliche Kalorien- und Makronährstoffaufnahme der letzten 7 Tage	x	x	König et al. (2025)

Anmerkung. Nur für die vorliegende Arbeit relevante Inhalte dargestellt; vollständiger Fragebogen umfasst zusätzliche Items für weitere Arbeiten.

Design und Ablauf

Der Ablauf basiert auf dem Studiendesign von König et al. (2025), auf dem die vorliegende Arbeit angelehnt ist. Es handelt sich um eine beobachtende Studie mit einem Within-Subjects-Design. Anhand zwei digitaler Fragebögen wurden relevante Outcomes zu 2 Zeitpunkten im Abstand von einer Woche erhoben. Zuerst füllten die Teilnehmenden den 1. Teil des Online-Fragebogen aus, in dem sie soziodemografische und anthropometrische Merkmale angaben, wofür sie ihre Ernährungs-App nutzen sowie die Aufzeichnungen ihrer Kilokalorien- und Makronährstoffaufnahme der letzten 7 Tage auflisteten. Anschließend gaben sie ihre E-Mail-Adresse an, mit der sie genau eine Woche später zum 2. Teil der Umfrage eingeladen wurden, in der sie erneut die Kilokalorien- und Makronährstoffzufuhr der letzten 7 Tage anführen sollten. Das Bewusstsein über die Forscherbeobachtung wurde somit für die folgende Woche induziert. Nach Abschluss des 2. Teils wurden die Teilnehmenden über die Intention der Studie aufgeklärt. Bis dahin war die einzige ihnen bekannte Information, dass sie ihre Nährwertinformationen an die Wissenschaft spenden.

Datenauswertung

Im folgenden Abschnitt wird das Vorgehen bei der Datenbereinigung, der Vorverarbeitung sowie der deskriptiven und inferenzstatistischen Analyse der erhobenen Daten beschrieben. Alle Analysen wurden in R (Version 4.5.0; R Core Team, 2025) durchgeführt.

Datenbereinigung

Zunächst wurden die Daten vor der Analyse manuell auf Tippfehler kontrolliert, beschreibende Angaben wie das Studienfach angeglichen (beispielsweise „Wiwi“ zu „Wirtschaftswissenschaften“) und Werte für die Berechnung des BMI (kg/m^2) vereinheitlicht. 24 Teilnehmende mit fehlenden Proteinangaben an mindestens 6 der 7 Tage von Fragebogen 1 und/oder den ersten beiden Tagen von Fragebogen 2 wurden aus der Analyse ausgeschlossen. Für die verbliebenen Daten wurden Tage ohne Einträge als fehlend gewertet, diese 0-Werte im Datensatz wurden entsprechend in *NA* umgewandelt.

Da Multilevel-Modelle grundsätzlich robust gegenüber fehlenden Daten sind, können bei teilweise fehlende Daten, etwa Proteinangaben einzelner Tage, trotzdem alle vorhandenen Daten der Person berücksichtigt werden (Hox et al., 2017). Die fehlenden Werte der täglichen Proteinaufnahmen für Woche 1 wurden in absoluten sowie relativen Werten geprüft. Die Proteinwerte der ersten beiden Tage von Woche 2 waren, wie in dem Abschnitt zur Datenbereinigung beschrieben, verpflichtend bei jedem/jeder Teilnehmer*in vollständig.

In Woche 1 lagen bei 8 Teilnehmenden geringe Anteile fehlender Werte vor (jeweils 1 oder 2 fehlende Messwerte; 12 insgesamt), während die restlichen 67 Teilnehmenden keine fehlenden Werte aufwiesen. Der Little's MCAR-Test, wie in der Präregistrierung erwähnt, wurde anschließend nicht durchgeführt, da die fehlenden Werte ausschließlich in Woche 1 in geringer Anzahl auftraten und weit unter 5 % der gesamten Proteindaten in Woche 1 ausmachten (2.3 %), wodurch die Ergebnisse nicht sinnvoll interpretierbar wären (Schafer & Graham, 2002).

Deskriptivstatistische Auswertung

Zu Beginn der statistischen Analyse wurde eine deskriptive Auswertung der in der Untersuchung erhobenen soziodemografischen Variablen vorgenommen. Das Alter der Teilnehmenden wurde mit Mittelwert und Standardabweichung beschrieben. Der Body-Mass-Index (BMI) wurde auf Grundlage der selbstberichteten Körpergröße in Meter und des Körpergewichts in Kilogramm berechnet und ebenfalls mit Mittelwert und Standardabweichung angegeben. Das Geschlecht wurde in den Kategorien weiblich, männlich und divers erfasst und prozentual ausgewertet. Für den höchsten Bildungsabschluss, den Ausbildungsabschluss sowie das monatliche Nettoeinkommen erfolgte eine deskriptive Darstellung in Form von Häufigkeiten und Prozentanteilen je Kategorie. Zudem wurde die Nutzung der verschiedenen Tracking-Apps erfasst und in absoluten sowie prozentualen Häufigkeiten angegeben.

Im Anschluss wurden die deskriptiven Kennwerte der ab- und unabhängigen Variablen sowie der Moderatorvariablen berechnet. Der durchschnittliche Proteinkonsum in Woche 1 (M & SD) und den ersten beiden Tagen von Woche 2 (M & SD) wurde jeweils angegeben, die gleichen deskriptiven Kennwerte wurden zudem nach Geschlecht berichtet. Die Trackingdauer (in Monaten) wurde des Weiteren als metrische Variable erfasst und ebenfalls mit Mittelwert und Standardabweichung beschrieben. Zur Veranschaulichung der deskriptiven Statistik wurde zusätzlich eine terzilibasierte Gruppierung der Trackingdauer vorgenommen, wobei auch hier der wöchentliche Proteinkonsum ausgewertet wurde. Überall wurde jeweils der Mittelwert sowie die Standardabweichung ausgerechnet.

Voraussetzungstests

Zur Überprüfung der Modellannahmen wurden für jedes Hypothesenmodell separate Analysen durchgeführt. Mithilfe von *DHARMA* (Hartig, 2024) erfolgte zunächst die Prüfung auf gleichmäßige Verteilung der standardisierten Residuen mittels Kolmogorov-Smirnov-Test. Anschließend wurde mit dem Dispersionstest auf Heteroskedastizität getestet, zur Identifikation potenzieller Ausreißer wurde ein binomialer Ausreißertest verwendet. Des Weiteren wurde zur Überprüfung der Angemessenheit

des Modellfits ein Vergleich der Modelle anhand der Informationsmaße AIC (Akaike-Informationskriterium) und BIC (Bayes'sches Informationskriterium) sowie mittels Likelihood-Ratio-Test durchgeführt. Zusätzlich wurde, sofern mehr als ein Prädiktor im Modell enthalten war (H2 & H3) Multikollinearität anhand von Varianzinflationsfaktoren (VIF) überprüft.

Das Modell zur Prüfung von Hypothese 1 zeigte eine gute Modellgüte, mit AIC = 1499,71 und BIC = 1511,75. Der Kolmogorov-Smirnov-Test ergab keinen Hinweis auf Abweichungen der standardisierten Residuen ($D = 0.092$, $p = .158$), der Dispersionstest zeigte außerdem keine Anzeichen für Heteroskedastizität (Dispersion = 1.012, $p = .864$). Der Ausreißertest ergab nur eine Beobachtung an den Randbereichen, was statistisch nicht signifikant war ($p = 1.00$). Eine Prüfung auf Multikollinearität wurde aufgrund des einzigen Prädiktors nicht durchgeführt. Der Singularitätscheck ergab keinen Hinweis auf Singularität.

Für das Modell der Hypothese 2 (AIC = 1462.88; BIC = 1480.94) zeigte der Kolmogorov-Smirnov-Test keine signifikanten Abweichungen von der Gleichverteilungsannahme der Residuen ($D = 0.065$, $p = .544$). Der Dispersionstest ergab keine Hinweise auf Heteroskedastizität (Dispersion = 1.011, $p = .864$). Der Ausreißertest zeigte zwei Beobachtungen an den Randbereichen, ohne signifikant zu sein ($p = .336$). Die Multikollinearitätsprüfung ergab niedrige VIF-Werte zwischen 1.16 und 1.79, was auf keine problematische Kollinearität hinweist (Hair et al., 2019). Der Singularitätscheck war unauffällig.

Das Modell zu Hypothese 3 (AIC = 1501.48; BIC = 1519.54) wies ebenfalls keine signifikanten Abweichungen von der Gleichverteilung der Residuen auf ($D = 0.084$, $p = .240$). Der Dispersionstest zeigte keine Hinweise auf Heteroskedastizität (Dispersion = 1.012, $p = .864$), der Ausreißertest identifizierte eine Beobachtung an den Randbereichen, jedoch ohne statistische Signifikanz ($p = 1.00$). Die Multikollinearitätsprüfung zeigte niedrige VIF-Werte zwischen 1.00 und 1.10 und damit keine problematische Kollinearität. Auch hier zeigte der Singularitätscheck keine Auffälligkeiten.

Modellvergleiche

Zur Überprüfung, ob komplexere Modellstrukturen eine signifikante Verbesserung der Modellpassung bewirken, wurden Likelihood-Ratio-Tests zwischen genesteten Modellen durchgeführt. Für H1 wurde zunächst ein Nullmodell mit Random Intercept (M1a) geschätzt und mit einem Modell mit Haupteffekt der Woche (M1b) verglichen. Für H2 wurde ein Modell mit Haupteffekt der Woche (M2a) einem Modell mit zusätzlicher Interaktion Woche $\times$ Geschlecht (M2b) gegenübergestellt. Für H3 wurde ein Modell mit Haupteffekt der Woche (M3a) mit einem Modell verglichen, das zusätzlich die Interaktion Woche $\times$ Trackingdauer enthielt (M3b).

Die Ergebnisse (siehe Tabelle 3) ergaben, dass nur das Modell mit der Interaktion zwischen Woche und Geschlecht (M2b) eine signifikant bessere Passung aufwies und somit für die Testung von Hypothese 2 verwendet wurde. Alle anderen Modellvergleiche ergaben keine signifikanten Verbesserungen und wurden daher verworfen.

Tabelle 3

Vergleich genesteter Modelle

Modell	df	-2LL	AIC	BIC	Δdf	χ^2	p
M1a	3	6239.0	6245.0	6258.5	-	-	-
M1b	4	6237.4	6245.4	6263.3	1	1.64	.200
M2a	4	6237.4	6245.4	6263.3	-	-	-
M2b	6	6197.3	6209.3	6236.2	2	40.09	<.001
M3a	4	6237.4	6245.4	6263.3	-	-	-
M3b	6	6236.1	6248.1	6275.0	2	1.27	.531

Anmerkung. M1a = Nullmodell mit Random Intercept (H1); M1b = Modell mit Haupteffekt Woche (H1); M2a = Modell mit Haupteffekt Woche (H2); M2b = Modell mit Interaktion Woche $\times$ Geschlecht (H2); M3a = Modell mit Haupteffekt Woche (H3); M3b = Modell mit Interaktion Woche $\times$ Trackingdauer (H3). -2LL = $-2 \times \log\text{-Likelihood}$.

Inferenzstatistische Auswertung

Zur Untersuchung der Forschungsfragen wurde der durchschnittliche tägliche Proteinkonsum in Gramm als abhängige Variable (AV) operationalisiert. Die unabhängige Variable (UV) „Woche“ wurde als dichotome Variable kodiert (0 = Woche 1, 1 = Tage 1 & 2 von Woche 2). Im Zuge der inferenzstatistischen Auswertung wurden lineare Multilevel-Modelle mit einem Random Intercept auf Personenebene spezifiziert, um die hierarchische Datenstruktur (2 Messzeitpunkte) innerhalb der Teilnehmenden adäquat abzubilden. Ausgewertet wurden die Tageswerte des Proteinkonsums in Gramm als Level-1-Beobachtungen, verschachtelt in Personen (Level-2). Die Modelle wurden mit *lme4* (Bates et al., 2025) berechnet, die dazugehörigen p-Werte für Fixed Effects mit *lmerTest* (Kuznetsova et al., 2020). Die allesamt gerichteten Hypothesen wurden jeweils zweiseitig getestet. Als Schätzmethode wurde die Maximum-Likelihood-Methode verwendet, um Modellvergleiche mittels Likelihood-Ratio-Tests durchführen zu können.

Für Hypothese 1 wurde ein Modell mit Woche als Prädiktor geschätzt und dann der Einfluss dieser UV auf den Proteinkonsum untersucht. Für Hypothese 2 wurde das Geschlecht als potenzieller Moderator in das Modell aufgenommen. Das Geschlecht wurde als dichotome Variable kodiert (0 = weiblich, 1 = männlich), nachfolgend wurde eine Cross-Level-Interaktion mit der UV Woche spezifiziert, um mögliche geschlechtsspezifische Unterschiede in der Veränderung des Proteinkonsums

zu erfassen. Da sich keine Person als divers identifizierte, wurde die Gruppe nicht für die Analyse berücksichtigt. Hypothese 3 prüfte die Rolle der bisherigen Erfahrung mit Ernährungstracking auf den Proteinkonsum. Hierzu wurde die Trackingdauer in Monaten als kontinuierliche Moderatorvariable verwendet, zentriert und ebenfalls als Cross-Level-Interaktion mit der UV Woche modelliert.

Die Signifikanz der Prädiktoren wurde anhand von Likelihood-Ratio-Tests bestimmt, wobei jeweils Modelle mit und ohne den relevanten Prädiktor verglichen wurden. Zusätzlich wurden Effektstärken in Form der Varianzaufklärer Within- R^2 , Between- R^2 und Total- R^2 mit *r2mlm* (Shaw et al., 2025) ermittelt, abweichend von der Präregistrierung statt marginalen und konditionalen R^2 .

Ergebnisse

Nachfolgend werden die Ergebnisse der Modellvoraussetzungstests, Deskriptivstatistiken sowie zur Beantwortung der Fragestellungen durchgeführten inferenzstatistischen Verfahren berichtet.

Deskriptive Statistiken

Der mittlere tägliche Proteinkonsum lag in Woche 1 bei 99.8 g ($SD = 43.7$), in den ersten beiden Tagen von Woche 2 bei 102.0 g ($SD = 49.6$). Für die vollständigen deskriptiven Statistiken siehe Tabelle 4.

Tabelle 4

Deskriptive Statistiken des täglichen Proteinkonsums (in Gramm)

Gruppe	Woche 1: M (SD)	Woche 2 ¹ : M (SD)
Gesamt	99.8 (43.7)	102.0 (49.6)
Geschlecht	-	-
Weiblich ($n = 46$)	78.1 (27.5)	78.8 (28.9)
Männlich ($n = 29$)	134.0 (42.8)	140.0 (52.8)
Trackingdauer ($M = 16.1$, $SD = 27.6$) ²	-	-
Unteres Terzil ($n = 25$; $M = 1.16$, $SD = 0.45$)	83.2 (32.7)	78.1 (31.9)
Mittleres Terzil ($n = 25$; $M = 4.96$, $SD = 1.81$)	91.3 (34.4)	102.0 (41.5)
Oberes Terzil ($n = 25$; $M = 42.0$, $SD = 35.7$)	125.0 (51.2)	127.0 (60.0)

Anmerkung. ¹ Ersten beiden Tage von Woche 2; ² Trackingdauer in Monaten.

Ergebnisse Hypothese 1: Effekt der Forscherbeobachtung auf den täglichen Proteinkonsum

Zur Prüfung der ersten Hypothese wurde untersucht, ob sich der durchschnittliche Proteinkonsum unter Forscherbeobachtung von Woche 1 im Vergleich zu den ersten beiden Tagen von Woche 2 systematisch veränderte. Der Regressionskoeffizient des linearen Multilevel-Modells wies zwar auf eine leichte Zunahme der täglichen Proteinaufnahme hin ($b = 2.69$), fiel jedoch statistisch nicht signifikant aus ($p = .399$). Die varianzaufklärenden Anteile deuten auf überwiegend interindividuelle Unterschiede hin (Total- $R^2 = .825$; Within- $R^2 = .005$; Between- $R^2 = .000$). Die vollständigen Modellparameter sind in Tabelle 5 dargestellt.

Tabelle 5*Parameter Hypothese 1*

Variable	<i>b</i>	95 % CI	<i>SE</i>	<i>t</i>	<i>df</i>	<i>p</i>	Within- <i>R</i> ²	Between- <i>R</i> ²	Total- <i>R</i> ²
Intercept	99.79	[89.28, 110.30]	5.36	18.61	89.29	< .001	-	-	-
Woche	2.69	[-3.53, 8.92]	3.18	0.85	75	.399	.005	.000	.825

Ergebnisse Hypothese 2: Interaktion von Forscherbeobachtung und Geschlecht

Für Hypothese 2 wurde ein Modell mit Interaktion zwischen Woche und Geschlecht aufgestellt, um zu untersuchen, ob Frauen stärker auf die induzierte Forscherbeobachtung reagieren als Männer. Ein Cross-Level-Interaktionseffekt mit der Woche ließ sich dabei nicht nachweisen, die Unterschiede zwischen den Geschlechtern blieben über beide Zeitpunkte hinweg stabil ($p = .419$). Unabhängig davon zeigten die Ergebnisse allerdings einen deutlichen Haupteffekt des Geschlechts: Männer berichteten über beide Wochen hinweg signifikant höhere Proteinaufnahmen als Frauen ($b = 56.02$; $p = < .001$). Das Modell erklärte fast keine Varianz innerhalb der Personen (Within- $R^2 = 0.0003$), während sich ein hoher Anteil der Varianzerklärung zwischen den Personen ergab (Between- $R^2 = 0.437$). Eine detaillierte Darstellung der Modellparameter findet sich in Tabelle 6.

Tabelle 6*Parameter Hypothese 2*

Variable	<i>b</i>	95 % CI	<i>SE</i>	<i>t</i>	<i>df</i>	<i>p</i>	Within- <i>R</i> ²	Between- <i>R</i> ²	Total- <i>R</i> ²
Intercept	78.13	[67.56, 88.71]	5.40	14.48	98.79	< .001	-	-	-
Woche	0.65	[-7.26, 8.57]	4.04	0.16	75.00	.872	-	-	-
Geschlecht	56.02	[39.02, 73.03]	8.68	6.46	98.79	< .001	.0003	.437	.820
Woche × Geschlecht	5.28	[-7.45, 18.01]	6.49	0.81	75.00	.419	-	-	-

Ergebnisse Hypothese 3: Interaktion von Forscherbeobachtung und Trackingdauer

Zur Überprüfung der Hypothese, dass eine kürzere Trackingdauer mit einer stärkeren Steigerung des Proteinkonsums in Woche 2 assoziiert ist, wurde ein Cross-Level-Interaktionsmodell spezifiziert. Die Ergebnisse zeigten weder einen signifikanten Haupteffekt der Trackingdauer ($p = .359$) noch einen signifikanten Interaktionseffekt ($p = .395$). Das Modell erklärte nur sehr geringe

Varianzanteile sowohl innerhalb ($\text{Within-}R^2 = .005$) als auch zwischen den Personen ($\text{Between-}R^2 = .01$). Eine Übersicht der Modellparameter ist in Tabelle 7 aufgeführt.

Tabelle 7

Parameter Hypothese 3

Variable	<i>b</i>	95 % CI	<i>SE</i>	<i>t</i>	<i>df</i>	<i>p</i>	Within- R^2	Between- R^2	Total- R^2
Intercept	99.79	[89.38, 110.20]	5.31	18.79	89.43	< .001	-	-	-
Woche	2.69	[-3.50, 8.89]	3.16	0.85	75.00	.397	-	-	-
Trackingdauer	0.18	[-0.20, 0.56]	0.19	0.92	89.43	.359	.005	.014	.825
Woche × Trackingdauer	0.10	[-0.13, 0.33]	0.12	0.86	75.00	.395	-	-	-

Diskussion

Die vorliegende Arbeit untersuchte, ob Messreaktivität im Rahmen digitaler Ernährungserfassung auftritt. Unter Anwendung einer neuartigen Methode, bei der die bereits verfügbaren Daten aus den Ernährungstracking-Apps der Teilnehmenden genutzt wurden, sollte die Rolle der Forscherbeobachtung bei reaktiven Effekten isoliert und die vorhandene Lücke von Studien zu dieser Thematik geschlossen werden. Konkret wurde geprüft, ob der Proteinkonsum zwischen der bereits vorhandenen ersten Woche und den ersten beiden Tagen der zweiten Woche unter Forscherbeobachtung anstieg (H1), ob Frauen einen größeren Anstieg des Proteinkonsums als Männer zu verzeichnen hatten (H2) und ob eine kürzere Vorerfahrung mit Ernährungstracking-Apps einen verstärkenden Einfluss auf das Ausmaß reaktiver Effekte hatte (H3).

Insgesamt zeigten die Ergebnisse, dass in der durchgeführten Untersuchung keine statistisch signifikanten Reaktivitätseffekte auftraten. Zwischen der unbeobachteten Woche 1 und den ersten beiden Tagen von Woche 2 unter Forscherbeobachtung ließ sich nur ein minimaler, nicht signifikanter Anstieg im Proteinkonsum feststellen. Hinsichtlich der reaktiven Geschlechtsunterschiede ergab sich zudem kein signifikanter Cross-Level-Interaktionseffekt, wodurch auch Hypothese 2 nicht bestätigt werden konnte. Darüber hinaus zeigte sich ein signifikanter Haupteffekt des Geschlechts, Männer berichteten unabhängig von der Forscherbeobachtung über beide Wochen hinweg höhere Proteinwerte. Nicht zuletzt konnte auch Hypothese 3 nicht bestätigt werden, die Trackingdauer erklärte keine Unterschiede der Reaktionen auf die Studienteilnahme.

Einordnung in den Forschungsstand und Interpretation

Die zentrale Hypothese der vorliegenden Arbeit ging durch die indizierte Forscherbeobachtung von einer Steigerung des Proteinkonsums von Woche 1 zu den ersten beiden Tagen von Woche 2 aus. Die Daten der durchgeführten Untersuchung lieferten hierfür jedoch keine Unterstützung. Die Ergebnisse zeigten zwar einen leichten Anstieg ($b = 2.69$), dieser fiel jedoch statistisch nicht signifikant aus ($p = .399$) und wies einen relativ hohen Standardfehler ($SE = 3.18$) auf. Das Modell erklärte fast ausschließlich Varianz zwischen den Personen ($Total-R^2 = .825$), während auf inter- und intraindividueller Ebene nahezu keine Varianz aufgeklärt wurde ($Between-R^2 = .000$; $Within-R^2 = .005$). In dieser Untersuchung lag somit keine Evidenz für eine systematische Veränderung des Proteinkonsums zwischen Woche 1 und Woche 2 vor, der Großteil der erklärten Varianz stammt aus individuellen Unterschieden im Proteinkonsum, welche unabhängig von der Forscherbeobachtung bestanden.

Da diese Studie die erste derartige Analyse der Rolle der Forscherbeobachtung bei digitaler Ernährungserfassung darstellt, können die Ergebnisse nicht direkt mit bestehender Evidenz verglichen

werden. Die Befunde fügen sich jedoch in das noch relativ vage Bild bezüglich Messreaktivität im digitalen, gesundheitsbezogenen Kontext ein. Während in bisherigen Studien überwiegend kleine bis mittlere Effekte berichtet wurden (König et al., 2022a) und damit zunächst den vorliegenden Ergebnissen widersprechen, beziehen sich diese eben größtenteils auf die Untersuchung physischer Aktivität. Eine aktuellere Studie von König et al. (2025) zu körperlicher Aktivität ergab allerdings ebenfalls, dass die Einführung von Forscherbeobachtung ohne zusätzlichen Einsatz eines ungewohnten Trackinggeräts keine relevanten reaktiven Effekte hervorruft. Auch in einer einjährigen Untersuchung des selbstberichteten Nahrungsergänzungsmittelkonsum zeigte sich kein signifikanter Studienteilnahmeeffekt (Steffen et al., 2021), was für die Heterogenität und Komplexität der untersuchten Thematik spricht. Dementsprechend erscheint es notwendig, die Bedingungen genauer zu bestimmen, unter denen digitale Messungen tatsächlich zu Reaktivität führen und zwischen Verhaltensdomänen zu differenzieren.

Generelle Erklärungsansätze

Ein möglicher Grund für das Ausbleiben signifikanter Reaktivitätseffekte könnte unter anderem darin liegen, dass Protein als Makronährstoff in der Alltagswahrnehmung vieler Menschen nicht dieselbe Salienz besitzt wie etwa Kalorien, Salz oder Zucker. Während diese Merkmale beispielsweise auf Lebensmittelverpackungen oder der Ernährungsberatung stark betont werden, spielt Protein in der Ernährungskommunikation und allgemeinen Wahrnehmung eine weniger prominente Rolle (Plasek et al., 2020; Song et al., 2021). Im Gegensatz dazu wird der Kalorienwert direkt mit Gewichtskontrolle und einer gesunden Ernährung assoziiert (Fernandes et al., 2019), weswegen er möglicherweise salienter als der Proteinwert wirkt und eher eine bewusste Verhaltensanpassung auslösen könnte. Wenn Teilnehmende also in Woche 2 insgesamt weniger Kalorien zu sich nahmen, wäre es widersprüchlich, gleichzeitig eine Erhöhung der Proteinaufnahme zu erwarten, da sich die Makronährstoffzufuhr grundsätzlich proportional zur Gesamtenergiezufuhr verändert (Sofi et al., 2025). Selbst eine geringfügige Verringerung der Kalorienaufnahme könnte somit bereits dazu geführt haben, dass ein möglicher relativer Anstieg des Proteinanteils nicht zu einem absoluten Anstieg der Proteinmenge führte.

Die Salienz der einzelnen Inhaltsstoffe in der genutzten Tracking-App besitzt eventuell ebenfalls einen Einfluss darauf, ob Nutzer*innen ihr Ernährungsverhalten aktiv regulieren. Informationen, die visuell hervorgehoben oder sozial stark betont werden, wirken sich stärker auf Entscheidungen aus als weniger präsente Aspekte (Dai et al., 2020). Wenn Protein in gängigen Ernährungstracking-Apps nicht so deutlich hervorgehoben wird wie etwa die Kalorien, bleibt sein Einfluss auf das Essverhalten möglicherweise begrenzt. Protein könnte also in Apps zur digitalen

Ernährungserfassung im Vergleich zu anderen Ernährungsparametern nicht in gleichem Maße als salient beziehungsweise gesundheitsrelevant wahrgenommen werden.

Ein weiterer plausibler Grund dafür, dass in der vorliegenden Studie keine signifikante Steigerung des Proteinkonsums zu beobachten war, besteht in der Stabilität etablierter Ess- und Kochroutinen. Essroutinen wie feste Mahlzeiten, bestimmte Zutaten oder bevorzugte Zubereitungsformen sind häufig in alltägliche Abläufe, Wertesysteme sowie soziale Kontexte integriert und die Veränderung solcher Ernährungspraktiken benötigt oft mehr Zeit oder wiederholte Interventionen (Jastran et al., 2009). Alltagspraktische Hürden wie zeitliche Einschränkungen, finanzielle Belastungen oder eine eingeschränkte Verfügbarkeit bestimmter proteinhaltiger Lebensmittel können Anpassungen zusätzlich erschweren (Pinho et al., 2018). Eine Steigerung des Proteinkonsums benötigt aufgrund dieser Einbettung in alltägliche Ernährungsroutinen beziehungsweise eventuell vorhandener Barrieren also möglicherweise mehr Zeit als andere reaktive Effekte, um sich in messbarem Verhalten niederzuschlagen.

Auch mangelndes Wissen über Protein und dessen Quellen stellen Schwierigkeiten dar, welche kurzfristige Veränderungen unter Forscherbeobachtung potenziell weiter erschweren können. Laut einer repräsentativen Umfrage des International Food Information Council (2025) kann 80% der US-amerikanischen Bevölkerung die empfohlene tägliche Proteinzufuhr nicht benennen, selbst in protein-affinen Gruppen wird dieser großteils unterschätzt (≤ 50 g/Tag). Viele Konsument*innen verwechseln zudem natürlich proteinreiche Lebensmittel mit künstlich proteinangereicherten Lebensmitteln (Banovic et al., 2018) oder stufen proteinhaltige Lebensmittel nur dann als relevant ein, wenn sie beispielsweise durch High-Protein-Etiketten explizit gekennzeichnet sind (Rramani et al., 2023). Aufgrund des relativ hohen Proteinkonsums in der vorliegenden Stichprobe ($M = 99.8$ g; $SD = 43.7$ in Woche 1) im Vergleich zu den für europäische Länder berichteten Durchschnittswerten (67–114 g/Tag bei Männern, 59–102 g/Tag bei Frauen; EFSA, 2012) kann zwar davon ausgegangen werden, dass dies auf einen Großteil der Studienteilnehmer*innen nicht zutrifft, für zukünftige Forschung mit unter Umständen abweichender Stichprobenzusammensetzung sollte dieser Aspekt jedoch nicht missachtet werden.

Im Gegensatz zu traditionellen Erhebungsmethoden ist die Forscherbeobachtung bei der Verwendung digitaler Geräte möglicherweise weniger salient, da die Teilnehmenden den Ernährungstracking-Prozess eventuell eher als Teil ihrer Alltagsroutine denn als Forschungsaktivität wahrnehmen. Hinzu kommt, dass digitale Erhebungen in der Regel zeitlich beziehungsweise räumlich flexibel erfolgen und weniger mit formalen Rahmenbedingungen wie etwa Laborsituationen, direkten Befragungen oder expliziten Instruktionen verknüpft sind. In der vorliegenden Untersuchung hatten die Teilnehmenden keinen direkten Kontakt mit den Forscher*innen selbst, des Weiteren musste nur

zweimal, jeweils zu Beginn und am Ende der Studie, ein kurzer Online-Fragebogen ausgefüllt werden, wodurch die Studienteilnahme in Kombination mit den routinierten digitalen Selbsterhebungen möglicherweise weniger bewusst erschien. Zudem wurde die Forscherbeobachtung in der durchgeführten Untersuchung nur zu Beginn und am Ende des 1. Fragebogens mit einer kurzen Erläuterung der weiteren Vorgehensweise induziert. Unter Umständen war diese Darstellung nicht ausreichend genug und Teilnehmende fassten die Mitteilung eher als organisatorische Information auf, wodurch die Beobachtung nicht in ausreichendem Maße salient wirkte, um eine bewusste Verhaltensanpassung auszulösen.

In diesem Zusammenhang könnte auch die bereits bestehende Vertrautheit der Teilnehmenden mit ihrer eigenen Tracking-App sowie deren fortgesetzte Nutzung einen abschwächenden Einfluss auf potenzielle Reaktivitätseffekte haben. Es lässt sich vermuten, dass Personen, die bereits regelmäßig ihr Gesundheitsverhalten dokumentieren, zu Beginn einer Studie keine erhöhte Aufmerksamkeit für dieses Verhalten entwickeln und damit weniger anfällig für Messreaktivität sind. Diese Annahme wurde bislang empirisch kaum geprüft (König et al., 2022a), in der Studie von König et al. (2025) zeigte sich allerdings, dass die bloße Einführung von Forscherbeobachtung ohne zusätzlichen Einsatz eines ungewohnten Messinstruments kaum zu nachweisbaren Reaktivitätseffekten im Hinblick auf körperliche Aktivität führt. In Bezug auf die digitale Ernährungserfassung erscheint es somit plausibel, dass die eigene Tracking-App von den Teilnehmenden nicht als neues oder ungewohntes Messinstrument wahrgenommen wurde und sie bereits mit der Funktionsweise und dem Feedback der App vertraut waren. Möglicherweise entstand aus diesem Grund zu Beginn der Studie keine gesteigerte Aufmerksamkeit für das eigene Verhalten, sodass keine Studienteilnahmeeffekte ausgelöst wurden, die das Ernährungsverhalten beziehungsweise den Proteinkonsum hätten beeinflussen können. Reaktivität tritt im Kontext der digitalen Ernährungserfassung wohl eher als Reaktion auf die Einführung einer Messung auf und weniger auf der Fortführung dieser. Signifikante Ergebnisse hätten im Umkehrschluss darauf hingedeutet, dass die Verhaltensänderungen in erster Linie durch den Studienkontext ausgelöst wurden. Es muss daher zumindest hinterfragt werden, ob in der vorliegenden Studie Forschende zusammen mit den bereits verwendeten Tracking-Apps zu ausreichend teilnehmenden Beobachter*innen im Alltag der Teilnehmenden wurden, um das Ernährungsverhalten tatsächlich zu beeinflussen.

Damit einhergehend stellen Deckeneffekte beziehungsweise unterschiedliche individuelle Ausgangsniveaus einen weiteren möglichen Mechanismus dar, welcher das Ausbleiben signifikanter reaktiver Effekte erklären könnte. Teilnehmende, die bereits zu Beginn der Erhebung ein verhältnismäßig hohes Proteinniveau aufwiesen, besitzen etwa aufgrund von zeitlichen, finanziellen,

oder physiologischen Faktoren, vermutlich nur begrenzten Spielraum für weitere Steigerungen. Der hohe Mittelwert der vorliegenden Stichprobe in Woche 1 deutet darauf hin, dass in dieser Erhebung ein erhöhtes Potenzial für derartige Deckeneffekte bestand. Selbst wenn diese Teilnehmenden durch die Forscherbeobachtung dazu geneigt gewesen wären, ihr Verhalten zu ändern, könnte ein bereits hoher Ausgangskonsum somit dazu geführt haben, dass eine weitere Erhöhung des alltäglichen Proteinkonsums nur schwierig möglich war. Dies passt ebenfalls zu den Ergebnissen von König et al. (2025), wo Teilnehmende mit niedriger körperlicher Aktivität vor Studienbeginn ihre Schrittzahl nach Beginn der Untersuchung verhältnismäßig steigerten, möglicherweise weil ihnen die Studienteilnahme den Abstand zwischen ihrem aktuellen Verhalten und gesellschaftlich kommunizierten Normen besonders salient machte. Umgekehrt reduzierten schon relativ aktive Teilnehmende ihre Aktivität während der Erhebung leicht, was als Hinweis darauf interpretiert werden kann, dass bereits hohe Ausgangswerte im Vergleich zu gängigen Normen eine Art „Sättigungseffekt“ hervorrufen könnten: Teilnehmende fühlen sich durch die Beobachtung in ihrem Verhalten bestätigt und sehen keinen Anlass, dieses weiter zu steigern. Man könnte somit annehmen, dass durch Messreaktivität induzierte mögliche Veränderungen in gewisser Hinsicht von individuellen Ausgangsniveaus abhängig sind.

Ein alternativer methodischer Erklärungsansatz ergibt sich aus der Datenerhebung selbst. Unter Forscherbeobachtung könnten die Teilnehmenden ihr Trackingverhalten verändert haben, ohne dass sich ihr tatsächliches Essverhalten veränderte. So erscheint es plausibel, dass die Dokumentation während der Studienteilnahme in Woche 2 sorgfältiger und vollständiger erfolgte als in Woche 1. Dadurch könnte der in der App berichtete Proteinkonsum steigen, ohne dass tatsächlich mehr Protein verzehrt wurde. Vielmehr ist sogar denkbar, dass der reale Proteinkonsum unter Umständen leicht zurückging, dieser Rückgang aber durch die größere Vollständigkeit der Dokumentation beziehungsweise Verbesserung der Trackinggenauigkeit ausgeglichen wurde. Es könnte sich somit um ein Messartefakt handeln, die erhaltenen Ergebnisse sollten daher nicht ausschließlich als direkte Abbildung des tatsächlichen Ernährungsverhaltens interpretiert werden.

Untersuchte Moderatoren

Für die zweite Hypothese wurde ein signifikanter Haupteffekt des Geschlechts gefunden ($p < .001$), wobei Männer über beide Messzeitpunkte hinweg rund 56 g mehr Protein pro Tag berichteten als Frauen. Der postulierte Interaktionseffekt zwischen der Woche und dem Geschlecht ($b = 5.28$, 95% CI [-7.45, 18.01]) erwies sich hingegen als nicht signifikant ($p = .419$), Frauen reagierten demnach nicht stärker auf die Forscherbeobachtung als Männer, der Effekt lief tendenziell sogar in die umgekehrte Richtung. Die Verteilung der erklärten Varianz verdeutlichte dies weiter: Während

innerhalb der Personen keine nennenswerte Varianz aufgeklärt wurde ($\text{Within-}R^2 = 0.0003$), erklärte das Modell fast die Hälfte der Varianz zwischen den Personen ($\text{Between-}R^2 = 0.437$). Das Geschlecht stellte in der durchgeführten Studie also einen Prädiktor für die bereits bestehende absolute Differenz im täglichen Proteinkonsum dar, nicht jedoch für die eigentlich angenommene Messreaktivität.

Der signifikante Haupteffekt, wonach Männer täglich mehr Protein dokumentierten als Frauen, deckt sich mit Befunden, in denen Männern im Allgemeinen ein gesteigerter Proteinkonsum zugeschrieben wird. Ein in der Literatur häufig diskutierter Erklärungsansatz sind grundsätzlich unterschiedliche Ernährungsverhalten, welche sich aus geschlechtsspezifischen Faktoren wie sozialen und biologischen Unterschieden ergeben. Generell verfügen Männer meist über mehr Masse sowie Muskelmasse als Frauen, was zu einem höheren Energiebedarf und dadurch gleichzeitig mehr konsumierten Protein führt (Kanter & Caballero, 2012). Männer neigen außerdem grundsätzlich dazu, einen höheren Proteinkonsum anzustreben, oft motiviert durch Fitness- oder Krafttraining-Ziele, wodurch proteinreiche Lebensmittel wie Fleisch stärker in den Ernährungsfokus rücken. Frauen hingegen sind tendenziell stärker auf die Reduktion von Kalorien fokussiert (Lombardo et al., 2024). Männer besitzen zudem ein positiveres, gesünderes Bild von Fleisch als Frauen und essen öfter genussorientiert (Lombardo et al., 2020), Frauen orientieren sich demgegenüber stärker an gesundheitsförderlichen Ernährungsweisen und sozialen Normen, wodurch sie durchschnittlich weniger Fleisch konsumieren. Schließlich könnten auch geschlechtsspezifische Essmuster dazu beitragen: Während Männer insgesamt wenige größere Portionen und energiereichere Mahlzeiten zu sich nehmen, konsumieren Frauen an einem Tag eher mehrere leichtere Speisen, wodurch ihr Proteinanteil tendenziell niedriger ausfallen könnte (Feraco et al., 2024).

Die für den signifikanten Haupteffekt angeführten geschlechtsspezifischen Erklärungsansätze können darüber hinaus in gewissem Maß herangezogen werden, um das Ausbleiben eines signifikanten Interaktionseffekt zwischen dem Geschlecht und der Forscherbeobachtung zu erklären. Die Annahme von Hypothese 2 stützte sich primär auf die Überlegung, dass Frauen tendenziell sensibler gegenüber sozialen Normen im Ernährungsverhalten sind und häufiger berichten, sich an gängigen Ernährungsempfehlungen zu orientieren. Wie im obigen Absatz ausgeführt, spielt Protein als Makronährstoff für Männer allerdings vermutlich eine größere Rolle als für Frauen. Die durch Messreaktivität angenommenen gesundheitsförderlichen Effekte könnten sich daher geschlechtsbezogen unterschiedlich äußern. Frauen fokussieren sich tendenziell stärker auf die Reduktion von Kalorien als auf einen hohen Proteinkonsum (Lombardo et al., 2024), da der Kilokalorienkonsum in der vorliegenden Arbeit nicht untersucht wurde, ist es möglich, dass diese Interaktion relevantere Ergebnisse als jene mit dem Proteinkonsum zeigt.

Der ausbleibende geschlechtsspezifische Interaktionseffekt reiht sich in die bislang uneinheitliche Evidenz bezüglich der Rolle des Geschlechts bei Messreaktivität ein. Abgesehen von den bereits erwähnten 6 identifizierten Studien in der Übersichtsarbeit zu Reaktivitätseffekten bei digitalen Messungen von König et al. (2022), bei denen in 5 Studien keine signifikanten Unterschiede zwischen Männern und Frauen gefunden wurden, fanden König et al. (2025) in ihrer Untersuchung einen größeren Anstieg an physischer Aktivität bei Männern. Die gefundenen Unterschiede in Messreaktivität zwischen Männern und Frauen wurden dabei unter anderem durch geschlechtsspezifische Lebensumstände und Motivationen erklärt. Frauen stehen im Alltag häufig vor der Herausforderung, berufliche und familiäre Verpflichtungen gleichzeitig zu bewältigen, was weniger Raum für gesundheitsbezogene Verhaltensweisen wie Ernährung oder körperliche Aktivität lässt und entsprechend als Hindernis wirken kann. Männer hingegen zeigen tendenziell ein stärkeres Interesse an Wettbewerben und könnten die Teilnahme an Studien eher als persönliche Herausforderung interpretieren, ihr Verhalten zu steigern. Dies könnte dazu beitragen, dass Männer in manchen Untersuchungen größere reaktive Effekte zeigen.

Auch die im Kontext von Hypothese 3 formulierte Interaktion zwischen Forscherbeobachtung und Trackingdauer erwies sich als nicht signifikant. Die Interaktion mit der Woche zeigte keinen Zusammenhang ($p = .395$) und auch der Haupteffekt der Trackingdauer war statistisch nicht signifikant ($p = .359$), der Regressionskoeffizient für die Interaktion ($b = 0.10$, 95% CI [-0.13, 0.33]) lag weiters nahe bei null. Auch hier zeigte sich, dass die erklärte Varianz fast vollständig auf bestehende Unterschiede zwischen den Personen zurückzuführen war ($\text{Total-}R^2 = .825$), während das Modell innerhalb und zwischen den Personen praktisch keine systematische Veränderung erklären konnte ($\text{Within-}R^2 = .005$; $\text{Between-}R^2 = .014$).

Ein naheliegender Erklärungsansatz stellt die Zusammenstellung der Stichprobe im Hinblick auf die Trackingdauer dar. Bei Betrachtung der zur Veranschaulichung gebildeten Terzile fällt die hohe durchschnittliche Trackingdauer sowie Varianz im oberen Terzil ($M = 42.0$, $SD = 35.7$) auf, welche auf eine deutliche Heterogenität innerhalb dieser erfahreneren Nutzer*innen hindeutet. Mit dem unteren ($M = 1.16$, $SD = 0.45$) und mittleren Terzil ($M = 4.96$, $SD = 1.81$) wiesen zwei Drittel der Teilnehmenden dagegen relativ ähnliche und deutlich geringere Nutzungswerte auf, wodurch sich vor allem Personen mit sehr viel als auch mit relativ wenig Erfahrung in der Stichprobe befanden.

Aus theoretischer Perspektive könnte die mitunter rasche Etablierung von Routinen bei der Nutzung von Selbstmonitoring-Tools eine mögliche Begründung für das Ausbleiben signifikanter Interaktionseffekte darstellen. Automatisierte Gewohnheiten bei gesundheitsbezogenen Verhaltensweisen können bei regelmäßiger Wiederholung bereits nach wenigen Wochen entstehen (Judah et al., 2013; Kaushal & Rhodes, 2015; Lally et al., 2010). Darüber hinaus scheint die Beziehung

zwischen Wiederholungsfrequenz und Gewohnheitsbildung nicht linear, sondern asymptotisch zu verlaufen, ab einem bestimmten Punkt führt zusätzliche Erfahrung also nicht mehr zu stärkeren Habitualisierungseffekten (Gardner et al., 2012; Lally et al., 2010). Vor diesem Hintergrund könnte die ursprüngliche Annahme einer kontinuierlich abnehmenden Reaktivität mit zunehmender Trackingdauer zu stark vereinfacht sein: Wenn Routinen bereits nach kurzer Zeit stabil werden, unterscheiden sich Teilnehmende mit wenigen Monaten Erfahrung möglicherweise kaum noch von jenen, die ihre Ernährung schon über mehrere Jahre dokumentieren.

Die Ergebnisse von König et al. (2025), in denen bei Personen, die bereits vor Studienbeginn regelmäßig Aktivitätstracker verwendeten, keine messbare Reaktivität infolge der Forscherbeobachtung festgestellt wurde, stützen diese Interpretation ebenfalls. In der Studie bestand die Stichprobe konträr zu der vorliegenden Arbeit überwiegend aus erfahrenen Nutzer*innen, von denen rund zwei Drittel (68.9 %) ihre Aktivität bereits seit mindestens einem Jahr mit dem eigenen Gerät aufzeichneten. Zusammengenommen mit der vergleichsweise unerfahreneren Stichprobe der vorliegenden Untersuchung, bei der ebenfalls keine reaktiven Effekte beobachtet wurden, spricht dies dafür, dass die Dauer der bisherigen Nutzung insgesamt einen eher geringen Einfluss auf das Auftreten von Reaktivitätseffekten ausübt. Entscheidend scheint weniger die Länge der Nutzung, sondern vielmehr die (Weiter-)Nutzung des vertrauten Messinstruments zu sein.

Stärken und Schwächen der vorliegenden Studie

Die bedeutendste Stärke dieser Arbeit liegt in der Anwendung einer in diesem Kontext neuartigen Methode, bei welcher die bereits vorhandenen Daten aus den Ernährungstracking-Apps der Teilnehmenden verwendet wurden, um die Rolle der Forscherbeobachtung auf digitale Ernährungserhebungen zu untersuchen. Durch den Einsatz von tatsächlich verwendeten, populären Apps unter alltäglichen Bedingungen konnten ökologisch valide Daten zum Proteinkonsum erhoben werden. Durch Anwendung dieser Methode konnten methodische Einschränkungen früherer Studien zu Messreaktivität umgangen werden, welche häufig auf den Vergleich verschiedener Messinstrumente angewiesen waren oder davon ausgingen, dass zu Beginn der Erhebung eine Verhaltensänderung auftritt, die sich im weiteren Studienverlauf wieder normalisiert (König et al., 2025). Das Design der vorliegenden Untersuchung erlaubte es hingegen, das Ernährungsverhalten im Rahmen der Studie direkt mit dem vorangegangenen Ernährungsverhalten zu vergleichen, womit die Effekte der Forscherbeobachtung von Effekten, die durch ein neuartiges Messinstrument entstehen, getrennt wurden. Die vorhandene Lücke von Studien zu dieser Thematik konnte somit durch erste Evidenz stellenweise adressiert werden, auch durch die Miteinbeziehung des Geschlechts und der Trackingdauer als bisher kaum untersuchte Moderatorvariablen.

Die vorliegende Arbeit besitzt dennoch einige Limitationen. Die Stichprobengröße umfasste aufgrund fehlender Pilotdaten und der noch laufenden Datenerhebung nur 75 Personen. Eine Fallzahl von weniger als 100 birgt das Risiko, dass Effekte mittlerer Größe nicht zuverlässig detektiert werden können (Cohen, 1988). Die statistische Power der Untersuchung war daher gewissermaßen eingeschränkt, wobei die erhaltenen Ergebnisse weit von statistischer Signifikanz beziehungsweise relevanten Effekten entfernt waren. Weiters befand sich die Mehrheit der Teilnehmenden in Ausbildung, Umschulung oder einem Studium (61.3 %), 45.3 % besaßen darüber hinaus bereits einen Hochschulabschluss. Zudem gab ein erheblicher Teil der studierenden Teilnehmenden ein psychologisches Studienfach an (59.5 %; 64.3 % wenn Kognitions- und Medienwissenschaft miteinbezogen wird). Dieses Muster deckt sich zumindest teilweise mit der allgemein häufig genannten Kritik, dass ein erheblicher Teil psychologischer Forschung auf Stichproben basiert, welche überwiegend westlich, gebildet, industrialisiert, wohlhabend und demokratisch geprägt sind (Perski et al., 2022). Diese Zusammensetzung resultiert in einer eingeschränkten Generalisierbarkeit dieser Untersuchung, womit sich die Ergebnisse so nur bedingt auf andere kulturelle oder sozioökonomische Kontexte übertragen lassen. Gleichzeitig könnte die Relevanz dieses Problems im konkreten Fall geringer sein, da Messreaktivität, sofern sie überhaupt auftritt, nach bisherigem Forschungsstand ein weitgehend populationsunabhängiges Phänomen darstellt und Moderatoren, wie auch in dieser Studie, selten signifikante Effekte ergeben (Arigo & König, 2024).

Trotz der beschriebenen Vorteile, welche durch die Weiternutzung der bisherigen Ernährungstracking-App entstanden, war außerdem als Folge der unterschiedlichen Tracking-Apps die Homogenität der verwendeten Messgeräte nicht gewährleistet. Da alle inkludierten Personen jedoch zwangsläufig eine App verwenden mussten, in welcher der tägliche Proteinkonsum erfasst werden kann und in Studien keine statistisch signifikanten Unterschiede in der Wirksamkeit verschiedener Funktionen unterschiedlicher Ernährungs-Apps festgestellt wurde (Pala et al., 2024), sollte dies jedoch kein methodisch relevantes Problem darstellen. Die Daten beruhen überdies notwendigerweise auf Selbstberichten, was keine objektive Einschätzung des Ernährungsverhaltens ermöglicht, self-reported dietary records weisen jedoch generell eine hohe Validität auf (Thompson & Subar, 2017).

Schließlich besteht eine weitere Limitation der vorliegenden Arbeit darin, dass die untersuchten Moderatoren – Geschlecht und Trackingdauer – selbst bei theoretisch signifikanten Ergebnissen nur dann sinnvoll interpretiert hätten werden können, wenn tatsächlich interindividuelle Unterschiede in der Reaktivität vorgelegen wären. Das Testen von Moderatoren auf ihre Relevanz, ohne zuvor nachgewiesen zu haben, dass in der untersuchten Stichprobe überhaupt bedeutsame Variabilität in den Reaktivitätsmustern existiert, stellt ein methodisches Problem dar, welches in vielen bestehenden Studien über Messreaktivität vorhanden ist (Arigo & König, 2024). Fehlt eine solche

Variabilität, können die untersuchten Moderatoren per se keine Unterschiede erklären. In der vorliegenden Untersuchung ist die Aussagekraft unabhängig von der nicht vorhandenen Signifikanz der getesteten Moderatoreffekte somit eingeschränkt, da keine Varianzkomponenten der individuellen Reaktivitätsverläufe geschätzt wurden.

Implikationen und zukünftige Forschung

Die vorliegenden Ergebnisse tragen unter anderem zur Forschung über Messreaktivität bei, indem sie implizieren, dass digitale Ernährungserhebungen mit vergleichsweise geringerem Risiko für starke Reaktivitätseffekte im Hinblick auf den Proteinkonsum eingesetzt werden können, solange die Teilnehmenden ihre bestehende Tracking-App weiterverwenden. Dies eröffnet praktisch die Möglichkeit, selbsterhobene Proteinangaben stärker in klinische und epidemiologische Studien zu integrieren, ohne erhebliche ökologische Validitätsverluste befürchten zu müssen. Dennoch sollte die Validität von digitalen self-reported dietary records mithilfe von objektiven Biomarkern wie *doubly-labelled water* (Burrows et al., 2019) sowohl allgemein als auch im Zusammenhang mit Studienteilnahmeeffekten weiter untersucht werden.

Wie in den Limitationen dargelegt, war die Stichprobe der vorliegenden Untersuchung relativ homogen, wodurch die Generalisierbarkeit der Resultate nur eingeschränkt möglich ist. Künftige Studien sollten daher größere und diversere Stichproben beinhalten, um die Übertragbarkeit der Ergebnisse auf unterschiedliche Populationen besser beurteilen zu können. Damit zusammenhängend ist bislang nicht gänzlich klar, in welchem Ausmaß unterschiedliche Bevölkerungsgruppen von Messreaktivität betroffen sind, speziell im digitalen Kontext. Die Nutzung von Smartphones und Wearables ist bei Jugendlichen und jungen Erwachsenen deutlich stärker verbreitet als bei Personen mittleren oder höheren Alters. Eine hohe Vertrautheit mit digitalen Technologien könnte dazu beitragen, dass bestimmte Mechanismen, die Messreaktivität auslösen, abgeschwächt werden (Maher et al., 2024). Vor diesem Hintergrund erscheint es sinnvoll, künftige Studien gezielt populationsspezifisch anzulegen, um mögliche Unterschiede zwischen Altersgruppen systematisch zu erfassen und besser zu verstehen.

In der vorliegenden Untersuchung wurden ausschließlich Proteinwerte analysiert, während andere Makronährstoffe sowie speziell das Gesamtkalorienniveau unberücksichtigt blieben. Dadurch bleibt vorerst offen, ob Messreaktivität bei anderen Ernährungsparametern aufgetreten sein könnte und sich die nicht nachgewiesenen Reaktivitätseffekte spezifisch auf den Proteinkonsum beschränken. Für zukünftige Forschung wäre es daher sinnvoll, neben Protein auch weitere Ernährungsaspekte wie etwa Fett, Zucker und insbesondere den Kalorienkonsum systematisch einzubeziehen, um ein umfassenderes Bild möglicher reaktiver Veränderungen bei der digitalen Ernährungserfassung zu

gewinnen. Die künftig verfassten Folgearbeiten auf Basis des erhobenen Datensatzes werden hierbei vermutlich zumindest teilweise Abhilfe schaffen können, auch durch die im vollständigen Fragebogen erfassten, jedoch in dieser Studie nicht berücksichtigten Variablen wie Ernährungsintentionen, Gewissenhaftigkeit oder qualitative Rückmeldungen zu wahrgenommenen Verhaltensänderungen. Damit könnten weitere wertvolle Informationen über das subjektive Erleben der Teilnehmenden sowie über potenzielle Verhaltensanpassungen gewonnen werden. Für zukünftige ähnlich angelegte Untersuchungen oder Replikationsstudien wäre es weiters sinnvoll, relative Makronährstoffanteile in Prozent der Gesamtenergiezufuhr zusätzlich zu den absoluten Mengen zu analysieren, um Veränderungen in Ernährungsmustern differenzierter abbilden zu können.

Das diskutierte eventuell aufgetretene Messartefakt, wo ein Rückgang des tatsächlich konsumierten Proteins in Woche 2 durch eine größere Vollständigkeit der Dokumentation ausgeglichen werden könnte, sollte durch Vergleiche der Trackinggenauigkeit zwischen Woche 1 und Woche 2 ebenfalls weiter untersucht werden. Man könnte argumentieren, dass diese gesteigerte Nutzung selbst eine Form von Messreaktivität darstellt, wenn auch nicht in der ursprünglich angedachten Veränderung des Verhaltens, sondern der Dokumentation. Da häufigeres und konsistenteres Tracking im Rahmen von selbsterhobenen Ernährungsdokumentationen mit einer höheren Wirksamkeit und genaueren Erfassung des Ernährungsverhaltens verbunden ist (Raber et al., 2021; Thompson & Subar, 2017; Turner-McGrievy et al., 2019), könnte diese potenzielle Manifestation von Messreaktivität unter Umständen als positiver Mechanismus der Selbstregulation verstanden und genutzt werden.

Schließlich sollten in zukünftigen Studien zu der Thematik Personen mit und ohne Erfahrung im digitalen Ernährungstracking systematisch verglichen werden, um die Aussagekraft der vorliegenden Untersuchung zu überprüfen und den Anwendungsbereich zu erweitern. Ergänzend dazu sollte bei weiterführender Forschung im Idealfall ein experimentelles Design wie das Solomon-Vier-Gruppen-Design (French et al., 2021a) eingesetzt werden. Dabei könnte durch die Kombination von Gruppen mit und ohne Baseline-Erhebung sowie mit und ohne Forscherbeobachtung noch gezielter geprüft werden, ob beobachtete Veränderungen im Ernährungsverhalten auf die Messung selbst oder auf die Intervention beziehungsweise den Studienkontext zurückzuführen sind.

Konklusion

Die vorliegende Arbeit untersuchte unter Verwendung bereits vorhandener Trackingdaten den Einfluss von Forscherbeobachtung auf die digitale Ernährungserfassung anhand des täglichen Proteinkonsums. Dabei ließ sich keine auftretende Messreaktivität feststellen, auch die getesteten Moderatoren Geschlecht und Trackingdauer erwiesen sich als nicht signifikant. Die Ergebnisse deuten

darauf hin, dass die Weiternutzung bereits vertrauter Ernährungstracking-Apps im Rahmen einer Studie zu keinen relevanten reaktiven Effekten führt, zumindest im Hinblick auf den Proteinkonsum. Die Arbeit leistet somit einen relevanten Beitrag zur Adressierung der bestehenden Forschungslücke hinsichtlich reaktiver Effekte bei app-basierter Ernährungserfassung. Zugleich unterstreichen die Befunde vor dem Hintergrund der Stichprobenzusammensetzung und des Studiendesigns die Notwendigkeit, in zukünftigen Studien andere Ernährungsparameter wie den Kalorienkonsum zu untersuchen sowie auch unerfahrene Teilnehmende und diversere Stichproben zu berücksichtigen.

Literaturverzeichnis

- Allmeta, A., Arigo, D., & König, L. M. (2022). *Easy or difficult? Investigating perceived ease to change aspects of eating behaviour and physical activity*. Abgerufen von osf.io/mqxyu
- Arentson-Lantz, E., Clairmont, S., Paddon-Jones, D., Tremblay, A., & Elango, R. (2015). Protein: A nutrient in focus. *Applied Physiology, Nutrition, and Metabolism*, 40(8), 755–761. <https://doi.org/10.1139/apnm-2014-0530>
- Arias, J. V., Guerrero, C. C., Espinoza, L. G., Jácome, M. Á. E., & Arizala, J. F. C. (2025). Healthy eating and nutrition, a review of the theoretical aspects of a healthy diet. *AG Salud*, 3, 112. <https://doi.org/10.62486/agsalud2025112>
- 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>
- Banovic, M., Arvola, A., Pennanen, K., Duta, D. E., Brückner-Gühmann, M., Lähteenmäki, L., & Grunert, K. G. (2018). Foods with increased protein content: A qualitative study on European consumer preferences and perceptions. *Appetite*, 125, 233–243. <https://doi.org/10.1016/j.appet.2018.01.034>
- Bardus, M., van Beurden, S. B., Smith, J. R., & Abraham, C. (2016). A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. *International Journal of Behavioral Nutrition and Physical Activity*, 13(1), 35. <https://doi.org/10.1186/s12966-016-0359-9>
- 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). Guildford Press.
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Grothendieck, G., & Green, P. (2025). *lme4: Linear mixed-effects models using Eigen and S4* (Version 1.1-37) [R package]. <https://cran.r-project.org/package=lme4>

- Bendtsen, M., & McCambridge, J. (2021). Causal models accounted for research participation effects when estimating effects in a behavioral intervention trial. *Journal of Clinical Epidemiology*, 136, 77–83. <https://doi.org/10.1016/j.jclinepi.2021.03.008>
- Bernardi, R. A., & Guptill, S. T. (2008). Social desirability response bias, gender, and factors influencing organizational commitment: An international study. *Journal of Business Ethics*, 81, 797–809. <https://doi.org/10.1007/s10551-007-9548-4>
- 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, 850. <https://doi.org/10.3389/fendo.2019.00850>
- Catania, J. A., Gibson, D. R., Chitwood, D. D., & Coates, T. J. (1990). Methodological problems in AIDS behavioral research: Influences on measurement error and participation bias in studies of sexual behavior. *Psychological Bulletin*, 108(3), 339–362. <https://doi.org/10.1037/0033-2909.108.3.339>
- 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>
- Clemes, S. A., & Deans, N. K. (2012). Presence and Duration of Reactivity to Pedometers in Adults. *Medicine and Science in Sports and Exercise*, 44(6), 1097–1101. <https://doi.org/10.1249/MSS.0b013e318242a377>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Routledge. <https://doi.org/10.4324/9780203771587>
- Dai, J., Cone, J., & Moher, J. (2020). Perceptual salience influences food choices independently of health and taste preferences. *Cognitive Research: Principles and Implications*, 5(1), 2–13. <https://doi.org/10.1186/s41235-019-0203-2>
- Drewnowski A., Moskowitz H., Reisner M., Krieger B. (2010). Testing consumer perception of nutrient content claims using conjoint analysis. *Public Health Nutrition*, 13(5), 688–694. <https://doi.org/10.1017/S1368980009993119>

- Eisele, G., Vachon, H., Lafit, G., Tuyaerts, D., Houben, M., Kuppens, P., Myin-Germeys, I., & Viechtbauer, W. (2023). A mixed-method investigation into measurement reactivity to the experience sampling method: The role of sampling protocol and individual characteristics. *Psychological assessment, 35*(1), 68–81. <https://doi.org/10.1037/pas0001177>
- Eldridge, A. L., Piernas, C., Illner, A.-K., Gibney, M. J., Gurinović, M. A., De Vries, J. H. M., & Cade, J. E. (2019). Evaluation of New Technology-Based Tools for Dietary Intake Assessment—An ILSI Europe Dietary Intake and Exposure Task Force Evaluation. *Nutrients, 11*(1), 55. <https://doi.org/10.3390/nu11010055>
- Fagerli, R. A., & Wandel, M. (1999). Gender Differences in Opinions and Practices with Regard to a „Healthy Diet“. *Appetite, 32*(2), 171–190. <https://doi.org/10.1006/appe.1998.0188>
- Feraco, A., Armani, A., Gorini, S., Camajani, E., Quattrini, C., Filardi, T., Karav, S., Strollo, R., Caprio, M., & Lombardo, M. (2024). Gender Differences in Dietary Patterns and Eating Behaviours in Individuals with Obesity. *Nutrients, 16*(23), 4226. <https://doi.org/10.3390/nu16234226>
- Fernandes, A. C., Rieger, D. K., & Proença, R. P. C. (2019). Perspective: Public Health Nutrition Policies Should Focus on Healthy Eating, Not on Calorie Counting, Even to Decrease Obesity. *Advances in Nutrition, 10*(4), 549–556. <https://doi.org/10.1093/advances/nmz025>
- Förg, H. (2025). *Messreaktivität im digitalen Ernährungstracking: Die Rolle von Ernährungszielen und wahrgenommener Forscherbeobachtung* (Unveröffentlichte Masterarbeit). Universität Wien.
- 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>
- French, D. P., Miles, L. M., Elbourne, D., Farmer, A., Gulliford, M., Locock, L., Sutton, S., & McCambridge, J., MERIT Collaborative Group. (2021a). Reducing bias in trials due to reactions to measurement: Experts produced recommendations informed by evidence. *Journal of Clinical Epidemiology, 139*, 130–139. <https://doi.org/10.1016/j.jclinepi.2021.06.028>

- French, D. P., Miles, L. M., Elbourne, D., Farmer, A., Gulliford, M., Locock, L., Sutton, S., & McCambridge, J., MERIT Collaborative Group. (2021b). Reducing bias in trials from reactions to measurement: The MERIT study including developmental work and expert workshop. *Health Technology Assessment*, 25(55), 1–72. <https://doi.org/10.3310/hta25550>
- Fritz, T., Huang, E. M., Murphy, G. C., & Zimmermann, T. (2014). Persuasive technology in the real world: a study of long-term use of activity sensing devices for fitness. In *CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 487–496). <https://doi.org/10.1145/2556288.2557383>
- Gardner, B., Lally, P., & Wardle, J. (2012). Making health habitual: the psychology of 'habit-formation' and general practice. *British Journal of General Practice*, 62(605), 664–666. <https://doi.org/10.3399/bjgp12X659466>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage.
- Harnack, L. J., & Pereira, M. A. (2017). Strategies for Improving the Validity of the 24-hour Dietary Recall and Food Record Methods. In D. A. Schoeller & M. S. Westerterp-Plantenga (Eds.), *Advances in the Assessment of Dietary Intake* (1st ed., pp. 67–84). CRC Press. <https://doi.org/10.1201/9781315152288-4>
- Hartig, F. (2024). *DHARMA: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models* (Version 0.4.7) [R package]. <https://cran.r-project.org/package=DHARMA>
- Healthy diet*. (2020). World Health Organization. <https://www.who.int/news-room/fact-sheets/detail/healthy-diet>
- Hebert, J. R., Ma, Y., Clemow, L., Ockene, I. S., Saperia, G., Stanek, E. J. III, Merriam, P. A., & Ockene, J. K. (1997). Gender differences in social desirability and social approval bias in dietary self-report. *American Journal of Epidemiology*, 146(12), 1046–1055. <https://doi.org/10.1093/oxfordjournals.aje.a009233>

- Ho, V., Simmons, R. K., Ridgway, C. L., van Sluijs, E. M., Bamber, D. J., Goodyer, I. M., Dunn, V. J., Ekelund, U., & Corder, K. (2013). Is wearing a pedometer associated with higher physical activity among adolescents? *Preventive Medicine*, 56(5), 273–277. <https://doi.org/10.1016/j.ypmed.2013.01.015>
- Hox, J., Moerbeek, M., & van de Schoot, R. (2017). *Multilevel Analysis: Techniques and Applications* (3rd ed.). Routledge. <https://doi.org/10.4324/9781315650982>
- IFIC Spotlight Survey: Americans' Perceptions of Protein. (2025). International Food Information Council. <https://ific.org/research/ific-spotlight-survey-americans-perceptions-of-protein>
- Jastran, M. M., Bisogni, C. A., Sobal, J., Blake, C., & Devine, C. M. (2009). Eating routines. Embedded, value based, modifiable, and reflective. *Appetite*, 52(1), 127–136. <https://doi.org/10.1016/j.appet.2008.09.003>
- Judah, G., Gardner, B., & Aunger, R. (2013). Forming a flossing habit: an exploratory study of the psychological determinants of habit formation. *British journal of health psychology*, 18(2), 338–353. <https://doi.org/10.1111/j.2044-8287.2012.02086.x>
- Kanter, R., & Caballero, B. (2012). Global Gender Disparities in Obesity: A Review. *Advances in Nutrition*, 3(4), 491–498. <https://doi.org/10.3945/an.112.002063>
- Kaushal, N., & Rhodes, R. E. (2015). Exercise habit formation in new gym members: a longitudinal study. *Journal of behavioral medicine*, 38(4), 652–663. <https://doi.org/10.1007/s10865-015-9640-7>
- Kiefer, I., Rathmanner, T., & Kunze, M. (2005). Eating and dieting differences in men and women. *Journal of Men's Health and Gender*, 2(2), 194–201. <https://doi.org/10.1016/j.jmhg.2005.04.010>
- Kimura, M., Moriyasu, A., & Makizako, H. (2022). Positive Association between High Protein Food Intake Frequency and Physical Performance and Higher-Level Functional Capacity in Daily Life. *Nutrients*, 14(1), 72. <https://doi.org/10.3390/nu14010072>

- 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>
- 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., Allmeta, A., Christlein, N., Van Emmenis, M., & Sutton, S. (2022a). 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. <https://doi.org/10.1080/17437199.2022.2047096>
- König, L. M., Attig, C., Franke, T., & Renner, B. (2021). Barriers to and Facilitators for Using Nutrition Apps: Systematic Review and Conceptual Framework. *JMIR mHealth and uHealth*, 9(6), e20037. <https://doi.org/10.2196/20037>
- König, L. M., Sproesser, G., Schupp, H. T., & Renner, B. (2018). Describing the Process of Adopting Nutrition and Fitness Apps: Behavior Stage Model Approach. *JMIR mHealth and uHealth*, 6(3), e55. <https://doi.org/10.2196/mhealth.8261>
- König, L. M., Van Emmenis, M., Nurmi, J., Kassavou, A., & Sutton, S. (2022b). 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. (2020). *lmerTest: Tests in Linear Mixed Effects Models* (Version 3.1-3) [R package]. <https://cran.r-project.org/package=lmerTest>
- Lally, P., van Jaarsveld, C. H. M., Potts, H. W. W., & Wardle, J. (2010). How are habits formed: Modelling habit formation in the real world. *European Journal of Social Psychology*, 40(6), 998–1009. <https://doi.org/10.1002/ejsp.674>
- Lane, M. M., Davis, J. A., Beattie, S., Gómez-Donoso, C., Loughman, A., O’Neil, A., Jacka, F., Berk, M., Page, R., Marx, W., & Rocks, T. (2021). Ultraprocessed food and chronic noncommunicable

- diseases: A systematic review and meta-analysis of 43 observational studies. *Obesity Reviews*, 22(3), e13146. <https://doi.org/10.1111/obr.13146>
- Leblanc, V., Bégin, C., Corneau, L., Dodin, S., & Lemieux, S. (2015). Gender differences in dietary intakes: what is the contribution of motivational variables? *Journal of Human Nutrition and Dietetics*, 28(1), 37–46. <https://doi.org/10.1111/jhn.12213>
- Lichtensteiner, M. (2025). *Messreaktivität beim Fettkonsum: Der Einfluss von Forscherbeobachtung und die moderierende Rolle von Geschlecht und Gewissenhaftigkeit* (Unveröffentlichte Masterarbeit). Universität Wien.
- Lombardo, M., Aulisa, G., Padua, E., Annino, G., Iellamo, F., Pratesi, A., Caprio, M., & Bellia, A. (2020). Gender differences in taste and foods habits. *Nutrition and Food Science*, 50(1), 229–239. <https://doi.org/10.1108/NFS-04-2019-0132>
- Lombardo, M., Feraco, A., Armani, A., Camajani, E., Gorini, S., Strollo, R., Padua, E., Caprio, M., & Bellia, A. (2024). Gender differences in body composition, dietary patterns, and physical activity: insights from a cross-sectional study. *Frontiers in nutrition*, 11, 1414217. <https://doi.org/10.3389/fnut.2024.1414217>
- Lucassen, D. A., Brouwer-Brolsma, E. M., Slotegraaf, A. I., Kok, E., & Feskens, E. J. M. (2022). Dietary ASSESSment (DIASS) Study: Design of an Evaluation Study to Assess Validity, Usability and Perceived Burden of an Innovative Dietary Assessment Methodology. *Nutrients*, 14(6), 1156. <https://doi.org/10.3390/nu14061156>
- 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). <https://doi.org/10.1123/jmpb.2023-0035>
- Manolis, A., Manolis, A., Manolis, T., & Melita, H. (2023). Features of a Balanced Healthy Diet With Cardiovascular and Other Benefits. *Current vascular pharmacology*. <https://doi.org/10.2174/1570161121666230327135916>

- Maugeri, A., & Barchitta, M. (2019). A systematic review of ecological momentary assessment of diet: Implications and perspectives for nutritional epidemiology. *Nutrients*, 11(11), 2696. <https://doi.org/10.3390/nu11112696>
- 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>
- Melaku, Y. A., Renzaho, A., Gill, T. K., Taylor, A. W., Dal Grande, E., de Courten, B., Gonzalez-Chica, D., Hyppönen, E., Shi, Z., Riley, M., Adams, R., & Kinfu, Y. (2019). Burden and trend of diet-related non-communicable diseases in Australia and comparison with 34 OECD countries, 1990–2015: Findings from the Global burden of disease study 2015. *European Journal of Nutrition*, 58(3), 1299–1313. <https://doi.org/10.1007/s00394-018-1656-7>
- Miles, L. M., Elbourne, D., Farmer, A., Gulliford, M., Locock, L., McCambridge, J., Sutton, S., & French, D. P. (2018). Bias due to MEasurement Reactions In Trials to improve health (MERIT): Protocol for research to develop MRC guidance. *Trials*, 19(1), 653–658. <https://doi.org/10.1186/s13063-018-3017-5>
- 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>
- Moon, J., & Koh, G. (2020). Clinical evidence and mechanisms of high-protein diet-induced weight loss. *Journal of obesity & metabolic syndrome*, 29(3), 166. <https://doi.org/10.7570/jomes20028>
- Morton, R. W., Murphy, K. T., McKellar, S. R., Schoenfeld, B. J., Henselmans, M., Helms, E., Aragon, A. A., Devries, M. C., Banfield, L., Krieger, J. W., & Phillips, S. M. (2017). A systematic review, meta-analysis and meta-regression of the effect of protein supplementation on resistance training-induced gains in muscle mass and strength in healthy adults. *British Journal of Sports Medicine*, 52(6), 376–384. <https://doi.org/10.1136/bjsports-2017-097608>

- Naska, A., Ligiou, A., & Ligiou, P. (2017). Dietary assessment methods in epidemiological research: Current state of the art and future prospects. *F1000Research*, 6, 926. <https://doi.org/10.12688/f1000research.10703.1>
- Noncommunicable diseases*. (2024). World Health Organization. <https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases>
- O'Neil, A., Jacka, F. N., Quirk, S. E., Cocker, F., Taylor, C. B., Oldenburg, B., & Berk, M. (2015). A shared framework for the common mental disorders and Non-Communicable Disease: Key considerations for disease prevention and control. *BMC Psychiatry*, 15(1), 15. <https://doi.org/10.1186/s12888-015-0394-0>
- Pala, D., Petrini, G., Bosoni, P., Larizza, C., Quaglini, S., & Lanzola, G. (2024). Smartphone applications for nutrition Support: A systematic review of the target outcomes and main functionalities. *International Journal of Medical Informatics*, 184, 105351. <https://doi.org/10.1016/j.ijmedinf.2024.105351>
- Patel, M. L., Wakayama, L. N., & Bennett, G. G. (2021). Self-Monitoring via Digital Health in Weight Loss Interventions: A Systematic Review Among Adults with Overweight or Obesity. *Obesity*, 29(3), 478–499. <https://doi.org/10.1002/oby.23088>
- Perski, O., Keller, J., Kale, D., Asare, B. Y.-A., Schneider, V., Powell, D., Naughton, F., ten Hoor, G., Verboon, P., & Kwasnicka, D. (2022). Understanding health behaviours in context: A systematic review and meta-analysis of ecological momentary assessment studies of five key health behaviours. *Health Psychology Review*, 16(4), 576–601. <https://doi.org/10.1080/17437199.2022.2112258>
- Pinho, M. G. M., Mackenbach, J. D., Charreire, H., Oppert, J. M., Bárdos, H., Glonti, K., Rutter, H., Compernelle, S., De Bourdeaudhuij, I., Beulens, J. W. J., Brug, J., & Lakerveld, J. (2018). Exploring the relationship between perceived barriers to healthy eating and dietary behaviours in European adults. *European Journal of Nutrition*, 57(5), 1761–1770. <https://doi.org/10.1007/s00394-017-1458-3>

- Plasek, B., Lakner, Z., & Temesi, Á. (2020). Factors that Influence the Perceived Healthiness of Food—Review. *Nutrients*, 12(6), 1881. <https://doi.org/10.3390/nu12061881>
- Protein. (2017). Deutsche Gesellschaft für Ernährung. <https://www.dge.de/wissenschaft/referenzwerte/protein>
- Raber, M., Liao, Y., Rara, A., Schembre, S. M., Krause, K. J., Strong, L., Daniel-MacDougall, C., & Basen-Engquist, K. (2021). A systematic review of the use of dietary self-monitoring in behavioural weight loss interventions: delivery, intensity and effectiveness. *Public Health Nutrition*, 24(17), 5885–5913. <https://doi.org/10.1017/S136898002100358X>
- Ram, N., Brinberg, M., Pincus, A. L., & Conroy, D. E. (2017). The Questionable Ecological Validity of Ecological Momentary Assessment: Considerations for Design and Analysis. *Research in Human Development*, 14(3), 253–270. <https://doi.org/10.1080/15427609.2017.1340052>
- 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>
- Rodrigues, A. M., O'Brien, N., French, D. P., Glidewell, L., & Sniehotta, F. F. (2015). The question-behavior effect: Genuine effect or spurious phenomenon? A systematic review of randomized controlled trials with meta-analyses. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, 34(1), 61–78. <https://doi.org/10.1037/hea0000104>
- Roethlisberger, F. J., & Dickson, W. J. (1939). *Management and the worker*. Harvard University Press.
- Rramani, Q., Barakat, Y., Jacob, G., Ohla, K., Lim, S. X. L., Schicker, D., Freiherr, J., Saruco, E., Pleger, B., Weber, B., & Schultz, J. (2023). Nutrition claims influence expectations about food attributes, attenuate activity in reward-associated brain regions during tasting, but do not impact pleasantness. *Brain and behavior*, 13(1), e2828. <https://doi.org/10.1002/brb3.2828>

- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147–177. <https://doi.org/10.1037/1082-989X.7.2.147>
- Scientific Opinion on Dietary Reference Values for protein. (2012). *EFSA Journal*, 10(2), 2557. <https://doi.org/10.2903/j.efsa.2012.2557>
- Shaw, M., Rights, J., Sterba, S., & Flake, J. (2025). *r2mlm: R-Squared Measures for Multilevel Models* (Version 0.3.8) [R package]. <https://cran.r-project.org/package=r2mlm>
- Shiffman, S. (2009). Ecological momentary assessment (EMA) in studies of substance use. *Psychological Assessment*, 21(4), 486–497. <https://doi.org/10.1037/a0017074>
- Soenen, S., & Westerterp-Plantenga, M. S. (2008). Proteins and satiety: implications for weight management. *Current Opinion in Clinical Nutrition and Metabolic Care*, 11(6), 747–751. <https://doi.org/10.1097/mco.0b013e328311a8c4>
- Sofi, S. A., Qureshi, F. H., Khawaja, S., Tilwani, S. A., Majeed, H., & Hassan, A. (2025). Examining the Impact of Macronutrient Composition on Caloric Value: A Structural Equation Modeling (SEM) Approach. *Journal of Information Systems Engineering and Management*, 10(51s), 482–498. <https://doi.org/10.52783/jisem.v10i51s.10420>
- Song, J., Brown, M. K., Tan, M., MacGregor, G. A., Webster, J., Campbell, N. R. C., Trieu, K., Ni Mhurchu, C., Cobb, L. K., & He, F. J. (2021). Impact of color-coded and warning nutrition labelling schemes: A systematic review and network meta-analysis. *PLOS Medicine*, 18(10), e1003765. <https://doi.org/10.1371/journal.pmed.1003765>
- Steffen, A. D., Wilkens, L. R., Yonemori, K. M., Albright, C. L., & Murphy, S. P. (2021). A dietary supplement frequency questionnaire correctly ranks nutrient intakes in US older adults when compared to a comprehensive dietary supplement inventory. *The Journal of Nutrition*, 151(8), 2486–2495. <https://doi.org/10.1093/jn/nxab140>
- Stokes, T., Hector, A. J., Morton, R. W., McGlory, C., & Phillips, S. M. (2018). Recent Perspectives Regarding the Role of Dietary Protein for the Promotion of Muscle Hypertrophy with Resistance Exercise Training. *Nutrients*, 10(2), 180. <https://doi.org/10.3390/nu10020180>

- Strath, S., Kaminsky, L., Ainsworth, B., Ekelund, U., Freedson, P., Gary, R., Richardson, C., Smith, D., & Swartz, A., American Heart Association Physical Activity Committee of the Council on Lifestyle and Cardiometabolic Health and Cardiovascular, exercise, Cardiac Rehabilitation and Prevention Committee of the Council on Clinical Cardiology, and Council. (2013). Guide to the assessment of physical activity: Clinical and research applications: A scientific statement from the American Heart Association. *Circulation*, 128(20), 2259–2279. <https://doi.org/10.1161/01.cir.0000435708.67487.da>
- Stubbs, R. J., O'Reilly, L. M., Whybrow, S., Fuller, Z., Johnstone, A. M., Livingstone, M. B. E., Ritz, P., & Horgan, G. W. (2014). Measuring the difference between actual and reported food intakes in the context of energy balance under laboratory conditions. *The British Journal of Nutrition*, 111(11), 2032–2043. <https://doi.org/10.1017/S0007114514000154>
- Sun, J., Buys, N., & Hills, A. (2014). Dietary pattern and its association with the prevalence of obesity, hypertension and other cardiovascular risk factors among chinese older adults. *International Journal of Environmental Research and Public Health*, 11(4), 3956–3971. <https://doi.org/10.3390/ijerph110403956>
- Thompson, F. E., & Subar, A. F. (2017). Dietary assessment methodology. *Nutrition in the Prevention and Treatment of Disease*, 5–48. <https://doi.org/10.1016/B978-0-12-802928-2.00001-1>
- Tierney, N., Cook, D., McBain, M., & Fay, C. (2024). *naniar: Data Structures, Summaries, and Visualisations for Missing Data* (Version 1.1.0) [R package]. <https://cran.r-project.org/package=naniar>
- 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>
- Verplanken, B., & Orbell, S. (2022). Attitudes, habits, and behavior change. *Annual Review of Psychology*, 73, 327–352. <https://doi.org/10.1146/annurev-psych-020821-011744>

- Villinger, K., Wahl, D. R., Boeing, H., Schupp, H. T., & Renner, B. (2019). The effectiveness of app-based mobile interventions on nutrition behaviours and nutrition-related health outcomes: A systematic review and meta-analysis. *Obesity Reviews*, 20(10), 1465–1484. <https://doi.org/10.1111/obr.12903>
- Wahl, D. R., Villinger, K., Blumenschein, M., König, L. M., Ziesemer, K., Sproesser, G., Schupp, H. T., & Renner, B. (2020). Why We Eat What We Eat: Assessing Dispositional and In-the-Moment Eating Motives by Using Ecological Momentary Assessment. *JMIR mHealth and uHealth*, 8(1), e13191. <https://doi.org/10.2196/13191>
- 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>
- Wood, W., & Neal, D. T. (2007). A new look at habits and the habit-goal interface. *Psychological Review*, 114(4), 843–863. <https://doi.org/10.1037/0033-295X.114.4.843>
- Wu, G. (2016). Dietary protein intake and human health. *Food & Function*, 7(3), 1251–1265. <https://doi.org/10.1039/C5FO01530H>
- Zhao, S., Zhang, H., Xu, Y., Li, J., Du, S., & Ning, Z. (2024). The effect of protein intake on athletic performance: A systematic review and meta-analysis. *Frontiers in Nutrition*, 11. <https://doi.org/10.3389/fnut.2024.1455728>