

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Messreaktivität beim Fettkonsum: Einfluss der Forscherbeobachtung und moderierende Rolle von Geschlecht und Gewissenhaftigkeit

verfasst von | submitted by

Marion Lichtenberger BSc

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

Wien | Vienna, 2026

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

UA 066 840

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

Masterstudium Psychologie

Betreut von | Supervisor:

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

Zusammenfassung

Messreaktivität beschreibt Veränderungen von Kognitionen, Emotionen oder Verhalten, die allein durch die Erhebung selbst ausgelöst werden und dadurch die Validität von Forschungsergebnissen beeinträchtigen können. Die Datenlage zu reaktiven Effekten bei der digitalen Ernährungserfassung ist bislang begrenzt. Ziel der vorliegenden Studie war es, zu prüfen, ob sich der tägliche Fettkonsum unter Forscherbeobachtung verändert und ob Geschlecht sowie Gewissenhaftigkeit diesen Zusammenhang moderieren.

Analysiert wurden Ernährungstracking-Daten von 118 Teilnehmenden, die ihre Nahrungsaufnahme mittels gängiger Apps dokumentierten und in einen Online-Fragebogen übertrugen. Die Studie folgte einem Beobachtungsdesign auf Basis einer Datenspende mit zwei Messzeitpunkten innerhalb derselben Person im Abstand von einer Woche. Der durchschnittliche tägliche Fettkonsum in der Woche vor der Induktion der Forscherbeobachtung wurde mit den Werten der ersten beiden Tage der darauffolgenden Woche verglichen. Zur Hypothesenprüfung wurden lineare Mehrebenenmodelle mit Random Intercepts auf Personenebene berechnet.

Die Ergebnisse zeigten keinen signifikanten Unterschied im Fettkonsum zwischen den beiden Erhebungsphasen. Auch die Interaktionseffekte von Geschlecht und Gewissenhaftigkeit erwiesen sich als nicht signifikant. Insgesamt liefern die Befunde keine Hinweise auf kurzfristige messreaktive Effekte im Kontext digitaler Ernährungserfassung. Weitere Forschung sollte insbesondere Kontextfaktoren sowie Personen ohne vorherige Tracking-Erfahrung berücksichtigen, um potenzielle Bedingungen reaktiver Effekte differenzierter zu untersuchen.

Abstract

Measurement reactivity refers to changes in cognitions, emotions, or behavior that are triggered solely by the act of measurement itself and may thereby threaten the validity of research findings. Empirical evidence on reactive effects in digital dietary assessment remains limited. The present study aimed to examine whether daily fat intake changes under researcher observation and whether gender and conscientiousness moderate this relationship.

Dietary tracking data from 118 participants who documented their food intake using commonly used apps and transferred the data into an online questionnaire were analyzed. The study employed an observational design based on data donation with two measurement points within the same individuals one week apart. Average daily fat intake during the week prior to the induction of researcher observation was compared to intake during the first two days of the following week. Linear multilevel models with random intercepts at the participant level were used to test the hypotheses.

The results indicated no significant difference in fat intake between the two assessment periods. Interaction effects of gender and conscientiousness were also not significant. Overall, the findings provide no evidence for short-term measurement reactivity in the context of digital dietary assessment. Future research should particularly consider contextual factors and include individuals without prior tracking experience to further examine the conditions under which reactive effects may occur.

Inhaltsverzeichnis

Einleitung	1
Methoden der Ernährungserfassung.....	2
Digitales Ernährungstracking	2
Messreaktivität	3
Messreaktivität im digitalen Ernährungstracking	4
Historische Einordnung der Messreaktivität	5
Erklärungsansätze der Messreaktivität	6
Moderatoren von Messreaktivität.....	8
Big-5-Persönlichkeitsfaktor Gewissenhaftigkeit	8
Geschlechterdifferenzen in Messreaktivität	9
Fettkonsum im Kontext digitaler Ernährungserfassung.....	11
Gesundheitliche Bedeutung des Fettkonsums	11
Epidemiologische Situation des Fettkonsums	12
Studienziel und Forschungsfragen.....	13
Forschungsfrage 1:	13
Forschungsfrage 2:	13
Forschungsfrage 3:	14
Methodik	15
Stichprobenplanung und Rekrutierung	15
Design und Ablauf	16
Maße und Materialien	17
Datenaufbereitung und Qualitätsprüfung	19
Modellprüfung.....	21
Deskriptivstatistische Auswertung	21
Inferenzstatistische Auswertung	22

Ergebnisse.....	24
Stichprobe	24
Ergebnisdarstellung	25
Der tägliche Fettkonsum ist in den ersten beiden Tagen unter Forscherbeobachtung signifikant kleiner als in der Woche ohne Forscherbeobachtung.....	25
Frauen zeigen eine stärkere Veränderung des täglichen Fettkonsums in den ersten beiden Tagen unter Forscherbeobachtung im Vergleich zur Woche ohne Forscherbeobachtung als Männer.....	26
Es wird untersucht, ob und auf welche Weise Gewissenhaftigkeit den Effekt der Forscherbeobachtung auf den täglichen Fettkonsum moderiert.	27
Diskussion	29
Theoretische Einordnung und Interpretation der Befunde	29
Effekt der Forscherbeobachtung auf den Fettkonsum (RQ1)	29
Geschlecht als Moderator der Messreaktivität (RQ2).....	33
Gewissenhaftigkeit als Moderator der Messreaktivität (RQ3)	34
Stärken	36
Limitationen	37
Implikationen	40
Conclusio.....	41
Literaturverzeichnis.....	42
Tabellenverzeichnis	46

Einleitung

Das Ernährungsverhalten stellt einen zentralen Einflussfaktor für die Gesundheit dar. Insbesondere ungünstige Ernährungsweisen zählen weltweit zu den führenden Risikofaktoren für Morbidität und Mortalität und stehen in engem Zusammenhang mit der Entstehung und dem Verlauf zahlreicher chronischer Erkrankungen, darunter kardiovaskuläre Erkrankungen, Diabetes mellitus sowie verschiedene Krebs- und neurodegenerative Erkrankungen (Afshin et al., 2019; Gunning et al., 2025). Diese Befunde verdeutlichen die Relevanz der Stärkung gesundheitsförderlichen Ernährungsverhaltens innerhalb der Gesundheitspsychologie. Untersucht werden dabei insbesondere die Wechselwirkungen zwischen psychologischen Einflussfaktoren, spezifischen Interventionskomponenten und dem Ernährungsverhalten mit dem Ziel, wirksame Strategien zur Verhaltensänderung zu entwickeln und zu optimieren (Michie et al., 2009; Sanderman & Morgan, 2025).

Die Entwicklung und Evaluation gesundheitspsychologischer Interventionen setzt jedoch valide empirische Erkenntnisse über das tatsächliche Ernährungsverhalten voraus (Naska et al., 2017). Diesbezüglich hat sich das Smartphone-basierte Ernährungstracking als eine zunehmend verbreitete Methode zur Erfassung der Nahrungsaufnahme etabliert (König et al., 2022b). Dabei dokumentieren Teilnehmende ihre Nahrungsaufnahme zeitnah zum Konsum direkt in einer App, wodurch im Vergleich zu retrospektiven Erhebungsverfahren potenziell genauere und weniger verzerrte Daten generiert werden können.

Trotz dieser methodischen Vorteile bleibt jedoch die Problematik bestehen, dass die Studienteilnahme selbst das Verhalten beeinflussen kann (French & Sutton, 2010; König et al., 2022a; Robinson et al., 2015). Dieses Phänomen wird als Messreaktivität bezeichnet und beschreibt Veränderungen im Verhalten oder Erleben von Teilnehmenden, die durch den Messprozess selbst ausgelöst werden (French & Sutton, 2010). Eine differenzierte Untersuchung potenzieller Messreaktivitätseffekte ist erforderlich, um die Validität der erhobenen Daten angemessen beurteilen zu können. Ziel der vorliegenden Studie ist es daher, den Einfluss von Messreaktivität im Forschungskontext der digitalen Ernährungserfassung zu untersuchen.

Methoden der Ernährungserfassung

Zur Erfassung der individuellen Nahrungsaufnahme stehen in der Ernährungsforschung verschiedene methodische Ansätze zur Verfügung (Thompson & Subar, 2017). Erinnerungsbasierte Verfahren wie der Food Frequency Questionnaire (FFQ), der die übliche Verzehrhäufigkeit bestimmter Lebensmittel über längere Zeiträume hinweg erfasst, sowie der 24-hour-Recall, der auf einer retrospektiven Rekonstruktion der am Vortag konsumierten Lebensmittel basiert, sind zwar ökonomisch, jedoch anfällig für Erinnerungsverzerrungen und damit für Messfehler (Thompson & Subar, 2017). Neben diesen Verfahren existieren weitere Erhebungsmethoden wie Wiegeprotokolle, Diet-History-Interviews oder biomarkerbasierte Ansätze, die jeweils spezifische Limitationen hinsichtlich Aufwand, Erinnerungsabhängigkeit oder ökologischer Validität aufweisen (Thompson & Subar, 2017).

Eine zentrale Methode zur Reduktion erinnerungsbedingter Verzerrungen ist der Dietary-Record-Ansatz (Thompson & Subar, 2017). Dabei handelt es sich um ein Selbstprotokollverfahren, bei dem Teilnehmende alle konsumierten Lebensmittel und Getränke einschließlich der jeweiligen Mengen über einen oder mehrere Tage dokumentieren. Die Erhebung erfolgt idealerweise unmittelbar vor oder nach dem Verzehr, wodurch Erinnerungsfehler reduziert und eine differenzierte Quantifizierung der aufgenommenen Lebensmittel ermöglicht wird. Portionsgrößen können gewogen oder mithilfe standardisierter Schätzverfahren erfasst werden. Aufgrund der zeitnahen Dokumentation gilt dieses Verfahren als vergleichsweise präzise, dennoch ist zu berücksichtigen, dass das aktive Protokollieren selbst das Ernährungsverhalten beeinflussen kann. Diese potenzielle Reaktivität ist eine methodische Besonderheit des Dietary-Record-Ansatzes (Thompson & Subar, 2017).

Digitales Ernährungstracking

Digitale Ernährungs-Apps basieren konzeptionell auf dem Dietary-Record-Ansatz und übertragen dessen Prinzipien in ein digitales Format. Über die Vorteile klassischer Ernährungstagebücher hinaus reduzieren sie den Erhebungsaufwand und ermöglichen durch unterstützende Funktionen eine präzisere Erfassung der Nährstoffaufnahme (Rollo et al., 2016). Die Abhängigkeit vom retrospektiven Erinnerungsvermögen wird durch die Möglichkeit der zeitnahen und alltagsintegrierten Dokumentation der Nahrungsaufnahme reduziert (Thompson & Subar, 2017). Durch automatisierte Nährwertberechnungen und standardisierte Eingabemasken, bei denen zunehmend auch KI-gestützte Funktionen zur Unterstützung der

Lebensmittelerkennung zum Einsatz kommen, stellen sie zudem eine praktikable und skalierbare Form der Datenerhebung dar, die sowohl in der Forschung als auch im Alltag zunehmend Anwendung findet (König et al., 2022b).

Messreaktivität

Die vorliegende Masterarbeit befasst sich mit dem Phänomen der Messreaktivität im Kontext der digitalen Erfassung von Ernährungsverhalten. Messreaktivität beschreibt den Effekt, dass kognitive, emotionale und behaviorale Prozesse allein durch die Teilnahme an einer Studie beeinflusst werden können. Die Untersuchung der Ursachen und Auswirkungen von Messreaktivität ist besonders relevant, da sie die Validität von Forschungsdaten beeinträchtigen kann, wenn Studien diesen Effekt nicht systematisch kontrollieren (French & Sutton, 2010). Die aus Studienergebnissen abgeleiteten Schlussfolgerungen können demnach durch Verzerrungen beeinträchtigt werden.

Messreaktivität ist jedoch nicht ausschließlich als methodisches Problem zu verstehen, sondern kann auch gezielt genutzt werden. Im Rahmen des Question-Behavior-Effects zeigt sich, dass bereits die wiederholte Befragung zu gesundheitsrelevanten Verhalten dieses verändern kann, etwa hin zu einem gesünderen Ernährungsverhalten (Miles et al., 2020).

Die Forschung zur Messreaktivität im Gesundheitsbereich konzentriert sich bislang überwiegend auf das Bewegungsverhalten, während Ernährungsaspekte vergleichsweise selten systematisch untersucht wurden (French et al., 2021; König et al., 2022a; Miles et al., 2020). Ein Großteil der empirischen Studien zeigt, dass bereits die Erfassung gesundheitsbezogener Verhaltensweisen, etwa mittels Schrittzählern, Wearables oder Selbstberichtsinstrumenten, mit Verhaltensveränderungen einhergehen kann (Bravata et al., 2007; Michie et al., 2009). Im Gegensatz dazu liegt für den Bereich der Ernährung bislang deutlich weniger empirische Evidenz vor, obwohl Ernährungsverhalten eine ebenso zentrale Rolle für die Gesundheit spielt (König et al., 2022a; Rollo et al., 2016).

Das ist auch deshalb relevant, da sich das Ernährungsverhalten in zentralen Merkmalen vom Bewegungsverhalten unterscheidet. So ist es stärker in alltägliche Routinen integriert, unterliegt vielfältigen situativen und sozialen Einflüssen und ist häufig weniger bewusst zugänglich (König et al., 2022b; Thompson & Subar, 2017). Zudem ist die Erfassung der Nahrungsaufnahme mit spezifischen methodischen Herausforderungen verbunden, etwa im

Hinblick auf Portionsschätzung oder die Komplexität von Mahlzeiten (Thompson & Subar, 2017; Illner et al., 2012). Diese Besonderheiten legen nahe, dass sich Mechanismen und Ausmaß von Messreaktivität im Ernährungsbereich potenziell von jenen im Bewegungsverhalten unterscheiden können (König et al., 2022b). Deshalb erscheint eine gezielte Untersuchung von Messreaktivität im Kontext des Ernährungsverhaltens notwendig, um bestehende Forschungslücken zu schließen und die Übertragbarkeit bisheriger Befunde kritisch zu prüfen (Miles et al., 2020; French et al., 2021).

Messreaktivität im digitalen Ernährungstracking

Im Bereich der Ernährung ist Messreaktivität von besonderer Relevanz, da digitale Ernährungstracking-Apps zunehmend sowohl in der Forschung als auch im Alltag eingesetzt werden, um Ernährungsgewohnheiten möglichst genau abzubilden. Wiederholte Selbstbeobachtung kann dazu bei gesundheitsrelevantem Verhalten kognitive und motivationale Prozesse aktivieren, etwa eine erhöhte Verhaltenssalienz, Zielaktivierung oder Selbstregulation, die kurzfristige Verhaltensanpassungen begünstigen können (French & Sutton, 2010). French und Sutton (2010) sowie Miles et al. (2020) argumentieren, dass reaktive Effekte speziell dann verstärkt auftreten können, wenn das erfasste Verhalten für die betreffende Person eine hohe subjektive Relevanz besitzt, mit bestehenden Veränderungsintentionen verknüpft ist oder normativ stark aufgeladen ist, etwa bei Themen wie Gesundheit oder Körpergewicht.

Empirische Untersuchungen zur Messreaktivität im Verhaltenskontext weisen darauf hin, dass sich Verhaltensmuster im Verlauf einer Studie systematisch verändern können (Arigo & König, 2024; König et al., 2022a). Häufig zeigt sich zu Beginn einer Erhebung eine verstärkte Ausprägung sozial erwünschter oder gesundheitsförderlicher Verhaltensweisen, die im weiteren Verlauf wieder abnimmt. Umgekehrt können weniger erwünschte Verhaltensweisen zu Beginn der Untersuchung reduziert auftreten und später wieder zunehmen. Arigo und König (2024) berichten zudem, dass Messreaktivität vor allem in den ersten ein bis zwei Tagen nach Induktion der Forscherbeobachtung am stärksten ausgeprägt sein kann. Übertragen auf das Ernährungsverhalten impliziert dieses Muster, dass gesundheitsförderliche Praktiken, wie beispielsweise eine Reduktion des Fettkonsums, zu Studienbeginn stärker ausgeprägt sein könnten, während im Zeitverlauf eine Rückkehr zu habitualisierten Verhaltensmustern erfolgt. Zudem weisen meta-analytische Befunde darauf hin, dass das Ausgangsniveau des Verhaltens

das Ausmaß reaktiver Effekte beeinflussen kann (Miles et al., 2020), wobei Personen mit geringer Ausprägung des Verhaltens oder problematischen Verhaltensmustern tendenziell stärkere kurzfristige Anpassungen zeigen.

Digitale Ernährungs-Apps nehmen in empirischen Untersuchungen eine besondere Stellung ein. Sie ermöglichen eine ökologisch valide, alltagsnahe und differenzierte Erfassung des Ernährungsverhaltens, da die Datenerhebung direkt im natürlichen Umfeld der Nutzenden erfolgt und auf tatsächlich verwendeten Anwendungen basiert (König et al., 2022b; Thompson & Subar, 2017). Dadurch können Verhaltensdaten unter realistischen Bedingungen erfasst werden, was für die Untersuchung von Alltagsverhalten von zentraler Bedeutung ist. Zugleich unterscheiden sich digitale Tracking-Ansätze methodisch von klassischen Erhebungsverfahren, da sie häufig auf bereits etablierte Selbstbeobachtungsprozesse zurückgreifen und nicht zwingend die Einführung eines neuen Messinstruments erfordern. Dies ermöglicht es, reaktive Effekte, die auf die Einführung eines Messinstruments zurückgehen, von Effekten der Forscherbeobachtung zu trennen und differenzierter zu analysieren (König et al., 2025).

Gleichzeitig werden digitale Ernährungs-Apps zunehmend nicht nur als Erhebungsinstrument, sondern auch als Mittel zur Verhaltensregulation und Prävention eingesetzt und weisen damit ein interventionsähnliches Potenzial auf (König et al., 2022b). Eine differenzierte Analyse potenzieller reaktiver Effekte im Kontext digitaler Ernährungserfassung erscheint daher erforderlich. Eine solche Untersuchung ermöglicht es, Verhaltensänderungen, die im Kontext der Messung auftreten, systematisch zu identifizieren und deren Einfluss auf die Aussagekraft der erhobenen Daten angemessen zu berücksichtigen. Zudem können Bedingungen untersucht werden, unter denen die Erhebung selbst mit Verhaltensänderungen einhergeht.

Historische Einordnung der Messreaktivität

Die wissenschaftliche Auseinandersetzung mit reaktiven Effekten im Forschungskontext lässt sich historisch auf die sogenannten Hawthorne-Studien in den Jahren 1924 bis 1932 zurückführen (Mayo, 2003). Im Rahmen dieser Untersuchungen wurde beobachtet, dass sich das Verhalten von Beschäftigten in den Hawthorne-Werken der Western Electric Company in Chicago (USA) allein durch das Wissen um ihre Teilnahme an einer Studie veränderte, eine beobachtete Verhaltensanpassung, die später als Hawthorne-Effekt bezeichnet wurde (French

et al., 2021). Der Hawthorne-Effekt beschreibt demnach Verhaltensänderungen, die auf die Wahrnehmung von Beobachtung zurückgeführt werden.

In der aktuellen Forschung wird er weniger als klar umrissenes Einzelphänomen verstanden, sondern vielmehr als historischer Ausgangspunkt für die systematische Untersuchung von Reaktivität im wissenschaftlichen Kontext. Er gilt als frühe konzeptionelle Vorform reaktiver Effekte, zu denen auch die Messreaktivität zählt (French et al., 2021). Obwohl die empirische Evidenz des klassischen Hawthorne-Effekts in der neueren Literatur differenziert bewertet wird (McCambridge et al., 2014), stellt er weiterhin einen zentralen Referenzpunkt für Theorien dar, die den Einfluss von Beobachtungs- und Messsituationen auf menschliches Verhalten erklären.

Erklärungsansätze der Messreaktivität

Zur Erklärung des Phänomens der Messreaktivität wurden in der bisherigen Forschung verschiedene theoretische und empirische Ansätze diskutiert. Ein zentraler Erklärungsansatz bezieht sich auf die erhöhte kognitive Zugänglichkeit bestimmter Einstellungen, Intentionen oder Normen infolge der Messung selbst. Das Beantworten von Fragen kann verhaltensrelevante Kognitionen, wie Einstellungen oder Intentionen, kognitiv zugänglicher machen (French & Sutton, 2010; Long et al., 2025). Übertragen auf den Kontext der vorliegenden Studie bedeutet dies, dass die wiederholte Erfassung des Ernährungsverhaltens dazu führen kann, dass gesundheitsbezogene Einstellungen oder Ernährungsziele stärker in den Fokus der Aufmerksamkeit rücken. Durch diese erhöhte Salienz gewinnen die aktivierten Kognitionen an handlungsleitender Bedeutung, was in der Folge das Auftreten entsprechender Verhaltensweisen begünstigen kann. In diesem Zusammenhang wird häufig vom sogenannten mere measurement effect gesprochen, bei dem das bloße Erfragen eines Verhaltens dessen Auftreten wahrscheinlicher machen kann (Long et al., 2025).

Ein weiterer bedeutender Erklärungsansatz betrifft Selbstmonitoring-Prozesse (Barta et al., 2012; Bravata et al., 2007; Zhang et al., 2023). Messinstrumente, die eine fortlaufende Selbstbeobachtung erfordern, wie etwa Ernährungstagebücher oder digitale Tracking-Anwendungen, können selbst als unbeabsichtigte Interventionen wirken (Barta et al., 2012). Bei der digitalen Ernährungserfassung kann demnach das wiederholte Dokumentieren der eigenen Nahrungsaufnahme mit Reflexions- und Selbstregulationsprozesse einhergehen (Zhang et al., 2023). Besonders im Kontext digitaler Selbsttracking-Methoden ist dieser

Mechanismus von hoher Relevanz, da diese Anwendungen explizit auf regelmäßige Selbstbeobachtung ausgelegt sind und dadurch kontinuierliche Rückmeldeschleifen erzeugen können (Bravata et al., 2007; Zhang et al., 2023).

Ein weiterer Ansatz zur Erklärung von Messreaktivität besteht in der Generierung neuer Kognitionen durch die Messung selbst (French et al., 2021). Das Ausfüllen von Fragebögen zum Ernährungsverhalten kann dazu führen, dass Personen erstmals explizit über bestimmte Aspekte ihres Verhaltens oder ihrer Einstellungen nachdenken. Die wiederholte Auseinandersetzung mit dem eigenen Ernährungsverhalten kann dabei neue Überzeugungen oder Bewertungen hervorbringen, die vor der Messung nicht oder nur implizit vorhanden waren (French et al., 2021). In einer Think-aloud-Studie konnten Darker und French (2009) zeigen, dass Teilnehmende während des Beantwortens von Fragebogenitems neue Gedanken und Überzeugungen konstruieren, was potenziell nachfolgendes Verhalten beeinflussen kann.

Neben kognitiven Prozessen werden auch emotionale Reaktionen als Erklärung für Messreaktivität diskutiert (French et al., 2021; French & Sutton, 2010). So konnte anhand einer Studie zum Beck Depression Inventory gezeigt werden, dass das Ausfüllen gesundheitsbezogener Fragebögen emotionale Reaktivität hervorrufen kann (Sharpe & Gilbert, 1998). Insbesondere Messungen, die sensible oder gesundheitsrelevante Themen betreffen, können kurzfristig emotionale Reaktionen wie Angst oder Besorgnis auslösen, welche das Antwortverhalten und gegebenenfalls auch das tatsächliche Verhalten beeinflussen können (French et al., 2021).

Schließlich wird Messung auch als implizite Handlungsaufforderung (Prompting) verstanden (French et al., 2021). Der Kontakt im Rahmen einer Studie oder das wiederholte Erinnern an ein bestimmtes Verhalten kann als Aufforderung interpretiert werden, dieses Verhalten auszuführen oder zu verändern (Edwards et al., 2002; French et al., 2021). In diesem Sinne fungiert die Messung selbst als situativer Hinweisreiz, der Verhaltensänderungen begünstigen kann (French et al., 2021). In der vorliegenden Studie könnte demnach die Ankündigung einer erneuten Datenerhebung eine Woche später dazu beitragen, dass Teilnehmende ihr Ernährungsverhalten bewusster wahrnehmen und gegebenenfalls anpassen.

Insgesamt deuten diese Erklärungsansätze darauf hin, dass Messreaktivität ein multifaktorielles Phänomen darstellt, das sowohl kognitive, emotionale als auch

selbstregulatorische Prozesse umfasst. Gleichzeitig betonen French et al. (2021), dass die zugrunde liegenden Mechanismen bislang nur unzureichend empirisch untersucht wurden, was die Notwendigkeit weiterer Forschung in diesem Bereich unterstreicht. Zudem weisen systematische Übersichtsarbeiten und Meta-Analysen auf eine ausgeprägte Heterogenität der berichteten Reaktivitätseffekte hin (French & Sutton, 2010; French et al. 2021; König et al., 2022a; Miles et al., 2020). Ausnahmen hierbei sind konsistente Effekte bezüglich des Question Behavior Effekts, sowie Messreaktivität beim Verwenden von Schrittzählern (French et al. 2021).

Moderatoren von Messreaktivität

Big-5-Persönlichkeitsfaktor Gewissenhaftigkeit

Ein möglicher individueller Einflussfaktor auf das Auftreten von Messreaktivität ist die Persönlichkeit der untersuchten Personen. Um Differenzen zwischen Individuen zu beschreiben, dominiert in der Persönlichkeitsforschung das Big-Five-Modell nach Costa und McCrae (1992). In der deutschen Literatur sind die Big Five unter den Bezeichnungen Neurotizismus oder Emotionale Stabilität, Extraversion, Offenheit oder Intellekt, Verträglichkeit und Gewissenhaftigkeit bekannt. Die Big Five beschreiben dabei relativ stabile interindividuelle Unterschiede, die sich in einer Vielzahl gesundheitsbezogener Verhaltensweisen widerspiegeln.

Der Big-Five-Persönlichkeitsfaktor Gewissenhaftigkeit umfasst dabei Facetten wie Ordnungsliebe, Selbstkontrolle, Verantwortungsbewusstsein und Fleiß (Roberts et al., 2005). In der bisherigen Forschung konnte wiederholt gezeigt werden, dass Gewissenhaftigkeit insbesondere mit gesundheitsförderlichem Verhalten in Zusammenhang steht. Goldberg und Strycker (2002) berichten beispielsweise, dass Personen mit höheren Ausprägungen in Gewissenhaftigkeit seltener ungesunde Ernährungsgewohnheiten zeigen und unter anderem den Konsum fettreicher Nahrungsmittel eher vermeiden. Insgesamt sind höhere Werte in Gewissenhaftigkeit mit einem Ernährungsverhalten assoziiert, das stärker an gesundheitlichen Empfehlungen orientiert ist, während geringere Ausprägungen mit einem vermehrten Konsum ungesunder Nahrungsmittel einhergehen. Zur Erklärung dieses Zusammenhangs verweisen Goldberg und Strycker (2002) auf zentrale Merkmale der Gewissenhaftigkeit wie eine erhöhte Selbstkontrolle, eine stärkere Orientierung an Regeln und Normen sowie eine langfristige

Ausrichtung auf gesundheitsbezogene Ziele, die sich auch im alltäglichen Ernährungsverhalten niederschlagen können.

Diese Merkmale legen nahe, dass Gewissenhaftigkeit nicht nur direkt mit dem Ernährungsverhalten zusammenhängt, sondern auch die Reaktion auf Mess- und Studiensituationen beeinflussen könnte. Besonders in Kontexten, in denen das eigene Verhalten regelmäßig erfasst und reflektiert wird, wie etwa bei der digitalen Ernährungserfassung, könnten gewissenhaftere Personen ihr Verhalten stärker an gesundheitsförderlichen Normen ausrichten. Dies würde bedeuten, dass Messreaktivität bei Personen mit höherer Gewissenhaftigkeit stärker ausgeprägt sein könnte. Da die bisherige Studienlage zur moderierenden Rolle von Gewissenhaftigkeit im Zusammenhang mit Messreaktivität jedoch begrenzt ist, wird dieser Zusammenhang in der vorliegenden Arbeit explorativ untersucht.

Geschlechterdifferenzen in Messreaktivität

Auch das Geschlecht wurde als möglicher Moderator der Messreaktivität diskutiert, allerdings ist die Studienlage diesbezüglich nach wie vor uneinheitlich (König et al., 2025). Die Frage nach psychologischen Geschlechtsunterschieden gilt generell als kontrovers und erfordert weiterhin differenzierte empirische Untersuchungen um fundierte Aussagen treffen zu können (Weisberg et al., 2011). Ziel der Geschlechterforschung ist es, durchschnittliche Unterschiede in Verhaltensmustern zwischen Männern und Frauen zu identifizieren und zu erklären, wobei zugleich berücksichtigt wird, dass selbst bei ausgeprägten Mittelwertsunterschieden häufig eine erhebliche Überlappung der Verteilungen beider Gruppen besteht (Hyde, 2005).

Eine systematische Übersichtsarbeit von König et al. (2022a) untersuchte unter anderem mögliche Geschlechtsunterschiede in der Messreaktivität bei gesundheitsbezogenen Verhaltensweisen, wobei der Schwerpunkt der einbezogenen Studien überwiegend auf physischer Aktivität lag. In insgesamt sechs der in die Übersichtsarbeit eingeschlossenen Studien wurde das Geschlecht als potenzieller Moderator von Messreaktivität berücksichtigt. In fünf dieser sechs Studien zeigten sich keine signifikanten Unterschiede zwischen männlichen und weiblichen Teilnehmenden. Eine dieser Studien berichtete einen geschlechtsspezifischen Effekt, wobei eine erhöhte Messreaktivität bei weiblichen, jedoch nicht bei männlichen Personen, beobachtet wurde (Ho et al., 2013). Insgesamt zeigt sich erneut, dass

Messreaktivität bislang überwiegend im Kontext physischer Aktivität untersucht wurde, während empirische Befunde zu möglichen reaktiven Effekten im Bereich des Ernährungsverhaltens nur in begrenztem Umfang vorliegen. Dahingegen konnten die Ergebnisse von Stubbs et al. (2014) zeigen, dass allein das Wissen um die Erfassung der Nahrungsaufnahme zu einer Verhaltensänderung führen kann, die sich insbesondere bei Frauen in einer signifikanten Reduktion der Fettaufnahme äußerte, während ein entsprechender Effekt bei Männern ausblieb.

Frauen weisen außerdem tendenziell höhere Ausprägungen in Gewissenhaftigkeit auf als Männer (Vecchione et al., 2012). Da Gewissenhaftigkeit unter anderem mit einer stärkeren Orientierung an Regeln und sozialen Normen sowie mit ausgeprägteren Selbstregulationsprozessen verbunden ist, lässt sich argumentieren, dass Frauen in Mess- und Studiensituationen eher zu Anpassungen ihres Verhaltens neigen. Dies gilt insbesondere dann, wenn das erfasste Verhalten in engem Zusammenhang mit gesundheitlichen Normen steht, was sich in einer ausgeprägteren Messreaktivität äußern könnte.

Zusätzlich wird in der Forschung diskutiert, dass soziale Erwünschtheit einen relevanten Einfluss auf reaktive Effekte ausüben kann (Barta et al., 2012). Soziale Erwünschtheit beschreibt die Tendenz, eigenes Verhalten oder Antworten in Richtung sozial akzeptierter Normen zu verändern oder darzustellen. Empirische Befunde deuten darauf hin, dass Frauen stärker zu sozial erwünschtem Antwortverhalten tendieren als Männer (Bernardi & Guptill, 2008). In Kombination mit einem insgesamt höheren Gesundheitsbewusstsein, speziell im Hinblick auf die Vermeidung ungesunder Ernährungsweisen wie fettreicher Nahrungsmittel (Goldberg & Strycker, 2002), könnte dies dazu führen, dass Frauen ihr Ernährungsverhalten im Rahmen von Studien stärker an gesundheitlichen Normen ausrichten.

Zu berücksichtigen ist, dass soziale Erwünschtheit nicht ausschließlich durch stabile Personenmerkmale bestimmt wird, sondern in erheblichem Maße auch durch Eigenschaften der Erhebungssituation beeinflusst werden kann (Barta et al., 2012). Befunde aus der Tagebuchforschung zeigen, dass eine erhöhte wahrgenommene Anonymität soziale Erwünschtheitseffekte reduzieren kann, da die Sorge vor negativer Bewertung durch andere verringert wird (Barta et al., 2012). Unter solchen Bedingungen berichten Teilnehmende tendenziell weniger verzerrt und orientieren sich in geringerem Ausmaß an sozialen Normen. Für den Kontext der vorliegenden Studie ist daher anzunehmen, dass die anonyme

Datenerhebung dazu beigetragen hat, den Einfluss sozial erwünschten Antwortverhaltens zumindest teilweise zu reduzieren. Gleichzeitig ist jedoch nicht auszuschließen, dass entsprechende Tendenzen weiterhin wirksam bleiben, da soziale Erwünschtheit nicht ausschließlich situativ bedingt ist, sondern auch dispositionelle Komponenten umfasst.

Das Geschlecht wird als potenzieller Moderator von Messreaktivität berücksichtigt, da die bisherigen Befunde hierzu uneinheitlich sind (König et al., 2022a). Dadurch soll geprüft werden, ob sich Unterschiede der Effekte der Forscherbeobachtung zwischen Frauen und Männern zeigen und inwiefern sich der heterogenen Forschungsstand weiter differenzieren lässt.

Fettkonsum im Kontext digitaler Ernährungserfassung

Gesundheitliche Bedeutung des Fettkonsums

Ein zentrales gesundheitsrelevantes Verhalten im Ernährungsbereich stellt die Fettzufuhr dar. Nahrungsfette stehen seit Jahrzehnten im Fokus ernährungswissenschaftlicher und präventivmedizinischer Forschung, da sowohl die absolute Menge als auch die qualitative Zusammensetzung der aufgenommenen Fettsäuren mit der Entstehung zahlreicher chronischer Erkrankungen assoziiert sind. Die evidenzbasierte Leitlinie der Deutschen Gesellschaft für Ernährung (DGE, 2015) verweist darauf, dass insbesondere Adipositas, Typ-2-Diabetes, Dyslipoproteinämien, arterielle Hypertonie, koronare Herzkrankheit, Schlaganfall, das metabolische Syndrom sowie bestimmte onkologische Erkrankungen im Zusammenhang mit einer ungünstigen Fettzufuhr untersucht wurden. Neben der Gesamtfettmenge spielt dabei insbesondere der Anteil gesättigter Fettsäuren sowie das Verhältnis zwischen gesättigten und mehrfach ungesättigten Fettsäuren eine zentrale Rolle. Fett nimmt damit eine bedeutende Stellung in der primären Prävention ernährungsmitbedingter Erkrankungen ein.

Für den deutschsprachigen Raum werden gemeinsame Referenzwerte für die Nährstoffzufuhr von der Deutschen Gesellschaft für Ernährung (DGE), der Österreichischen Gesellschaft für Ernährung (ÖGE) und der Schweizerischen Gesellschaft für Ernährung (SGE) herausgegeben (D-A-CH-Referenzwerte). Für Jugendliche und Erwachsene wird ein Richtwert von 30 % der täglichen Gesamtenergiezufuhr aus Fett empfohlen (DGE, 2015). Diese Empfehlung berücksichtigt sowohl die hohe Energiedichte von Fetten als auch deren potenziellen Einfluss auf Körpergewicht, Lipidstoffwechsel und kardiovaskuläre Gesundheit.

Epidemiologische Situation des Fettkonsums

Für die empirische Erfassung des Ernährungsverhaltens ist neben der prozentualen Angabe eine Darstellung in absoluten Mengen erforderlich. Die Darstellung in Gramm pro Tag ist im Kontext der vorliegenden Studie von besonderer Bedeutung, da der tägliche Fettkonsum in absoluten Mengen erhoben wurde. Während prozentuale Energieanteile für bevölkerungsbezogene Empfehlungen geeignet sind, ermöglichen absolute Mengenangaben eine differenzierte Analyse individueller Verhaltensmuster sowie potenzieller intraindividuellere Veränderungen im Zeitverlauf.

Da Fett eine Energiedichte von 9 kcal pro Gramm aufweist, lässt sich die Empfehlung von 30 % der Gesamtenergiezufuhr in Gramm umrechnen. Bei einer Energiezufuhr von 1800 kcal entspricht dies etwa 60 g Fett pro Tag, bei 2000 kcal rund 67 g, bei 2200 kcal etwa 73 g und bei 2500 kcal etwa 83 g Fett pro Tag. Für die erwachsene Allgemeinbevölkerung ergibt sich damit in Abhängigkeit vom individuellen Energiebedarf ein Orientierungsbereich von ungefähr 60 bis 80 g Fett pro Tag.

Die tatsächliche Fettzufuhr in Österreich liegt jedoch über diesen Richtwerten. Der Österreichische Ernährungsbericht 2017 berichtet eine mittlere Energiezufuhr von 1815 kcal pro Tag bei Frauen und 2453 kcal pro Tag bei Männern (Bundesministerium für Gesundheit und Frauen, 2017). Gleichzeitig beträgt der durchschnittliche Fettanteil 36,8 % der Gesamtenergiezufuhr (Bundesministerium für Gesundheit und Frauen, 2017). Rechnerisch entspricht dies etwa 74 g Fett pro Tag bei Frauen und rund 100 g Fett pro Tag bei Männern und liegt somit über dem aus den D-A-CH-Referenzwerten abgeleiteten Orientierungsbereich. Zudem wird berichtet, dass insbesondere die Zufuhr gesättigter Fettsäuren im Mittel erhöht ist, während mehrfach ungesättigte Fettsäuren häufig unter den Empfehlungen liegen (Bundesministerium für Gesundheit und Frauen, 2017). Diese Befunde verdeutlichen, dass sowohl die Quantität als auch die Qualität der Fettzufuhr ernährungsphysiologisch relevant sind. Während die durchschnittliche Fettzufuhr bereits über den Referenzwerten liegt, zeigt eine Übersichtsarbeit von Fekete und Weyers (2016) zudem, dass Ernährungsverhalten sozial ungleich verteilt ist und bestimmte Bevölkerungsgruppen in besonderem Maße von ungünstigen Ernährungsmustern betroffen sind.

Vor dem Hintergrund der hohen epidemiologischen Bedeutung des Fettkonsums ist es sinnvoll, mögliche reaktive Effekte im Rahmen der Ernährungserfassung differenziert zu

untersuchen. Eine valide Einschätzung solcher Effekte ist nicht nur methodisch relevant, sondern auch für die Interpretation populationsbezogener Ernährungsdaten von zentraler Bedeutung.

Studienziel und Forschungsfragen

Ziel der vorliegenden Studie ist es, den kurzfristigen Effekt der Forscherbeobachtung auf den täglichen Fettkonsum in einem direkten Vergleich mit dem Ernährungsverhalten ohne Forscherbeobachtung zu untersuchen. Auf Grundlage der dargestellten theoretischen Überlegungen sowie der bisherigen empirischen Befundlage werden zwei konfirmatorische Forschungsfrage sowie eine explorative Forschungsfrage abgeleitet.

Forschungsfrage 1:

Im Zentrum der Untersuchung steht die Frage, ob und in welchem Ausmaß die Teilnahme an einer Studie zur digitalen Ernährungserfassung mit Veränderungen im alltäglichen Ernährungsverhalten einhergeht. Als zentrale Untersuchungsvariable dient der tägliche Fettkonsum. Daraus ergibt sich die erste Forschungsfrage sowie die zugehörige Hypothese:

RQ1: Zeigt sich eine intraindividuelle Veränderung des durchschnittlichen täglichen Fettkonsums zwischen der Baseline-Woche ohne Forscherbeobachtung und den ersten beiden Tagen unter Forscherbeobachtung?

H1: Der durchschnittliche tägliche Fettkonsum ist in den ersten beiden Tagen unter Forscherbeobachtung geringer als in der Baseline-Woche.

Forschungsfrage 2:

Darüber hinaus wird berücksichtigt, dass Messreaktivität nicht bei allen Personen in gleichem Maße auftritt. Neben dem Haupteffekt der Forscherbeobachtung werden daher auch interindividuelle Unterschiede in den Blick genommen. Eine potenzielle Moderatorvariable stellt das Geschlecht dar. Daraus ergibt sich die zweite Forschungsfrage sowie die zugehörige Hypothese:

RQ2: Moderiert das Geschlecht den Effekt der Forscherbeobachtung auf den Fettkonsum?

H2: Frauen zeigen eine stärkere Veränderung des täglichen Fettkonsums in den ersten beiden Tagen unter Forscherbeobachtung im Vergleich zur Woche ohne Forscherbeobachtung als Männer.

Forschungsfrage 3:

Zuletzt wird explorativ geprüft, ob die Gewissenhaftigkeit den Zusammenhang zwischen Messzeitpunkt und Fettkonsum moderiert. Aufgrund der bislang begrenzten empirischen Evidenz wird hierfür keine gerichtete Hypothese formuliert.

RQ3: Moderiert die Gewissenhaftigkeit den Effekt der Forscherbeobachtung auf den Fettkonsum?

Methodik

Die Datenerhebung erfolgte im Rahmen eines übergeordneten Forschungsprojekts am Arbeitsbereich Gesundheitspsychologie der Universität Wien, das sich mit reaktiven Effekten im Kontext digitaler Selbstvermessung befasst (König & Ruf, 2024). Die vorliegende Arbeit stellt eigenständige Fragestellungen innerhalb dieses Projekts dar und nutzt die Variablen und Erhebungszeitpunkte, die zur Prüfung der spezifischen Hypothesen zum Fettkonsum und dessen potenziellen Moderatoren erforderlich sind. Parallel zu dieser Arbeit werden weitere Abschlussarbeiten auf Basis desselben Studiendesigns durchgeführt, wobei unterschiedliche gesundheitsbezogene Variablen im Fokus stehen (Förg, 2026; Lackner, 2025).

Zur Sicherstellung methodischer Transparenz und zur Minimierung forschungspraktischer Verzerrungen wurde die Studie vor Zugriff auf die Daten präregistriert. Die finale Version der Präregistrierung wurde im November 2025 in der arbeitsbereichsinternen Studiendatenbank hinterlegt. Darin sind die Forschungsfragen, Hypothesen, Operationalisierungen, Ein- und Ausschlusskriterien sowie der geplante Analyseansatz detailliert spezifiziert. Sämtliche für diese Arbeit relevanten Materialien, einschließlich der eingesetzten Fragebögen, der Datensätze sowie der Auswertungsskripte, sind in der arbeitsbereichsinternen Datenbank archiviert. Dieses Vorgehen dient der Nachvollziehbarkeit der Analyseschritte und unterstützt eine potenzielle Replizierbarkeit der Ergebnisse. Etwaige Abweichungen von den in der Präregistrierung festgelegten Vorgehensweisen werden im jeweiligen Abschnitt explizit kenntlich gemacht und begründet, um die Transparenz des Forschungsprozesses zu wahren. Da die Datenerhebung im Kontext eines übergeordneten Forschungsprojekts erfolgte, umfasste der Fragebogen zusätzliche Variablen, die nicht Bestandteil der hier geprüften Hypothesen sind. Im Methodenteil werden daher ausschließlich die für die Analysen relevanten Variablen näher beschrieben.

Stichprobenplanung und Rekrutierung

Da es sich bei der vorliegenden Studie um eine erstmalige Untersuchung von Messreaktivität im Zusammenhang mit der täglichen Nahrungsaufnahme mithilfe einer Ernährungstracking-App handelt, lagen keine vergleichbaren empirischen Befunde vor, die eine fundierte a-priori-Stichprobenumfangsplanung erlaubt hätten. Entsprechend wurde eine angestrebte Stichprobengröße von $N = 200$ Teilnehmenden festgelegt. Da diese Zielgröße bis

Ende November 2025 nicht erreicht wurde, basieren die Analysen der vorliegenden Arbeit auf den bis zu diesem Zeitpunkt erhobenen Daten.

In die finale Stichprobe eingeschlossen wurden Personen ab einem Mindestalter von 18 Jahren, die bereits mindestens eine Woche vor Studienbeginn eine Ernährungstracking-App nutzten. Ausgeschlossen wurden Teilnehmende, die 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 Plattform surveycircle.com, soziale Medien, das Verteilen von Flyern sowie durch Mundpropaganda. Um systematische Verzerrungen des alltäglichen Essverhaltens durch veränderte Routinen während Feiertagen oder Ferienzeiten zu minimieren, wurde die Rekrutierung jeweils in einem Zeitraum von einer Woche vor bis eine Woche nach größeren Feiertagen beziehungsweise Ferien (Weihnachten und Ostern) pausiert. Darüber hinaus wurde die Studie eine Woche vor Semesterende geschlossen und zwei Wochen nach Semesterbeginn erneut geöffnet, um typische Ernährungsgewohnheiten während des Semesters abzubilden. Als Kompensation erhielten die Teilnehmenden vier LABS-Credits für Studierende der Psychologie an der Universität Wien oder die Möglichkeit zur Teilnahme an einer Verlosung von drei Gutscheinen im Wert von jeweils 100 €.

Design und Ablauf

Das Studiendesign orientiert sich an dem Ansatz von König et al. (2025) und wurde für die vorliegenden Fragestellungen entsprechend adaptiert. Methodisch handelt es sich um eine Beobachtungsstudie im Within-Subjects-Design mit zwei Messzeitpunkten. Die Datenerhebung erfolgte vollständig online mittels zweier standardisierter Fragebögen, die im Abstand von einer Woche bearbeitet wurden. Alle Teilnehmenden bearbeiteten dieselben Fragebögen in identischer Reihenfolge. Da das Ziel darin bestand, intraindividuelle Veränderungen des Ernährungsverhaltens unter Forscherbeobachtung zu erfassen, war eine Randomisierung der Reihenfolge weder vorgesehen noch sinnvoll.

Zum ersten Erhebungszeitpunkt wurden zunächst soziodemografische Variablen wie Alter und Geschlecht sowie Informationen zur verwendeten Ernährungs-Tracking-App erhoben. Darüber hinaus wurden die Big-Five-Persönlichkeitsmerkmale erfasst. Die Teilnehmenden

übertrugen ihre in der jeweiligen App dokumentierten Werte zur täglichen Energieaufnahme in Kilokalorien sowie zur Makronährstoffzufuhr hinsichtlich Fett, Kohlenhydraten und Proteinen für die vergangenen sieben Tage in den Fragebogen. Die Datenerhebung stützte sich somit auf retrospektiv berichtete, jedoch app-basierte Aufzeichnungen und nicht auf subjektiven Schätzungen des Ernährungsverhaltens. Am Ende des ersten Fragebogens hinterlegten die Teilnehmenden ihre E-Mail-Adresse, um eine Woche später automatisiert zur zweiten Befragung eingeladen zu werden. In der Zwischenzeit wurden sie instruiert, ihre Ernährung wie gewohnt weiter in ihrer Tracking-App zu dokumentieren. Durch die Ankündigung der erneuten Datenerhebung wurde gezielt ein Bewusstsein für die fortgesetzte wissenschaftliche Beobachtung geschaffen, wodurch die potenzielle Reaktivität der Messung zusätzlich begünstigt wurde.

Beim zweiten Erhebungszeitpunkt gaben die Teilnehmenden erneut ihre app-basierten Angaben zur Energie- und Makronährstoffzufuhr der vergangenen sieben Tage an. Für die vorliegende Arbeit wurden dabei ausschließlich die ersten beiden Tage dieser zweiten Woche berücksichtigt, da frühere Befunde darauf hindeuten, dass reaktive Effekte insbesondere zu Beginn einer Beobachtungsphase am stärksten ausgeprägt sind (Arigo & König, 2024). Zusätzlich wurden die Teilnehmenden im zweiten Fragebogen gefragt, ob sie subjektiv Veränderungen ihres Ernährungsverhaltens wahrgenommen hatten und wie sie diese gegebenenfalls erklärten. Eine vollständige Aufklärung über die eigentliche Zielsetzung der Studie erfolgte erst nach Abschluss des zweiten Fragebogens im Rahmen eines Debriefings. Bis zu diesem Zeitpunkt war den Teilnehmenden lediglich bekannt, dass ihre Ernährungsdaten zu wissenschaftlichen Zwecken erhoben und ausgewertet werden.

Maße und Materialien

Zur Erfassung der für die vorliegende Arbeit relevanten Variablen wurden Selbstberichtsdaten sowie app-basierte Angaben zum Ernährungsverhalten herangezogen. Im Folgenden werden die zentralen Variablen und deren Operationalisierung beschrieben.

Das Alter wurde in Jahren als kontinuierliche Variable erfasst. Zusätzlich gaben die Teilnehmenden die von ihnen verwendete Ernährungstracking-App an (König et al., 2018). Diese Angaben wurden ausschließlich deskriptiv ausgewertet und nicht in die inferenzstatistischen Analysen einbezogen.

Das Geschlecht wurde als soziodemografische Variable mittels Selbstauskunft erhoben. Die Teilnehmenden konnten zwischen den Kategorien weiblich, männlich und divers wählen. Für die Moderationsanalyse wurde das Geschlecht dichotom kodiert (0 = weiblich, 1 = männlich). Personen mit der Angabe „divers“ wurden aufgrund der geringen Fallzahl nicht in die geschlechtsspezifischen Analysen einbezogen.

Die Gewissenhaftigkeit wurde mithilfe einer Kurzsкала auf Basis des Big-Five-Inventars (BFI-10; Rammstedt & John, 2007) erfasst. Die Skala bestand aus zwei Items („Ich erledige Aufgaben gründlich“ und „Ich bin bequem, neige zur Faulheit“ [invertiert]), die auf einer fünfstufigen Likert-Skala von 1 = trifft gar nicht zu bis 5 = trifft völlig zu beantwortet wurden. Für die Datenanalyse wurde das negativ gepolte Item umkodiert. Anschließend wurde ein Mittelwert über beide Items gebildet, wobei höhere Werte eine höhere Ausprägung an Gewissenhaftigkeit anzeigen. Zur besseren Interpretierbarkeit der Interaktionseffekte wurde die Variable vor den Analysen zentriert.

Der tägliche Fettkonsum wurde in Gramm erfasst und basierte auf den in den jeweiligen Ernährungstracking-Apps dokumentierten Angaben der Teilnehmenden (König et al., 2025). Der durchschnittliche tägliche Fettkonsum in Gramm wurde als abhängige Variable in den Analysen verwendet. Hierzu wurde der durchschnittliche Fettkonsum getrennt für die sieben Tage der ersten Woche sowie für die ersten beiden Tage der zweiten Woche berechnet, um Unterschiede in der Beobachtungsdauer angemessen zu berücksichtigen. Fehlende Einträge wurden nicht als Nullwerte interpretiert, sondern als fehlende Werte kodiert. Die unabhängige Variable „Woche“ operationalisierte die Beobachtungsbedingung und wurde dichotom kodiert (0 = Woche 1/Baseline; 1 = erste beiden Tage von Woche 2 unter Forscherbeobachtung).

Eine Übersicht der Fragebogeninhalte kann der Tabelle 1 entnommen werden.

Tabelle 1

Übersicht der Studienmaterialien

Erhobene Daten	1. Teil	2. Teil	Quelle
Soziodemographische Daten (Alter, Geschlecht)	x		
Erfahrung und Nutzen von Ernährungsapp (Verwendete App)	x		König et al. (2018)

Big-Five-Persönlichkeitsmerkmale (Offenheit, Gewissenhaftigkeit, Extraversion, Verträglichkeit und Neurotizismus)	x		Rammstedt & John (2007)
Tägliche Kalorien- und Makronährstoffaufnahme der letzten 7 Tage	x	x	König et al. (2025)

Anmerkung. In der Tabelle sind ausschließlich jene Variablen und Items aufgeführt, die für die Fragestellungen der vorliegenden Arbeit relevant sind.

Datenaufbereitung und Qualitätsprüfung

Nach Abschluss der Datenerhebung wurden mehrere, vorab definierte Ausschlusskriterien angewendet.

Eine Person wurde ausgeschlossen, da die Erhebung der Gewissenhaftigkeit unvollständig war und nicht beide Items beantwortet wurden. Darüber hinaus wurden 252 Teilnehmende von den Analysen ausgeschlossen, bei denen die Fettangaben an mindestens sechs der sieben Tage im ersten Fragebogen und/oder an den ersten beiden Tagen des zweiten Fragebogens fehlten. Diese Kriterien waren in der Präregistrierung festgelegt und stellen sicher, dass der tägliche Fettkonsum ausreichend valide operationalisiert werden kann. Für die verbleibenden Datensätze wurden Tage ohne dokumentierte Einträge nicht als tatsächliche Nullwerte interpretiert, sondern als fehlende Werte behandelt. Entsprechend wurden im Datensatz vorhandene 0-Angaben, die auf nicht erfolgte Einträge zurückzuführen waren, in fehlende Werte umkodiert. Teilweise fehlende Angaben einzelner Tage führten nicht automatisch zum Ausschluss einer Person. Da lineare Mehrebenenmodelle gegenüber unvollständigen Daten auf Ebene wiederholter Messungen als vergleichsweise robust gelten, können vorhandene Beobachtungen innerhalb einer Person in die Schätzung einbezogen werden, selbst wenn einzelne Tageswerte fehlen (Hox et al., 2017). Voraussetzung war allerdings, dass die Fettwerte der ersten beiden Tage der zweiten Woche vollständig vorlagen, da diese für die Hypothesenprüfung zwingend erforderlich waren.

Zusätzlich wurden neun weitere Personen ausgeschlossen, weil sie den zweiten Fragebogen nicht innerhalb des vorgegebenen Zeitfensters von 24 Stunden nach Erhalt der Einladung bearbeitet hatten. Diese zeitliche Einschränkung war notwendig, um die Vergleichbarkeit der Messzeitpunkte zu gewährleisten.

Zur Überprüfung der Robustheit der Ergebnisse wurde eine Sensitivitätsanalyse durchgeführt, die der systematischen Kontrolle diente, ob einzelne extreme Werte oder Personen die Ergebnisse der Hauptanalysen maßgeblich beeinflussen. Dabei wurden Versuchspersonen identifiziert, die mindestens einen extremen Tageswert des Fettkonsums aufwiesen. Als Kriterium zur Bestimmung extremer Werte wurde die $1.5 \times \text{IQR}$ -Regel herangezogen. Auf diese Weise wurden insgesamt 12 Personen mit mindestens einem auffälligen Tageswert identifiziert. Anschließend wurde geprüft, ob ein inhaltlich begründeter Ausschluss einzelner Personen gerechtfertigt war. Im Rahmen der Sensitivitätsanalyse wurde ein Fall identifiziert, der durch durchgehend extrem hohe Werte des täglichen Fettkonsums auffiel ($M = 361.1 \text{ g/Tag}$, Spannweite: 200–500 g). Diese Werte lagen deutlich außerhalb der Verteilung der Stichprobe und überschritten den durchschnittlichen Fettkonsum um ein Vielfaches. Darüber hinaus widerspricht eine derart hohe Fettzufuhr ernährungsphysiologischen Referenzwerten, nach denen die empfohlene tägliche Fettaufnahme für Erwachsene typischerweise im Bereich von etwa 60 bis 80 g liegt (DGE, 2015). Die beobachteten Werte entsprechen einer Energiezufuhr von bis zu 4500 kcal allein aus Fett und übersteigen damit den üblichen täglichen Gesamtenergiebedarf deutlich. Da die erhöhten Werte konsistent über alle Erhebungstage hinweg auftraten und somit nicht auf einzelne fehlerhafte Eingaben zurückgeführt werden konnten, wurde dieser Fall als inhaltlich unplausibel eingestuft und aus den Analysen ausgeschlossen. Für die verbleibenden 11 Personen lagen zwar einzelne extreme Werte vor, diese erschienen jedoch inhaltlich plausibel und ließen sich nicht eindeutig als Mess- oder Eingabefehler interpretieren. Da für diese Fälle kein theoretisch oder methodisch begründeter Ausschlussgrund vorlag, wurden sie in den statistischen Analysen beibehalten.

Nach Anwendung aller Ausschlusskriterien ergab sich eine finale Analysestichprobe von $N = 118$ Personen. Eine Person mit non-binärer Geschlechtsangabe wurde in der geschlechtsspezifischen Moderationsanalyse aufgrund der zu geringen Gruppengröße nicht berücksichtigt, verblieb jedoch in den übrigen Analysen.

Da die Skala zur Erfassung von Gewissenhaftigkeit nur aus zwei Items bestand, wurde zur Reliabilitätsprüfung die Inter-Item-Korrelation berechnet. Die Korrelation zwischen den beiden Items war gering ($r = .09$) und erreichte nicht das in der Präregistrierung festgelegte Kriterium von $r \geq .30$. Die Skala weist somit in der vorliegenden Stichprobe keine zufriedenstellende interne Konsistenz auf.

Zur Prüfung der Annahme vollständig zufällig fehlender Werte (Missing Completely at Random, MCAR) wurde der Little's MCAR-Test durchgeführt. Der Test ergab kein signifikantes Ergebnis, $\chi^2(90) = 102.00$, $p = .184$. Somit kann davon ausgegangen werden, dass die fehlenden Werte vollständig zufällig auftreten. Entsprechend der Präregistrierung wurden die Analysen ohne zusätzliche Imputation der fehlenden Werte durchgeführt.

Modellprüfung

In der Präregistrierung der Masterarbeit wurden keine spezifischen Annahmenprüfungen festgelegt. Zur methodischen Absicherung wurden die Modellannahmen des linearen Mehrebenenmodells dennoch grafisch überprüft. Hierzu wurden Q-Q-Plots der Residuen, Residuen-gegen-Vorhersage-Plots sowie Diagnostikplots zur Identifikation potenziell einflussreicher Beobachtungen herangezogen. Die grafische Modellprüfung ergab keine Hinweise auf gravierende Verletzungen der Modellannahmen. Es zeigten sich leichte Abweichungen von der Normalverteilung im oberen Residuenbereich sowie eine geringfügige Varianzheterogenität, die jedoch als unkritisch eingestuft wurden.

Deskriptivstatistische Auswertung

Im ersten Schritt der statistischen Auswertung mit R (Version 4.5.2; R Core Team, 2025) erfolgte eine deskriptive Beschreibung der Stichprobe sowie der zentralen Untersuchungsvariablen. Zur Beschreibung der soziodemografischen Variablen wurde das Alter der Teilnehmenden anhand von Mittelwert (M) und Standardabweichung (SD) berichtet. Das Geschlecht wurde in den Kategorien weiblich, männlich und divers erfasst und in Form absoluter und relativer Häufigkeiten dargestellt. Darüber hinaus wurde die verwendete Ernährungstracking-App in absoluten als auch in prozentualen Häufigkeiten dokumentiert.

Anschließend wurden die zentralen Variablen zur Hypothesenprüfung deskriptiv ausgewertet. Für den täglichen Fettkonsum wurden getrennt für Woche 1 (Baseline) sowie für die ersten beiden Tage von Woche 2 unter Forscherbeobachtung Mittelwert (M), Standardabweichung (SD) sowie Minimal- und Maximalwerte berichtet. Zusätzlich wurden die entsprechenden deskriptiven Statistiken getrennt nach Geschlecht dargestellt. Die Moderatorvariable Gewissenhaftigkeit wurde ebenfalls mittels Mittelwert und Standardabweichung beschrieben. Insgesamt bilden diese deskriptiven Analysen die Grundlage für die anschließenden inferenzstatistischen Prüfungen der Hypothesen und tragen

zur transparenten Darstellung der Stichprobencharakteristika sowie der Variablenverteilungen bei.

Inferenzstatistische Auswertung

Die inferenzstatistischen Analysen wurden mit R (Version 4.5.2; R Core Team, 2025) durchgeführt. Zur Prüfung der Hypothesen kamen lineare Mehrebenenmodelle (LMM) mit zufälligen Intercepts pro Person zum Einsatz, spezifiziert über das Paket *lme4*. Da mehrere Tagesmessungen innerhalb derselben Personen vorlagen, wurde die hierarchische Datenstruktur berücksichtigt. Die einzelnen Tageswerte des Fettkonsums stellten Level-1-Beobachtungen dar, die in Personen (Level 2) verschachtelt waren. Um der intraindividuellen Abhängigkeit der Messwerte angemessen abzubilden, wurden lineare Mixed-Effects-Modelle mit zufälligem Intercept auf Personenebene spezifiziert. Dadurch wird erlaubt, dass sich die Ausgangsniveaus des Fettkonsums zwischen den Personen unterscheiden, während der Effekt der Woche als fixer Effekt modelliert wird. Die Analysen wurden in R unter Verwendung des Pakets *lme4* und *lmer* (Bates et al., 2025) berechnet. Die Modelle wurden mittels Maximum-Likelihood-Schätzung (ML) geschätzt, um Modellvergleiche auf Basis von Likelihood-Ratio-Tests zu ermöglichen. Obwohl die Hypothesen gerichtet formuliert waren, erfolgten die Signifikanztests konservativ zweiseitig bei einem α -Niveau von .05.

Zur Überprüfung von Hypothese 1 wurde zunächst ein Basismodell spezifiziert, in dem „Woche“ als Prädiktor des täglichen Fettkonsums aufgenommen wurde. Der fixe Effekt von Woche bildet dabei die durchschnittliche intraindividuelle Veränderung des Fettkonsums zwischen Woche 1 und den ersten beiden Tagen von Woche 2 ab. Für Hypothese 2 wurde das Geschlecht als Moderator ergänzt. Zusätzlich zum Haupteffekt von Geschlecht wurde eine Interaktion zwischen Woche (Level 1) und Geschlecht (Level 2) analysiert. Diese Cross-Level-Interaktion prüft, ob sich die Veränderung des Fettkonsums zwischen den Wochen systematisch zwischen Frauen und Männern unterscheidet. Im Rahmen der explorativen Fragestellung (Hypothese 3) wurde Gewissenhaftigkeit als weiterer Moderator in das Modell integriert. Analog zu Hypothese 2 wurde eine Interaktion zwischen Woche und Gewissenhaftigkeit modelliert, um zu prüfen, ob die intraindividuelle Veränderung des Fettkonsums in Abhängigkeit vom Ausmaß der Gewissenhaftigkeit variiert.

Abweichend von der Präregistrierung wurde zur Quantifizierung der erklärten Varianz Effektstärken in Form von Within- R^2 , Between- R^2 und Total- R^2 berechnet. Die gewählte

Differenzierung erlaubt eine präzise Trennung zwischen aufgeklärter Varianz auf der Personenebene und innerhalb von Personen über die Messzeitpunkte hinweg.

Ergebnisse

Stichprobe

Die finale Stichprobe umfasste 118 Teilnehmende. Das durchschnittliche Alter betrug $M = 27.40$ Jahre ($SD = 8.62$) bei einer Altersspanne von 18 bis 65 Jahren. In Bezug auf das Geschlecht identifizierten sich 66.9 % der Teilnehmenden als weiblich ($n = 79$), 32.2 % als männlich ($n = 38$) und 0.8 % als divers ($n = 1$). Die Teilnehmenden nutzten eine Vielzahl unterschiedlicher Ernährungstracking-Apps. Am häufigsten wurde Yazio verwendet (47.9 %, $n = 56$), gefolgt von MyFitnessPal (19.7 %, $n = 23$) und Lifesum (7.7 %, $n = 9$). Eine vollständige Übersicht der deskriptiven Statistiken der Stichprobe ist in Tabelle 2 zu finden.

Tabelle 2

Stichprobenbeschreibung der Teilnehmenden (N = 118)

Variable	<i>n</i>	%	<i>M</i>	<i>SD</i>
Alter (Jahre)	118	—	27.40	8.62
Geschlecht				
Weiblich	79	66.9	—	—
Männlich	38	32.2	—	—
Divers	1	0.8	—	—
Ernährungstracking-App				
Yazio	56	47.9	—	—
MyFitnessPal	23	19.7	—	—
Lifesum	9	7.7	—	—
FDDB	7	6.0	—	—
FatSecret	5	4.3	—	—
Foodvisor	4	3.4	—	—
Samsung Health	3	2.6	—	—
Cronometer	2	1.7	—	—
Sonstiges	8	6.8	—	—

Anmerkung. Ernährungstracking-Apps, die nur einmal genannt wurden, wurden zur Kategorie Sonstiges zusammengefasst.

Ergebnisdarstellung

Der tägliche Fettkonsum ist in den ersten beiden Tagen unter Forscherbeobachtung signifikant kleiner als in der Woche ohne Forscherbeobachtung.

Zur deskriptiven Beschreibung des Fettkonsums wurden für beide Zeiträume Mittelwerte über die jeweils erfassten Tage berechnet, um Unterschiede in der Beobachtungsdauer zu berücksichtigen. Die Ergebnisse der Deskriptivstatistik der H1 sind in Tabelle 3 dargestellt.

Tabelle 3

Deskriptivstatistiken des Fettkonsums in Woche 1 vs. erste zwei Tage Woche 2

Zeitraum	<i>M</i>	<i>SD</i>	Min	Max
Woche 1	68.50	25.90	21.40	172.00
Woche 2 (Tage 1–2)	67.40	31.40	4.50	222.00

Zur Überprüfung der H1 wurde ein lineares Mehrebenenmodell mit zufälligem Intercept für die Versuchspersonen gerechnet. Der tägliche Fettkonsum wurde dabei als abhängige Variable modelliert, wobei die Woche (Woche 1 vs. erste zwei Tage der zweiten Woche) als Prädiktor in das Modell aufgenommen wurde. Der Effekt der Woche auf den täglichen Fettkonsum war nicht signifikant ($p = .496$). Damit zeigte sich kein statistisch bedeutsamer Unterschied im durchschnittlichen Fettkonsum zwischen Woche 1 und Woche 2. Das Modell erklärte insgesamt 55.2 % der Varianz im täglichen Fettkonsum ($\text{Total-}R^2 = .552$). Die H1 konnte nicht bestätigt werden. Die Parameter der Hypothese sind in Tabelle 4 dargestellt.

Tabelle 4

Feste Effekte des Mehrebenen-Regressionsmodells zur Prüfung von Hypothese 1

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	68.52	[63.77, 73.27]	2.42	28.29	123.02	< .001	-	-	-
Woche	-1.13	[-4.39, 2.12]	1.66	-0.68	922.53	.496	.000	.000	.552

Anmerkung. Within- R^2 bezeichnet den erklärten Varianzanteil innerhalb von Personen (Ebene der Messzeitpunkte), Between- R^2 den erklärten Varianzanteil zwischen Personen und Total- R^2 den gesamten erklärten Varianzanteil des Modells. Woche: 0 = Woche 1, 1 = erste zwei Tage der zweiten Woche.

Frauen zeigen eine stärkere Veränderung des täglichen Fettkonsums in den ersten beiden Tagen unter Forscherbeobachtung im Vergleich zur Woche ohne Forscherbeobachtung als Männer.

Es wurde eine non-binäre Person aufgrund der zu kleinen Gruppengröße von der geschlechtsspezifischen Hypothese ausgeschlossen. Zur deskriptiven Beschreibung des Fettkonsums der Geschlechter „männlich“ und „weiblich“ wurden für beide Zeiträume und beide Geschlechter Mittelwerte über die jeweils erfassten Tage berechnet, um Unterschiede in der Beobachtungsdauer zu beachten. Die Ergebnisse der Deskriptivstatistik der H2 sind in Tabelle 5 dargestellt.

Tabelle 5

Deskriptive Statistiken des täglichen Fettkonsums getrennt nach Geschlecht für Woche 1 und die ersten beiden Tage unter Forscherbeobachtung

Geschlecht	Zeitraum	<i>M</i>	<i>SD</i>	Min	Max
Weiblich	Woche 1	62.60	19.40	21.40	119.00
	Woche 2 (Tage 1–2)	60.80	26.00	4.50	139.00
Männlich	Woche 1	80.70	33.10	26.40	172.00
	Woche 2 (Tage 1–2)	81.60	37.20	25.00	222.00

Anmerkung. Die Analysen nach Geschlecht beziehen sich ausschließlich auf männliche und weibliche Teilnehmende

Die H2 prüfte, ob Frauen eine stärkere Veränderung des täglichen Fettkonsums in den ersten beiden Tagen unter Forscherbeobachtung im Vergleich zur Woche ohne Forscherbeobachtung als Männer aufzeigen. Der Haupteffekt der Woche war nicht signifikant ($p = .355$). Dagegen ergab sich ein signifikanter Haupteffekt des Geschlechts, wobei männliche Teilnehmende insgesamt einen höheren Fettkonsum berichteten als weibliche Teilnehmende ($b = 18.13, p < .001$).

Entscheidend für die H2 war der Interaktionseffekt zwischen Woche und Geschlecht. Dieser erwies sich als nicht signifikant ($p = .451$). Daraus ergibt sich keine Moderation des Effekts der Forscherbeobachtung durch das Geschlecht. Das Modell erklärt insgesamt 56 % der Varianz (Total- $R^2 = .554$). Die Parameter der H2 sind in Tabelle 6 dargestellt.

Tabelle 6*Feste Effekte des Mehrebenen-Regressionsmodells zur Prüfung von 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	62.67	[57.16, 8.18]	2.81	22.30	121.36	< .001	—	—	—
Woche	-1.88	[-5.87, 2.10]	2.03	-0.93	915.37	.355	—	—	—
Geschlecht	18.13	[8.46, 27.81]	4.94	3.67	121.80	< .001	.068	.068	.554
Woche × Geschlecht	2.70	[-4.31, 9.70]	3.57	0.75	915.75	.451	—	—	—

Anmerkung. Within-*R*² bezeichnet den erklärten Varianzanteil innerhalb von Personen (Ebene der Messzeitpunkte), Between-*R*² den erklärten Varianzanteil zwischen Personen und Total-*R*² den gesamten erklärten Varianzanteil des Modells. Woche: 0 = Woche 1, 1 = erste zwei Tage der zweiten Woche. Geschlecht: 0 = weiblich, 1 = männlich.

Es wird untersucht, ob und auf welche Weise Gewissenhaftigkeit den Effekt der Forscherbeobachtung auf den täglichen Fettkonsum moderiert.

Die Gewissenhaftigkeit der Stichprobe lag im Mittel bei $M = 3.64$ ($SD = 0.68$) auf einer fünfstufigen Likert-Skala.

In einer explorativen Analyse wurde geprüft, ob Gewissenhaftigkeit den Effekt der Forscherbeobachtung auf den täglichen Fettkonsum moderiert. Hierzu wurde ein lineares Mehrebenenmodell mit zufälligem Intercept für Personen gerechnet, das neben den Haupteffekten der Woche und der Gewissenhaftigkeit auch deren Interaktion enthielt. Die Interaktion zwischen Woche und Gewissenhaftigkeit war nicht signifikant ($p = .313$). Auch der Haupteffekt der Gewissenhaftigkeit auf den täglichen Fettkonsum war nicht signifikant ($p = .303$). Das Modell erklärte insgesamt 55.4 % der Varianz im täglichen Fettkonsum (Total- $R^2 = .554$). Die vollständigen Parameter der explorativen Analyse sind in Tabelle 7 dargestellt.

Tabelle 7*Feste Effekte des Mehrebenen-Regressionsmodells zur Prüfung der explorativen Fragestellung*

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	68.52	[63.77, 73.28]	2.42	28.27	121.96	< .001	—	—	—

Woche	-1.13	[-4.39, 2.12]	1.66	-0.68	921.53	.495	—	—	—
Gewissenhaftigkeit	-3.71	[-10.75, 3.33]	3.59	-1.03	122.10	.303	.005	.005	.554
Woche × Gewissenhaftigkeit	2.48	[-2.34, 7.31]	2.46	1.01	921.42	.313	—	—	—

Anmerkung. Within-R² bezeichnet den erklärten Varianzanteil innerhalb von Personen (Ebene der Messzeitpunkte), Between-R² den erklärten Varianzanteil zwischen Personen und Total-R² den gesamten erklärten Varianzanteil des Modells. Woche: 0 = Woche 1, 1 = erste zwei Tage der zweiten Woche. Gewissenhaftigkeit wurde als Mittelwert der entsprechenden Items berechnet und zur Analyse am Stichprobenmittelwert zentriert.

Diskussion

Die vorliegende Arbeit untersuchte das Auftreten von Messreaktivität im Kontext digitaler Ernährungserfassung, verstanden als systematische Veränderungen des Ernährungsverhaltens infolge der Studienteilnahme. Für diesen Anwendungsbereich liegt bislang kaum empirische Evidenz vor, weshalb die Untersuchung darauf abzielte, einen Beitrag zur Schließung dieser Forschungslücke zu leisten.

Grundlage bildete eine Beobachtungsstudie auf Basis von Datenspenden, in der innerhalb derselben Personen zwei Messzeitpunkte im Abstand von einer Woche miteinander verglichen wurden. Im Zentrum stand die Frage, ob sich der tägliche Fettkonsum in den ersten beiden Tagen unter Forscherbeobachtung im Vergleich zu einer Woche ohne Beobachtung verändert (RQ1). Darüber hinaus wurde geprüft, ob ein möglicher reaktiver Effekt durch das Geschlecht (RQ2) sowie durch individuelle Unterschiede in Gewissenhaftigkeit (RQ3) moderiert wird.

Die Ergebnisse zeigten weder einen signifikanten kurzfristigen Rückgang des Fettkonsums unter Forscherbeobachtung noch Hinweise auf moderierende Effekte von Geschlecht oder Gewissenhaftigkeit. Insgesamt liefern die Befunde somit keine Evidenz für eine ausgeprägte Messreaktivität im Kontext digitaler Ernährungserfassung in der untersuchten Stichprobe. Im Folgenden werden die zentralen Ergebnisse der Studie anhand bestehender theoretischer Ansätze sowie empirischer Befunde eingeordnet und diskutiert.

Theoretische Einordnung und Interpretation der Befunde

Effekt der Forscherbeobachtung auf den Fettkonsum (RQ1)

Die erste Hypothese der Studie ging davon aus, dass der tägliche Fettkonsum in den ersten beiden Tagen unter Forscherbeobachtung signifikant geringer ausfallen würde als in der Woche ohne Beobachtung. In der untersuchten Stichprobe zeigte sich jedoch kein Hinweis auf einen kurzfristigen Rückgang des Fettkonsums infolge der Studienteilnahme, sodass die Hypothese nicht bestätigt werden konnte.

Dieses Ergebnis steht nur eingeschränkt im Einklang mit zentralen Annahmen zur Messreaktivität, denen zufolge bereits die Erfassung eines Verhaltens Prozesse der Selbstbeobachtung und Selbstregulation aktivieren kann, die kurzfristige Verhaltensanpassungen begünstigen (French et al., 2021). Empirische Befunde aus dem

Verhaltenskontext deuten entsprechend darauf hin, dass sich Verhaltensmuster im Verlauf einer Studie systematisch verändern können (Arigo & König, 2024; König et al., 2022a). Häufig zeigt sich zu Beginn einer Erhebung eine verstärkte Ausprägung sozial erwünschter oder gesundheitsförderlicher Verhaltensweisen, die im weiteren Verlauf wieder abnimmt (Arigo & König, 2024). Umgekehrt können weniger erwünschte Verhaltensweisen initial reduziert auftreten und im Zeitverlauf wieder zunehmen. Ausgehend von diesen Befunden wäre ein zumindest geringer Rückgang des Fettkonsums unter Forscherbeobachtung theoretisch erwartbar gewesen (Arigo & König, 2024; König et al., 2022a).

Kontextabhängige Ausprägung von Messreaktivität

Gleichzeitig relativiert der bisherige empirische Forschungsstand die Erwartung eindeutig nachweisbarer reaktiver Effekte. So zeigen systematische Übersichtsarbeiten und empirische Studien, dass Messreaktivität in vielen Fällen nur in kleiner bis moderater Ausprägung auftritt und zudem eine erhebliche Variabilität aufweist, die in Abhängigkeit vom untersuchten Verhalten, dem Studiendesign sowie den eingesetzten Messinstrumenten variiert (French & Sutton, 2010; French et al., 2021; König et al., 2022a).

Die vorliegende Studie beschränkte sich auf die ersten beiden Tage unter Forscherbeobachtung, da frühere Befunde darauf hindeuten, dass reaktive Effekte insbesondere zu Beginn einer Beobachtungsphase auftreten können (König et al., 2022a). Gleichzeitig ist jedoch zu berücksichtigen, dass der zeitliche Verlauf von Messreaktivität nicht einheitlich ist. Entsprechende Effekte können unter bestimmten Bedingungen sehr kurzfristig auftreten, beispielsweise ausschließlich am ersten Tag der Messung, oder sich erst im Verlauf mehrerer Tage oder Wochen entwickeln. In beiden Fällen besteht die Möglichkeit, dass sich reaktive Effekte im gewählten Vergleichszeitraum nicht eindeutig abbilden lassen.

Weiteres ist zu beachten, dass mögliche Verhaltensanpassungen im Bereich des Ernährungsverhaltens nicht ausschließlich von der Aktivierung kognitiver oder motivationaler Prozesse abhängen, sondern in hohem Maße durch alltagspraktische Rahmenbedingungen beeinflusst werden. Zeitliche Einschränkungen, finanzielle Belastungen oder eine begrenzte Verfügbarkeit bestimmter Lebensmittel können potenzielle Anpassungen zusätzlich erschweren (Pinho et al., 2018). Solche Faktoren wirken insbesondere dann einschränkend, wenn Veränderungen des Ernährungsverhaltens mit konkreten Anpassungen in der Lebensmittelauswahl oder der Zubereitung verbunden sind und somit über bloße intentionale

Prozesse hinausgehen. So kann es sein, dass eine Veränderung des Fettkonsums mehr Zeit benötigt als andere Formen reaktiver Verhaltensänderungen, die sich stärker auf kurzfristig modifizierbare Verhaltensaspekte beziehen. Kurzfristige Messzeiträume sind daher möglicherweise nicht ausreichend, um entsprechende Veränderungen valide abzubilden, da sich diese erst verzögert in messbaren Verhaltensindikatoren niederschlagen könnten.

Darüber hinaus konzentriert sich die aktuelle Studienlage zur Messreaktivität im Bereich digitaler, alltagsnaher Erhebungen überwiegend auf körperliche Aktivität, während Studien zur Nahrungsaufnahme vergleichsweise selten sind (König et al., 2022a). Die Übertragbarkeit der Ergebnisse von Untersuchungen bezüglich der physischen Aktivität auf das Ernährungsverhalten ist daher nur eingeschränkt anzunehmen. Dies erscheint deshalb relevant, da sich Ernährungsverhalten in zentralen Merkmalen vom Bewegungsverhalten unterscheidet und stärker von situativen, sozialen und routinisierten Faktoren geprägt ist (Jastran et al., 2009). Das Ausbleiben einer signifikanten Steigerung des Fettkonsums könnte demnach mit der Stabilität etablierter Ernährungsgewohnheiten zusammenhängen. Ess- und Kochroutinen, etwa in Form regelmäßiger Mahlzeiten, bevorzugter Zutaten oder typischer Zubereitungsweisen, sind häufig fest in alltägliche Abläufe integriert und eng mit individuellen Wertvorstellungen sowie sozialen Kontexten verbunden. Veränderungen solcher Routinen erfolgen daher in der Regel nicht kurzfristig, sondern setzen oft wiederholte Impulse oder einen längeren Anpassungsprozess voraus (Jastran et al., 2009).

Insgesamt wird Messreaktivität in der Literatur als kontextabhängiges Phänomen beschrieben, das unter bestimmten Bedingungen stärker ausgeprägt ist als unter anderen (French & Sutton, 2010; French et al., 2021). Angesichts der Komplexität von Messreaktivität erscheint es plausibel, dass kurzfristige Effekte nicht in allen Stichproben oder Untersuchungskontexten gleichermaßen nachweisbar sind, sondern in Abhängigkeit von spezifischen Rahmenbedingungen variieren.

Einfluss habituellder Selbstmonitoring-Prozesse auf Messreaktivität

Eine weitere mögliche Erklärung für das Ausbleiben signifikanter Effekte liegt in der bereits vor Studienbeginn etablierten Nutzung von Ernährungstracking-Apps innerhalb der untersuchten Stichprobe. In der gesundheitspsychologischen Literatur wird angenommen, dass wiederholtes Erfassen gesundheitsbezogener Verhaltensweisen Prozesse wie erhöhte Verhaltenssalienz oder Intentionsaktivierung auslösen kann, die Verhaltensanpassungen

begünstigen (French & Sutton, 2010; French et al., 2021). Die kontinuierliche Dokumentation des eigenen Ernährungsverhaltens stellt somit eine Form des Selbstmonitorings dar, das als zentraler Mechanismus von Messreaktivität gilt (Barta et al., 2012; Bravata et al., 2007).

Voraussetzung für die Studienteilnahme war, dass die Teilnehmenden bereits vor Studienbeginn über mindestens eine Woche hinweg eine Ernährungstracking-App genutzt hatten. Es ist daher davon auszugehen, dass ein Großteil der untersuchten Stichprobe diese Form der Protokollierung der Nahrungsaufnahme nicht erstmals im Rahmen der Studie, sondern bereits über einen längeren Zeitraum hinweg praktizierte. Infolgedessen ist anzunehmen, dass zentrale Prozesse der Selbstbeobachtung, Reflexion und Selbstregulation bereits vor Beginn der Untersuchung aktiviert waren und somit nicht erst durch die Teilnahme an der Studie ausgelöst wurden (Barta et al., 2012; French et al., 2021).

Theoretische Überlegungen sowie empirische Befunde legen nahe, dass Messreaktivität insbesondere zu Beginn einer Messphase ausgeprägt sein kann, wenn die Verhaltensmessung neu eingeführt wird und entsprechende Verhaltensaspekte erstmals verstärkt in den Fokus der Aufmerksamkeit rücken (French & Sutton, 2010; French et al., 2021; König et al., 2022a; Miles et al., 2020). In der vorliegenden Stichprobe, in der die Erfassung des Ernährungsverhaltens bereits etabliert war, könnte ein solcher initialer Reaktivitätseffekt daher weitgehend ausgeblieben sein.

Darüber hinaus zeigen empirische Befunde, dass Personen mit ungünstigerem Ausgangsniveau tendenziell stärkere Verhaltensänderungen im Zuge von Messungen aufweisen (French et al., 2021; Miles et al., 2020). Daher ist denkbar, dass Teilnehmende, die ihre Ernährung schon über längere Zeit hinweg erfassten, ihr Ernährungsverhalten stärker reflektierten und potenziell entsprechende Anpassungen bereits vorgenommen hatten. Dies könnte dazu geführt haben, dass das Ausgangsniveau des untersuchten Verhaltens innerhalb der Stichprobe vergleichsweise günstiger war und somit weniger Spielraum für weitere kurzfristige Veränderungen bestand.

Zusätzlich ist zu berücksichtigen, dass digitales Ernährungstracking unter den vorliegenden Bedingungen möglicherweise primär als etablierte Alltagsroutine und weniger als forschungsbezogene Messung wahrgenommen wurde (French et al., 2021; König et al., 2022b). Digitale Selbsttracking-Anwendungen werden häufig in alltägliche Routinen integriert und über die Zeit habitualisiert genutzt, wodurch die Erfassung des eigenen Verhaltens

zunehmend automatisiert erfolgt und weniger bewusste Aufmerksamkeit erfordert (Clark et al., 2024). In diesem Sinne verschiebt sich die Wahrnehmung der Datenerhebung von einer expliziten Messsituation hin zu einem selbstverständlichen Bestandteil alltäglicher Handlungen. Eine solche habitualisierte Nutzung könnte dazu beitragen, dass die Erhebungssituation weniger als neue oder besonders saliente Beobachtungssituation erlebt wurde, wodurch die Wahrscheinlichkeit reaktiver Anpassungsprozesse weiter reduziert wird. Wenn das Tracking nicht primär als externe Beobachtung, sondern als Teil der eigenen Routine wahrgenommen wird, verliert die Messung potenziell ihren auffordernden Charakter, der für die Auslösung von Verhaltensanpassungen zentral ist.

Insgesamt lässt sich daraus ableiten, dass zusätzliche Verhaltensanpassungen im Rahmen der Studienteilnahme nur in begrenztem Ausmaß zu erwarten waren. Die vorliegenden Ergebnisse liefern dementsprechend keine Hinweise auf einen kurzfristigen Rückgang des Fettkonsums infolge der Forscherbeobachtung. Vielmehr sprechen die Befunde dafür, dass Forscherbeobachtung im Kontext digitaler Ernährungserfassung in einer Population, die an kontinuierliches Selbstmonitoring gewöhnt ist, nicht zwangsläufig mit systematischen kurzfristigen Verhaltensveränderungen einhergeht.

Geschlecht als Moderator der Messreaktivität (RQ2)

Entgegen der theoretisch begründeten Annahme zeigte sich kein signifikanter moderierender Effekt des Geschlechts auf den Zusammenhang zwischen Forscherbeobachtung und täglichem Fettkonsum. Zwar ergab sich ein signifikanter Haupteffekt des Geschlechts, wonach männliche Teilnehmende insgesamt höhere Fettkonsumwerte berichteten als weibliche Teilnehmende. Die Veränderung des Fettkonsums zwischen Woche 1 und den ersten beiden Tagen unter Forscherbeobachtung unterschied sich jedoch nicht signifikant zwischen den Geschlechtern.

Dieser Befund steht zum Teil im Einklang mit der bisherigen uneinheitlichen Forschung zu geschlechtsspezifischen Unterschieden in der Messreaktivität. So berichten König et al. (2022a) in ihrer Meta-Analyse zu digitalem Ecological Momentary Assessment gesundheitsbezogener Verhaltensweisen, dass in der Mehrzahl der untersuchten Studien keine signifikanten Unterschiede zwischen männlichen und weiblichen Teilnehmenden festgestellt wurden. Zwar existieren einzelne Befunde, die auf stärkere reaktive Effekte bei

Frauen hindeuten (Ho et al., 2013), insgesamt scheint das Geschlecht jedoch kein konsistenter Moderator von Messreaktivität zu sein.

Vor dem Hintergrund theoretischer Überlegungen, wie aufgrund höherer durchschnittlicher Gewissenhaftigkeit bei Frauen (Vecchione et al., 2012), einer stärkeren Orientierung an gesundheitsbezogenen Normen oder einer erhöhten Tendenz zu sozial erwünschtem Antwortverhalten (Bernardi & Guptill, 2008), wäre eine stärkere Anpassung des Ernährungsverhaltens bei Frauen grundsätzlich zu erwarten gewesen. Gleichzeitig ist jedoch zu berücksichtigen, dass Frauen im Durchschnitt häufiger ein gesundheitsförderlicheres Ernährungsverhalten aufweisen als Männer (Goldberg & Strycker, 2002). Dadurch ist denkbar, dass in der vorliegenden Stichprobe bei weiblichen Teilnehmenden bereits ein vergleichsweise günstigeres Ausgangsniveau des Ernährungsverhaltens vorlag. Das könnte den Spielraum für weitere kurzfristige Anpassungen im Rahmen der Studienteilnahme eingeschränkt haben (Miles et al., 2020). Unter diesen Bedingungen wären zusätzliche Veränderungen des Fettkonsums weniger wahrscheinlich beziehungsweise würden in geringerem Ausmaß auftreten, wodurch sich potenzielle geschlechtsspezifische Unterschiede in der Messreaktivität weniger deutlich zeigen.

Außerdem ist zu berücksichtigen, dass Messreaktivität in vielen Fällen stärker durch situative und kontextuelle Faktoren als durch stabile Personenmerkmale beeinflusst wird (French et al., 2021). In der vorliegenden Studie könnten insbesondere die geringe Salienz der Forscherbeobachtung sowie die bereits routinisierte Nutzung von Ernährungstracking-Apps dazu beigetragen haben, dass sich mögliche geschlechtsspezifische Unterschiede weniger deutlich zeigen. Interaktionseffekte sind in der Regel schwieriger nachzuweisen als Haupteffekte und erfordern häufig größere Stichproben oder längere Messzeiträume (Leon & Heo, 2009). Es ist daher möglich, dass potenzielle geschlechtsspezifische Unterschiede in der Reaktivität im gewählten Design nicht ausreichend sensitiv erfasst werden konnten. Insgesamt sprechen die Ergebnisse jedoch dafür, dass sich ein kurzfristiger Effekt der Forscherbeobachtung auf den Fettkonsum nicht systematisch zwischen Frauen und Männern unterscheidet.

Gewissenhaftigkeit als Moderator der Messreaktivität (RQ3)

Bezüglich der explorativen Fragestellung zeigt sich in der vorliegenden Untersuchung kein signifikanter moderierender Effekt der Gewissenhaftigkeit auf den Zusammenhang zwischen

Forscherbeobachtung und täglichem Fettkonsum. Weder der Haupteffekt der Gewissenhaftigkeit noch die Interaktion zwischen Woche und Gewissenhaftigkeit erreichten statistische Signifikanz. Damit ergaben sich keine Hinweise darauf, dass gewissenhaftere Personen im Rahmen der digitalen Ernährungserfassung stärker oder schwächer auf die Messsituation reagierten als weniger gewissenhafte Personen.

Im Gegensatz dazu wird Gewissenhaftigkeit in der bisherigen Forschung mit gesundheitsförderlichem Verhalten in Verbindung gebracht (Goldberg & Strycker, 2002). Frühere Arbeiten konzentrieren sich jedoch überwiegend auf direkte Zusammenhänge zwischen Persönlichkeitsmerkmalen und Essgewohnheiten, während die vorliegende Arbeit die moderierende Rolle von Gewissenhaftigkeit im Kontext von Messreaktivität betrachtete. Es ist daher möglich, dass Gewissenhaftigkeit zwar mit stabilen Unterschieden im Ernährungsverhalten einhergeht, jedoch nicht zwangsläufig beeinflusst, in welchem Ausmaß Personen ihr Verhalten kurzfristig als Reaktion auf eine Mess- oder Studiensituation anpassen.

Eine mögliche Erklärung hierfür liegt darin, dass gewissenhafte Personen ihr Verhalten bereits unabhängig von der Studienteilnahme stärker regulieren und an persönlichen Standards ausrichten. Gewissenhafte Personen tendieren zu einem gesunden Ernährungsverhalten, was auf einen geringen Fettkonsum als Ausgangsniveau schließt (Goldberg & Strycker, 2002). Wenn das Ernährungsverhalten schon vor Beginn der Untersuchung relativ stabil und zielgerichtet organisiert ist, könnte der zusätzliche Einfluss einer kurzfristigen Messsituation gering ausfallen. Aufgrund des günstigeren Ausgangsniveaus wäre zu erwarten, dass Messreaktivität insgesamt schwächer ausgeprägt ist und sich Unterschiede zwischen Personen mit hoher und niedriger Gewissenhaftigkeit nur begrenzt zeigen (Miles et al., 2020).

Darüber hinaus könnte die geringe Reliabilität der eingesetzten Kurzskala zur Erfassung von Gewissenhaftigkeit (Rammstedt & John, 2007) anhand von nur zwei Items die Nachweisbarkeit moderierender Effekte eingeschränkt haben. Die niedrige Inter-Item-Korrelation weist auf einen erhöhten Messfehler hin, der insbesondere bei der Untersuchung von Interaktionseffekten zu einer Unterschätzung tatsächlicher Zusammenhänge führen kann. Es ist daher nicht auszuschließen, dass potenzielle Effekte der Gewissenhaftigkeit aufgrund begrenzter Messpräzision nicht erfasst wurden. Entsprechend sollten Befunde zu

Gewissenhaftigkeit vorsichtig interpretiert werden, da Messfehler die Stärke möglicher Zusammenhänge reduziert haben könnten.

Insgesamt unterstreicht das vorliegende nicht signifikante Ergebnis der explorativen Fragestellung zur Moderation von Gewissenhaftigkeit die Notwendigkeit weiterer Forschung, die reliablere Erfassungen von Persönlichkeitsmerkmalen einbezieht.

Stärken

Eine zentrale Stärke der vorliegenden Arbeit liegt in der Anwendung eines methodisch innovativen Ansatzes (König & Ruf, 2024), bei dem auf bereits vorhandene Daten aus den von den Teilnehmenden genutzten Ernährungstracking-Apps zurückgegriffen wurde. Dadurch war es möglich, Ernährungsverhalten unter realen Alltagsbedingungen zu erfassen und auf dieser Grundlage die Rolle der Forscherbeobachtung im Kontext digitaler Ernährungserhebung zu untersuchen. Im Unterschied zu vielen bisherigen Studien, die auf selbst eingeführte Messinstrumente oder experimentelle Settings zurückgreifen, basiert die Datenerhebung somit auf tatsächlich genutzten, etablierten Anwendungen, was zu einer hohen ökologischen Validität der erhobenen Daten beiträgt.

Ein wesentlicher Vorteil dieses Vorgehens besteht darin, dass methodische Einschränkungen früherer Untersuchungen (French & Sutton, 2010; French et al., 2021, König et al., 2022a; König et al., 2025) zur Messreaktivität reduziert werden konnten. Durch das gewählte Erhebungsdesign konnte ausgeschlossen werden, dass potenzielle Verhaltensänderungen auf die Einführung eines neuen Messinstruments zurückzuführen sind. Stattdessen erlaubte das Studiendesign einen direkten Vergleich des Ernährungsverhaltens vor und während der Studienteilnahme innerhalb derselben Personen. Auf diese Weise konnten Effekte der Forscherbeobachtung gezielter von Veränderungen getrennt werden, die durch die erstmalige Nutzung eines Erhebungsinstruments entstehen könnten (König et al., 2025).

Darüber hinaus trägt die Studie dazu bei, eine bestehende Forschungslücke zu adressieren, indem Messreaktivität im Bereich des Ernährungsverhaltens untersucht wird. Während sich ein Großteil der bisherigen Forschung auf körperliche Aktivität konzentriert, liegen für die Nahrungsaufnahme bislang vergleichsweise wenige Befunde vor (König et al., 2022a). Die vorliegenden Ergebnisse leisten somit einen Beitrag zum Verständnis der Validität digitaler

Ernährungserfassung und erweitern den bisherigen Forschungsstand um ein bislang wenig berücksichtigtes Verhalten.

Ein weiterer methodischer Vorteil liegt im Einbezug potenzieller Moderatorvariablen. Durch die Berücksichtigung von Geschlecht und Gewissenhaftigkeit konnten interindividuelle Unterschiede in möglichen reaktiven Effekten untersucht werden. Dies ermöglicht eine differenziertere Betrachtung von Messreaktivität und liefert Hinweise darauf, unter welchen Bedingungen und für welche Personengruppen entsprechende Effekte auftreten könnten. Die Einbeziehung solcher Moderatorvariablen ist insbesondere im Kontext von Messreaktivität von Bedeutung, da dieses Phänomen als stark kontextabhängig beschrieben wird (French & Sutton, 2010; French et al., 2021; König et al., 2022a). Reaktive Effekte treten nicht einheitlich auf, sondern variieren in Abhängigkeit individueller und situativer Bedingungen. Die Berücksichtigung von Moderatoren ermöglicht es daher, diese Unterschiede systematisch zu erfassen und die Befunde differenzierter einzuordnen.

Hinsichtlich der statistischen Analyse weist die Studie ebenfalls Stärken auf. Durch den Einsatz von Mehrebenenmodellen konnten Unterschiede zwischen Personen als auch intraindividuelle Veränderungen im Zeitverlauf berücksichtigt werden. Diese Modellierung ist der hierarchischen Datenstruktur angemessen und ermöglicht eine präzisere Schätzung der Effekte. Insgesamt erlaubt die Kombination aus alltagsnaher Datenerhebung, innerhalb-Personen-Vergleich und differenzierter statistischer Analyse eine fundierte Untersuchung von Messreaktivität.

Limitationen

Den dargestellten Stärken der vorliegenden Studie stehen einige Einschränkungen gegenüber, die bei der Interpretation der Ergebnisse berücksichtigt werden sollten. Eine zentrale Limitation besteht darin, dass die beobachteten Befunde nicht eindeutig auf einzelne Aspekte des Studiendesigns zurückgeführt werden können. Im Vergleich zu Studien, in denen Messreaktivität nachgewiesen wurde (French & Sutton, 2010; French et al., 2021; König et al., 2022a), unterscheidet sich die vorliegende Untersuchung in mehreren methodischen Merkmalen gleichzeitig. Diese Mehrdeutigkeit erschwert die Einordnung der Ergebnisse im Kontext der bestehenden Forschung.

So wurde kein neues Messinstrument eingeführt, sondern auf Datenspenden aus bereits genutzten Ernährungstracking-Apps zurückgegriffen. Während in vielen Studien reaktive Effekte zumindest teilweise auf die erstmalige Nutzung eines Erhebungsinstruments zurückgeführt werden können, entfällt dieser Einfluss im vorliegenden Design. Zudem erfolgte eine stärkere Trennung zwischen Mess- und Forschungskontext, da das Ernährungsverhalten bereits vor Studienbeginn erfasst wurde. Veränderungen können somit gezielter der Forscherbeobachtung zugeordnet werden, anstatt mit Effekten der Messung selbst konfundiert zu sein.

Gleichzeitig fand die Datenerhebung unter alltagsnahen Bedingungen statt, ohne direkte Interaktion mit den Forschenden und bei vergleichsweise geringer Salienz der Erhebungssituation. Diese Rahmenbedingungen unterscheiden sich deutlich von stärker kontrollierten oder experimentellen Settings und könnten das Ausmaß messreaktiver Effekte reduziert haben. Auch wenn diese methodischen Besonderheiten einerseits als Stärke zu bewerten sind, erschweren die damit verbundenen Unterschiede zur bestehenden Forschungslage die Vergleichbarkeit der Ergebnisse. In der Folge bleibt unklar, welche dieser Faktoren maßgeblich dazu beigetragen haben, dass keine messreaktiven Effekte festgestellt wurden.

Ein weiterer Unterschied betrifft die Zusammensetzung der Stichprobe, die ausschließlich aus Personen bestand, die bereits Erfahrung mit digitalem Ernährungstracking hatten. Im Gegensatz zu Studien mit unerfahrenen Teilnehmenden ist davon auszugehen, dass Selbstmonitoring-Prozesse in dieser Gruppe bereits etabliert waren. Zudem beschränkte sich die Untersuchung auf einen sehr kurzen Beobachtungszeitraum von zwei Tagen, wodurch insbesondere längerfristige Anpassungsprozesse möglicherweise nicht erfasst werden konnten. Aufgrund dieser kumulativen Unterschiede ist die Vergleichbarkeit mit bisherigen Studien nur eingeschränkt gegeben und es bleibt daher offen, in welchem Ausmaß das Ausbleiben messreaktiver Effekte auf spezifische Merkmale des Studiendesigns zurückzuführen ist.

Darüber hinaus ist die Generalisierbarkeit der Befunde durch die Zusammensetzung der Stichprobe eingeschränkt. Untersucht wurden ausschließlich Personen, die ihre Ernährung bereits vor Studienbeginn regelmäßig mithilfe von Apps erfassten. Es ist daher anzunehmen, dass es sich um eine spezifische Teilpopulation mit erhöhter Selbstbeobachtung und

möglicherweise höherem Gesundheitsbewusstsein handelt, deren Ernährungsverhalten zudem durch das aktive Protokollieren selbst beeinflusst sein kann (Thompson & Subar, 2017). Die Übertragbarkeit der Ergebnisse auf Personen ohne Erfahrung im Ernährungstracking oder auf die Allgemeinbevölkerung ist entsprechend begrenzt.

Auch die Stichprobengröße stellt eine Einschränkung dar. Mit 118 Teilnehmenden wurde die angestrebte Zielgröße von 200 Personen nicht erreicht. Vor diesem Hintergrund ist nicht auszuschließen, dass kleinere, aber potenziell relevante Effekte aufgrund begrenzter statistischer Power unentdeckt geblieben sind. Dies gilt insbesondere für Interaktionseffekte, deren Nachweis in der Regel größere Stichproben erfordert.

Eine weitere Limitation betrifft die Verlässlichkeit der erhobenen Ernährungsdaten. Da kein direkter Zugriff auf die Rohdaten der verwendeten Apps bestand, musste davon ausgegangen werden, dass die Teilnehmenden ihre Einträge zeitnah dokumentiert und korrekt in die Befragung übertragen haben. In welchem Ausmaß dies tatsächlich der Fall war, lässt sich jedoch nicht überprüfen. Es ist denkbar, dass fehlende Angaben nachträglich geschätzt wurden oder dass Übertragungsfehler zwischen App und Fragebogen aufgetreten sind. Entsprechend kann nicht ausgeschlossen werden, dass die Daten durch Erinnerungs- oder Eingabefehler verzerrt sind.

Eine methodische Einschränkung betrifft die Analyse von Moderatorvariablen. Damit Moderatoreffekte sinnvoll interpretiert werden können, müssten überhaupt Unterschiede zwischen den Personen in ihrer Reaktion auf die Messsituation vorliegen. Wenn sich die Teilnehmenden in ihrem Ausmaß an Messreaktivität kaum unterscheiden, können Moderatoren diese Unterschiede auch nicht erklären. Dadurch ist die Aussagekraft der Moderationsanalysen eingeschränkt, unabhängig davon, ob die entsprechenden Effekte statistisch signifikant ausfielen oder nicht (Leon & Heo, 2009).

Schließlich ist auch die Operationalisierung der Gewissenhaftigkeit als Limitation zu berücksichtigen. Die verwendete Kurzsкала basierte auf lediglich zwei Items und wies eine geringe Inter-Item-Korrelation auf, was auf eine eingeschränkte Reliabilität hindeutet. Eine reduzierte Messpräzision kann insbesondere bei der Untersuchung von Interaktionseffekten dazu führen, dass bestehende Zusammenhänge unterschätzt oder nicht erkannt werden.

Implikationen

Die vorliegenden Befunde haben sowohl methodische als auch praktische Implikationen für die Forschung zur Messreaktivität und für den Einsatz digitaler Ernährungserfassung. Zunächst legen die Ergebnisse nahe, dass digitale Ernährungstracking-Apps unter bestimmten Bedingungen nicht zwangsläufig mit ausgeprägten kurzfristigen reaktiven Verhaltensänderungen einhergehen. Dies ist insbesondere für die Interpretation von Daten aus solchen Anwendungen relevant, da es darauf hindeutet, dass die Validität der erhobenen Daten in Populationen mit bereits etabliertem Selbstmonitoring möglicherweise weniger stark durch Messreaktivität beeinträchtigt ist.

Gleichzeitig verdeutlichen die Befunde, dass Messreaktivität im Kontext digitaler Ernährungserfassung stark von spezifischen Rahmenbedingungen abhängt. Insbesondere die habitualisierte Nutzung von Tracking-Apps sowie die geringe Salienz der Erhebungssituation scheinen potenzielle reaktive Effekte abzuschwächen. Für zukünftige Studien ergibt sich daraus die Notwendigkeit, diese Kontextfaktoren systematisch zu untersuchen und Designs gezielt darauf auszurichten, unterschiedliche Bedingungen der Messung zu variieren. Hierzu zählen beispielsweise Vergleiche zwischen erfahrenen und unerfahrenen Nutzenden, variierende Beobachtungszeiträume oder gezielte Manipulationen der Salienz der Forscherbeobachtung.

Darüber hinaus unterstreichen die Ergebnisse die Bedeutung einer differenzierten Betrachtung von Messreaktivität über verschiedene Verhaltensbereiche hinweg. Die eingeschränkte Übertragbarkeit von Befunden aus dem Bereich der körperlichen Aktivität auf das Ernährungsverhalten legt nahe, dass verhaltensspezifische Eigenschaften stärker in den Fokus zukünftiger Forschung rücken sollten. Die Verankerung des Ernährungsverhaltens in stabilen Alltagsroutinen könnte dazu beitragen, dass reaktive Effekte weniger kurzfristig und weniger ausgeprägt auftreten als in anderen Gesundheitsverhaltensdomänen.

Für die praktische Anwendung digitaler Ernährungstracking-Apps ergibt sich daraus, dass diese nicht nur als Messinstrumente, sondern auch als potenziell verhaltensverändernde Tools differenziert betrachtet werden müssen. Während in bestimmten Kontexten eine reaktive Wirkung im Sinne einer Verhaltensverbesserung erwünscht sein kann, deuten die vorliegenden Ergebnisse darauf hin, dass eine solche Wirkung bei bereits regelmäßiger Nutzung nicht automatisch eintritt. Dies hat Implikationen für die Entwicklung und den Einsatz

digitaler Gesundheitsanwendungen, da zusätzliche Interventionselemente erforderlich sein könnten, um gezielte Verhaltensänderungen zu fördern.

Insgesamt tragen die Ergebnisse dazu bei, ein differenzierteres Verständnis der Bedingungen zu entwickeln, unter denen Messreaktivität im Kontext digitaler Ernährungserfassung auftritt, und liefern Ansatzpunkte für die Weiterentwicklung zukünftiger Studien sowie für den Einsatz Ernährungstracking-Apps.

Conclusio

Die vorliegende Arbeit untersuchte das Auftreten von Messreaktivität im Kontext digitaler Ernährungserfassung und leistet damit einen Beitrag zu einem bislang wenig erforschten Bereich der gesundheitspsychologischen Forschung. Entgegen der theoretischen Annahme zeigte sich kein signifikanter kurzfristiger Effekt der Forscherbeobachtung auf den Fettkonsum. Ebenso ergaben sich keine Hinweise darauf, dass Geschlecht oder Gewissenhaftigkeit den Zusammenhang zwischen Messsituation und Ernährungsverhalten moderieren.

Die Ergebnisse legen nahe, dass Messreaktivität im Kontext digitaler Ernährungstracking-Apps nicht als universell auftretendes Phänomen verstanden werden kann, sondern in hohem Maße von spezifischen Rahmenbedingungen abhängt. Vor allem die bereits etablierte Nutzung von Tracking-Apps sowie die geringe Salienz der Erhebungssituation könnten dazu beigetragen haben, dass reaktive Effekte in der untersuchten Stichprobe ausblieben.

Die vorliegenden Befunde unterstreichen die Notwendigkeit, Messreaktivität differenziert zu betrachten. Gleichzeitig liefern sie erste Hinweise darauf, dass digitale Ernährungserfassung unter alltagsnahen Bedingungen valide Verhaltensdaten liefern kann, ohne zwangsläufig mit systematischen kurzfristigen Verhaltensveränderungen einherzugehen. Weitere Forschung ist erforderlich, um die Bedingungen, unter denen Messreaktivität im Ernährungsbereich auftritt, präziser zu bestimmen und die Übertragbarkeit der Befunde auf andere Populationen und Erhebungskontexte zu prüfen.

Literaturverzeichnis

- Afshin, A., Sur, P. J., Fay, K. A., Cornaby, L., Ferrara, G., Salama, J. S., Mullany, E. C., Abate, K. H., Abbafati, C., Abebe, Z., Afarideh, M., Aggarwal, A., Agrawal, S., Akinyemiju, T., Alahdab, F., Bacha, U., Bachman, V. F., Badali, H., Badawi, A., . . . Murray, C. J. L. (2019). Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, 393(10184), 1958–1972. [https://doi.org/10.1016/s0140-6736\(19\)30041-8](https://doi.org/10.1016/s0140-6736(19)30041-8)
- 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 and Health*, 39(3), 319–335. <https://doi.org/10.1080/08870446.2022.2055024>
- Barta, W. D., Tennen, H., & Litt, M. D. (2012). *Measurement Reactivity in Diary Research*. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 108–123). The Guilford Press.
- 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>
- 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(4), 797–809. <https://doi.org/10.1007/s10551-007-9548-4>
- Bravata, D. M., Smith-Spangler, C., Sundaram, V., Gienger, A. L., Lin, N., Lewis, R., Stave, C. D., Olkin, I., & Sirard, J. R. (2007). Using pedometers to increase physical activity and improve health: A systematic review. *JAMA*, 298(19), 2296–2304. <https://doi.org/10.1001/jama.298.19.2296>
- Bundesministerium für Gesundheit und Frauen. (2017). *Österreichischer Ernährungsbericht 2017*. <https://www.bmgf.gv.at/>
- Clark, M., Southerton, C., & Driller, M. (2024). Digital self-tracking, habits and the myth of discontinuance: It doesn't just 'stop'. *New Media & Society*, 26(4), 2168–2188. <https://doi.org/10.1177/14614448221083992>
- Costa, P. T., & McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. *Journal of Personality Disorders*, 6(4), 343–359. <https://doi.org/10.1521/pedi.1992.6.4.343>
- Darker, C. D., & French, D. P. (2009). What sense do people make of a theory of planned behaviour questionnaire?: A think-aloud study. *Journal of Health Psychology*, 14(7), 861–871. <https://doi.org/10.1177/1359105309340983>
- Deutsche Gesellschaft für Ernährung e. V. (2015). *Evidenzbasierte Leitlinie: Fettzufuhr und Prävention ausgewählter ernährungsmitbedingter Krankheiten* (2. Version). <https://www.dge.de/>
- Edwards, P., Roberts, I., Clarke, M., DiGuseppi, C., Pratap, S., Reinhard, W., & Kwan, I. (2002). Increasing response rates to postal questionnaires: Systematic review. *British Medical Journal*, 324(7347), 1183–1191. <https://doi.org/10.1136/bmj.324.7347.1183>

- Fekete, C., & Weyers, S. (2016). Soziale Ungleichheit im Ernährungsverhalten. *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz*, 59(2), 197–205. <https://doi.org/10.1007/s00103-015-2279-2>
- Förg, H. (2026). *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. (2021). Reducing bias in trials from reactions to measurement: The MERIT study including developmental work and expert workshop. *Health Technology Assessment*, 25(55). <https://doi.org/10.3310/hta25550>
- Goldberg, L. R., & Strycker, L. A. (2002). Personality traits and eating habits: The assessment of food preferences in a large community sample. *Personality and Individual Differences*, 32(1), 49–65. [https://doi.org/10.1016/S0191-8869\(01\)00005-8](https://doi.org/10.1016/S0191-8869(01)00005-8)
- Gunning, J. A., Converse, M. F., Gudarzi, B., Lotfallah, W., & Racette, S. B. (2025). Dietary Patterns Influence Chronic Disease Risk and Health Outcomes in Older Adults: A Narrative Review. *Nutrients*, 17(24), 3910. <https://doi.org/10.3390/nu17243910>
- Ho, V., Simmons, R. K., Ridgway, C. L., van Sluijs, E. M. F., 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>
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60(6), 581–592. <https://doi.org/10.1037/0003-066X.60.6.581>
- Illner, A.-K., Freisling, H., Boeing, H., Huybrechts, I., Crispim, S. P., & Slimani, N. (2012). Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. *International Journal of Epidemiology*, 41(4), 1187–1203. <https://doi.org/10.1093/ije/dys105>
- 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>
- 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., van Emmenis, M., Nurmi, J., Kassavou, K., & Sutton, S. (2022b). Characteristics of smartphone-based dietary assessment: A systematic review. *Health Psychology Review*, 16(4), 526–550. <https://doi.org/10.1080/17437199.2021.2016066>
- König, L. M., Pasko, K., Baga, K., Harsora, R., & Arigo, D. (2025). Isolating the role of researcher observation on reactivity to the measurement of physical activity. *Applied Psychology: Health and Well-Being*, 17(1), e12630. <https://doi.org/10.1111/aphw.12630>
- König, L. M., 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>
- Lackner, S. (2025). *Messreaktivität bei der digitalen Ernährungserfassung: Verändert Forscherbeobachtung den alltäglichen Proteinkonsum?* (Masterarbeit). Universität Wien.
- Leon, A. C., & Heo, M. (2009). Sample sizes required to detect interactions between two binary fixed-effects in a mixed-effects linear regression model. *Computational statistics & data analysis*, 53(3), 603–608. <https://doi.org/10.1016/j.csda.2008.06.010>
- Long, P. A., Huberts, A. S., di Torrero, A. N., Otto, L. R., Rogge, A. A., Ritschl, V., & Stamm, T. A. (2025). The mere-measurement effect of patient-reported outcomes: A systematic review and meta-analysis. *Quality of Life Research*, 34(5), 1211–1220. <https://doi.org/10.1007/s11136-025-03909-y>
- Mayo, E. (2003). *The human problems of an industrial civilization*. Routledge. (Original work published 1933)
- McCambridge, J., Witton, J., & Elbourne, D. R. (2014). Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. *Journal of Clinical Epidemiology*, 67(3), 267–277. <https://doi.org/10.1016/j.jclinepi.2013.08.015>
- Michie, S., Abraham, C., Whittington, C., McAteer, J., & Gupta, S. (2009). Effective techniques in healthy eating and physical activity interventions: A meta-regression. *Health Psychology*, 28(6), 690–701. <https://doi.org/10.1037/a0016136>
- Miles, L. M., Rodrigues, A. M., Sniehotta, F. F., & French, D. P. (2020). Asking questions changes health-related behavior: An updated systematic review and meta-analysis. *Journal of Clinical Epidemiology*, 123, 59–68. <https://doi.org/10.1016/j.jclinepi.2020.03.014>
- Naska, A., Lagiou, A. & Lagiou, 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>
- 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>
- Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41(1), 203–212. <https://doi.org/10.1016/j.jrp.2006.02.001>

- R Core Team. (2025). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology*, 58(1), 103–139. <https://doi.org/10.1111/j.1744-6570.2005.00301.x>
- Robinson, E., Hardman, C. A., Halford, J. C. G., & 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>
- Rollo, M. E., Williams, R. L., Burrows, T., Kirkpatrick, S. I., Bucher, T., & Collins, C. E. (2016). What are they really eating? A review on new approaches to dietary intake assessment and validation. *Current Nutrition Reports*, 5(4), 307–314. <https://doi.org/10.1007/s13668-016-0182-6>
- Sanderman, R., & Morgan, K. (Eds.). (2025). *The Routledge International Handbook of Health Psychology: Global and Contemporary Issues*. Routledge. <https://doi.org/10.4324/9781003318026>
- Sharpe, J. P., & Gilbert, D. G. (1998). Effects of repeated administration of the Beck Depression Inventory and other measures of negative mood states. *Personality and Individual Differences*, 24(4), 457–463. [https://doi.org/10.1016/S0191-8869\(97\)00219-7](https://doi.org/10.1016/S0191-8869(97)00219-7)
- 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>
- Thompson, F. E., & Subar, A. F. (2017). Dietary assessment methodology. In A. M. Coulston, C. J. Boushey, & M. G. Ferruzzi (Eds.), *Nutrition in the prevention and treatment of disease* (4th ed., pp. 5–48). Academic Press. <https://doi.org/10.1016/B978-0-12-802928-2.00001-1>
- Vecchione, M., Alessandri, G., Barbaranelli, C., & Caprara, G. (2012). Gender differences in the Big Five personality development: A longitudinal investigation from late adolescence to emerging adulthood. *Personality and Individual Differences*, 53(6), 740–746. <https://doi.org/10.1016/j.paid.2012.05.033>
- Weisberg, Y. J., DeYoung, C. G., & Hirsh, J. B. (2011). Gender differences in personality across the ten aspects of the Big Five. *Frontiers in Psychology*, 2, 178. <https://doi.org/10.3389/fpsyg.2011.00178>
- Zhang, Y. D., Zhang, H. L., Xie, J. Q., & Zhang, C. B. (2023). The influence of self-quantification on individual's participation performance and behavioral decision-making in physical fitness activities. *Humanities and Social Sciences Communications*, 10(1), 592. <https://doi.org/10.1057/s41599-023-02103-0>

Tabellenverzeichnis

Tabelle 1	18
Tabelle 2	24
Tabelle 3	25
Tabelle 4	25
Tabelle 5	26
Tabelle 6	27
Tabelle 7	27