IE UNIVERSIDAD

TESIS DOCTORAL/ DOCTORAL DISSERTATION

Intervenciones para reducir el consumo digital problemático / Interventions to Reduce Problematic Digital Consumption

Quang Duc Pham

SEGOVIA, 2024

IE UNIVERSIDAD

TESIS DOCTORAL/ DOCTORAL DISSERTATION

Intervenciones para reducir el consumo digital problemático / Interventions to Reduce Problematic Digital Consumption

Quang Duc Pham

Doctoral Thesis Advisor: Dr. Antonios Stamatogiannakis

Doctoral Thesis Co-advisor: Dr. Jeeva Somasundaram

Abstract

This dissertation examines interventions for reducing problematic digital consumption, specifically, smartphone and social media usage, and methods for enhancing interventions' effectiveness. Chapters 1 and 2 present a dual-target intervention designed to sustain behavior change following the discontinuation of monetary incentives. By incentivizing the reduction of undesirable consumption patterns while promoting beneficial alternative activities, this approach was tested in two field studies targeting social media and smartphone usage. Results indicate that participants who met set targets consistently during treatment exhibited lower social media / smartphone usage compared to those with a single consumption reduction target. This finding offers policymakers a tool to enhance the long-term impact of behavioral interventions against digital addiction and holds potential for the application of habit formation in other domains that focus on behavior change. Additionally, it presents opportunities for social marketers to introduce products aiding in addictive behavior management.

Chapter 3 compares the effectiveness of existing interventions targeting problematic smartphone and social media use. Through a grounded framework, causal evidence from intervention studies is synthesized using multivariate, multilevel meta-analytical models. The findings indicate that interventions had a small-to-medium effect on reducing consumption and improving consequences (such as depression, anxiety, or sleep quality). The analysis revealed significant heterogeneities in effect sizes, suggesting that factors such as age and gender may influence intervention design in meaningful ways. This research synthesis provides valuable guidance for future interventions, emphasizing the importance of considering diverse demographics and study factors in intervention planning and implementation.

Resumen

Esta tesis examina las intervenciones para reducir el consumo digital problemático, en concreto, el uso de teléfonos inteligentes y redes sociales, y los métodos para mejorar la eficacia de las intervenciones. Los capítulos 1 y 2 presentan una intervención de doble objetivo diseñada para mantener el cambio de comportamiento después de la interrupción de los incentivos monetarios. Al incentivar la reducción de los patrones de consumo no deseados, mientras que la promoción de actividades alternativas beneficiosas, este enfoque se puso a prueba en dos estudios de campo dirigidos a los medios de comunicación social y el uso de teléfonos inteligentes. Los resultados indican que los participantes que cumplieron los objetivos fijados de forma sistemática durante el tratamiento hicieron un menor uso de las redes sociales y los teléfonos inteligentes en comparación con los que se fijaron un único objetivo de reducción del consumo. Este hallazgo ofrece a los responsables políticos una herramienta para mejorar el impacto a largo plazo de las intervenciones conductuales contra la adicción digital y tiene potencial para la aplicación de la formación de hábitos en otros dominios que se centran en el cambio de comportamiento. Además, presenta oportunidades para que los vendedores sociales introduzcan productos que ayuden en la gestión del comportamiento adictivo.

El capítulo 3 compara la eficacia de las intervenciones existentes dirigidas al uso problemático de teléfonos inteligentes y redes sociales. A través de un marco fundamentado, se sintetizan las pruebas causales de los estudios de intervención mediante modelos metaanalíticos multinivel y multivariantes. Los resultados indican que las intervenciones tuvieron un efecto entre pequeño y medio en la reducción del consumo y la mejora de las consecuencias (como la depresión, la ansiedad o la calidad del sueño). El análisis reveló heterogeneidades significativas en los tamaños del efecto, lo que sugiere que factores como la edad y el sexo pueden influir en el diseño de la intervención de manera significativa. Esta síntesis de la investigación proporciona una valiosa orientación para futuras intervenciones, haciendo hincapié en la importancia de considerar diversos factores demográficos y de estudio en la planificación y ejecución de la intervención.

This page is intentionally left blank.

Acknowledgement

I would like to thank, first and foremost, my two advisors, Antonios and Jeeva, for their unwavering support throughout my Ph.D. journey. Your advice and words of encouragement mean the world to me.

To Dr. Laura Zimmermann, thank you for always bringing a positive perspective to things.

I would like to express my sincerest gratitude to Dr. Simona Botti and Dr. Sanchayan Banerjee for their valuable comments, which have been instrumental in the completion of this dissertation.

I am grateful to the Doctoral Program coordinators at IE for their prompt support over the past four-plus years.

I dedicate this dissertation to my two most important people in the world, Dad and Mom. Thank you, Dad, for always believing in me. Thank you, Mom, for always silently supporting me.

Digital Addiction Interventions: The Role of Alternative Activities in Reducing Smartphone and Social Media Consumption

Table of Contents

Introduction *(English version)*

Excessive smartphone and social media usage has become a growing concern for many consumers and policymakers across the world. Recent research has shown that there is a negative link between smartphone screen time and various well-being indicators (e.g., Twenge & Campbell, 2019; Hisler et al., 2020, Liu et al., 2020). The average consumer spends around four hours per day looking at their smartphone screen (App Annie, 2021). Assuming the average number of waking hours is 17, this amount of smartphone usage would account for almost a quarter of a person's day; within a year, the cumulative usage would be about 1,460 hours, or roughly 60 full-days. While many consumers nowadays have the goal of reducing their mobile screen time, this goal is seldom achieved, arguably because this type of digital consumption is a highly habitual behavior (Anderson & Wood, 2021).

Habits are notoriously hard to change, and consumers typically struggle to achieve personal goals when these require changing a habitual behavior pattern (Wood & Neal, 2009). For example, consumers wanting to lose weight often fail to reduce their daily calorie intake. Those who want to reduce time spent on social media apps still find themselves scrolling mindlessly through these platforms. Due to the force of habit, these behaviors are difficult to change and often result in failures to achieve one's consumption goals (Wood & Neal, 2016). How can this problem be addressed?

While previous research has predominately concentrated on tackling a habitual behavior in isolation, I draw on the idea that changing a habit requires a comprehensive approach, as it is inextricably linked to other daily activities (Marlatt & George, 1984; Witkietvitz & Kirouac, 2016). When consumers try to reduce an undesired habit, especially one that occupies a significant amount of time in their daily routine, they are confronted with having more "idle time". Building on previous research (Kushlev & Leitao, 2020; Adriannse et al., 2011; Laibson,

2001) I argue that if this time is occupied by an alternate activity, there is a lower likelihood of the old habit relapsing. Thus, engaging in an alternative activity, ideally one that is beneficial, might be helpful for achieving lasting behavior change. This constitutes the main research questions that the first two chapters address: How effective are interventions that reduce an undesired habitual behavior while building a beneficial alternative activity, in the long run? Is it more effective than just reducing the undesired habitual behavior in isolation?

To answer these questions, I designed an intervention to help consumers reduce undesired habits while at the same time encouraging the uptake of new beneficial activities in their daily routines. Chapter 1 and 2 of the current research detail two field studies, in which consumers who are monetarily incentivized to reduce an undesired habit while at the same time engaging more in a beneficial activity, are more successful in sustaining their behavior change, than consumers who are only incentivized to reduce the undesired habit. I investigate the persistence of behavior change in the short term (immediately after the intervention) and in the long term (40 days after the intervention ended).

I chose social media usage in Chapter 1, and overall smartphone usage in Chapter 2, as the focal behaviors. Not only can these behaviors be measured objectively with screen time applications, but they are also highly habitual, which makes them prime candidates for testing this intervention. Importantly, many consumers are concerned about their digital media consumption and, importantly, would like to reduce it (Anderson & Wood, 2021; Deloitte, 2019).

As alternative activities, I chose learning a new language (Chapter 1) and engaging in physical activity (Chapter 2). Like smartphone usage, both activities can be measured objectively with tracking applications. In addition, they are considered beneficial due to their positive impact on health, cognitive function, and emotional well-being (e.g., Lavie et al., 2019; Nieman & Wentz, 2019, Zimmermann & Chakravarti, 2022; De la Rosa et al., 2020; Egger et al., 2019, Chekroud et al., 2018; Kvam et al., 2016). Importantly, both are commonly reported as activities that consumers would like to engage in more (Hoare et al., 2017; Young, 2014).

Results show that integrating a positive alternative activity (i.e., learning a new language or increasing physical activity) into the daily routine while reducing an undesired behavior (i.e., excessive phone or social media usage) can reduce the risk of relapse. This finding applies to individuals who more consistently performed the positive activity while reducing the undesired behavior. By tying two distinct behaviors to a single monetary incentive (hereafter known as the 'dual-target' approach), we can mitigate the risk of relapse, offering a more cost-effective intervention compared to traditional interventions that focus solely on one behavior. Although it may appear counterintuitive since achieving two behavioral targets for one incentive is more challenging, I demonstrate that the higher effort is worthwhile. Successful adherence to both targets, rather than only one target, improves the chance of achieving persistent behavior change in the long run.

Together, Chapter 1 and 2 have several contributions. First, my approach builds on and extends previous behavior change interventions which tackle one behavior at a time as well as research investigating compensatory addictions (i.e., substituting an undesired habit with other similar and negative ones; Dai & Fishbach, 2014; Dey et al., 2019; Melumad & Pham, 2020). My approach is different from temptation bundling where vices are tied to virtues (e.g., watching your favorite series while exercising at the gym) to support habit formation (Milkman et al., 2014). Instead, I am encouraging consumers to use the time gained by reducing a vice for engaging in a virtuous alternative. Thus, both activities cannot be performed together.

Furthermore, the dual-target intervention is related to, but distinct from, behavioral spillover (e.g., Dolan & Galizzi, 2015; Galizzi & Whitmarsh, 2019). A spillover means that the 'bad' primary behavior (e.g., social media usage) would be unintentionally reduced by just increasing the 'good' alternative activities (e.g. Duolingo usage). Some previous RCTs are directly related to this phenomenon (e.g., Brailovskaia et al., 2022), where they either (1) asked participants to do the alternative behaviors, and measured both target and alternative behaviors, or (2) suggested participants to do both behaviors, but no incentives are provided. During the dual-target intervention, both good and bad behaviors are actively targeted and incentivized. Thus, the current research occupies a related, but different area where the reduction of one and the building up of another are a lot more deliberate and, for lack of a better word, calculated[1].

Previous marketing research has shown that companies are motivated to keep consumers constantly and promptly engaged through social media, because it enhances consumer-brand relationships (Hudson et al., 2016; Dhaoui & Webster, 2021) and enlists consumers as contributors to brand's success (Saboo et al., 2016). However, scholars have recently started emphasizing the importance of sustainable and moderate consumption in marketing for fostering long-term brand loyalty, improving corporate reputation, and addressing societal concerns (White et al., 2019; Cornil et al., 2022; Somasundaram et al., 2023). Curbing excessive consumption can even be beneficial for firms offering digital products. As Nevskaya and Albuquerque (2019) show, curbing excessive gaming time benefits firms' profit with longer subscriptions.

Thus, the current work offers a practical intervention for reducing excessive consumption habits and presents new marketing avenues for businesses to capitalize on by offering beneficial alternatives. These alternatives serve to fill the gap of needs left behind by the reduction of the undesired consumption. For example, while reducing undesired technology consumption, consumers have a need to occupy their newfound time (i.e. the need gap), and

[1] Because behavioral spillover is a broad phenomenon that includes highly heterogeneous effects and mechanisms (Galizzi & Whitmarsh, 2019), I will discuss how it can be incorporated into extending the current research in dual-target intervention in the General Discussion of Chapter 1 and 2.

thus might seek alternative products and services. Marketers can leverage this gap to introduce new alternatives that not only align with their financial goals but also address the growing demand for social responsibility, in line with social marketing (Kotler & Zaltman, 1971; Lee & Kotler, 2015).

Several recent studies have employed monetary incentives to reduce smartphone usage (Stanley et al., 2022; Collis & Eggers, 2022; Allcott et al., 2022; Allcott et al., 2020). Similarly, some studies have investigated how recommending an alternative activity affects behavior change (Reed et al., 2023; Brailovskaia et al., 2022; Esmaeili & Ahmadi, 2018). The current research adds to this body of literature as it combines both approaches to prevent relapse (see Appendix A1 for a detailed comparison of this research with the relevant literature). Rather than relying on consumers' self-initiative to engage in an alternative activity, I set concrete, actionable targets for both reducing an undesired habit and increasing an alternative activity, and tie both targets to a single incentive to encourage behavior change on a daily basis which is crucial for habit formation (Somasundaram et al., 2023; Somasundaram et al., 2024). Finally, I overcome a methodological drawback by measuring all behaviors objectively.

Recent research has shown that there is a negative link between mobile screen time and academic performance (Giunchiglia et al. 2018), performance at work (Liu et al., 2020) as well as well-being (Twenge & Campbell 2019) and sleep (Hisler et al., 2020). Chapter 1 and 2 provide a timely response to the emerging need for an effective solution to reduce smartphone/social media usage. However, Chapter 1 and 2's scope is limited to only the class of interventions that promote behavioral change using incentives and/or alternative activities. There exists a diverse range of other approaches to tackle problematic smartphone and social media consumption. Therefore, I was inspired to explore other intervention types that aim to alleviate this problem in Chapter 3, by means of an exhaustive literature search and meta-analysis. Thus, Chapter 3 seeks to address the following research questions: How effective are interventions at reducing smartphone / social media usage and improving wellbeing? Are there differences between intervention types? How do demographic factors (such as age and gender) and study characteristics (such as study design and intervention length) affect the effectiveness of these interventions?

In Chapter 3, I identify what strategies have been evaluated for their effectiveness in reducing problematic usage, and how they measure up against one another in two types of outcomes: consumption (i.e., smartphone screen time, social media usage time), and consequence (i.e., addiction, well-being, ill-being, sleep quality, productivity) improvement. Using a grounded approach, I classify interventions into three broad types: *Limit*, *Limit Plus*, and *Non-Limit*, so that they can be meaningfully compared in meta-analytical models. I expand the scope of my analyses by examining the influence of relevant study characteristics (e.g., intervention design, treatment duration) and population characteristics (e.g., age, gender) on interventions' effect sizes.

My analysis in Chapter 3 demonstrates that, on average, interventions have a medium effect on reducing consumption of social media and smartphone, and improve consequence outcomes by a small-to-medium amount. Additionally, on average, there is no difference in effect sizes among the three types of interventions. However, depending on the main goal (either consumption reduction or consequence improvement) of the intervention type, Non-Limit interventions work better on consequence outcomes, while Limit and Limit Plus interventions help reduce consumption more. I also gained insights into how various factors may influence the effectiveness of interventions. For example, while younger people benefit more from interventions with well-being and ill-being improvements, older people experience higher reduction in consumption of social media and smartphone usage through interventions. Most notably, women seem to benefit less from interventions than men, as evidenced by effect sizes ranging from 2 to 3 times smaller than those of males, for both consumption and consequence outcomes.

With this set of findings, Chapter 3 makes the following contributions. First, it provides a framework that broadly characterizes interventions based on whether and how they impose direct restrictions to consumption of mobile technology consumption. This allows me to compare effect sizes across broadly different intervention categories, which expands the scope of research synthesis of causal evidence, in contrast to past studies which have only investigated a narrow set of interventions. Second, this framework acknowledges the heterogeneity across intervention studies, a concern that was raised by past research in this domain (Valkenburg, 2022), by incorporating multiple outcome, population, and study characteristics. Statistically, I address this issue by using multi-level, multivariate, random-effects meta-analytical models, which account for potential confounds that might bias average effect size estimates yielded by previous meta-analyses using conventional two-level and three-level intercept-only models for estimation. This constitutes a major methodological contribution of the current work to the literature. Third, my literature search results in a much larger number of studies compared to the previous meta-analyses and systematic review which have summarized causal evidence in this domain. Thus, my analyses provide a more comprehensive evaluation of available evidence of interventions' effectiveness. Finally, based on the findings of this chapter, I give specific, actionable advice that can guide the design of future interventions, which may interest policymakers and researchers. Together with the meta-analyses, the compilation of interventions that have been tested (Appendix D3) may also serve as a useful reference source for other stakeholders, such as parents of young consumers who are looking for evidence-based approaches to help guide their consumption of social media and smartphone use. This is especially relevant in light of the recent rising number of concerned parents in the UK and elsewhere who have congregated into special interest groups and called for a blanket-ban of social media for children under 16 (Banfield-Nwachi, 2024).

Introducción *(versión en español)*

El uso excesivo de smartphones y redes sociales se ha convertido en una preocupación creciente para muchos consumidores y responsables políticos de todo el mundo. Investigaciones recientes han demostrado que existe una relación negativa entre el tiempo de pantalla del teléfono inteligente y diversos indicadores de bienestar (por ejemplo, Twenge y Campbell, 2019; Hisler et al., 2020, Liu et al., 2020). El consumidor medio pasa unas cuatro horas al día mirando la pantalla de su smartphone (App Annie, 2021). Suponiendo que el número medio de horas de vigilia sea 17, esta cantidad de uso del smartphone supondría casi una cuarta parte del día de una persona; en un año, el uso acumulado sería de unas 1.460 horas, o aproximadamente 60 días completos. Aunque hoy en día muchos consumidores se proponen reducir el tiempo que pasan frente a la pantalla del móvil, este objetivo rara vez se consigue, probablemente porque este tipo de consumo digital es un comportamiento muy habitual (Anderson & Wood, 2021).

Los hábitos son notoriamente difíciles de cambiar, y los consumidores suelen tener dificultades para alcanzar sus objetivos personales cuando éstos requieren cambiar un patrón de comportamiento habitual (Wood & Neal, 2009). Por ejemplo, los consumidores que quieren perder peso a menudo no consiguen reducir su ingesta diaria de calorías. Los que quieren reducir el tiempo que pasan en las aplicaciones de las redes sociales siguen desplazándose sin pensar por estas plataformas. Debido a la fuerza de la costumbre, estos comportamientos son difíciles de cambiar y a menudo hacen que no se alcancen los objetivos de consumo (Wood & Neal, 2016). ¿Cómo puede abordarse este problema?

Mientras que las investigaciones anteriores se han centrado predominantemente en abordar un comportamiento habitual de forma aislada, yo me baso en la idea de que cambiar un hábito requiere un enfoque integral, ya que está inextricablemente ligado a otras actividades cotidianas (Marlatt & George, 1984; Witkietvitz & Kirouac, 2016). Cuando los consumidores intentan reducir un hábito no deseado, especialmente uno que ocupa una cantidad significativa de tiempo en su rutina diaria, se enfrentan a tener más «tiempo ocioso». Basándome en investigaciones anteriores (Kushlev & Leitao, 2020; Adriannse et al., 2011; Laibson, 2001) sostengo que si este tiempo es ocupado por una actividad alternativa, hay una menor probabilidad de que el antiguo hábito recaiga. Por lo tanto, dedicarse a una actividad alternativa, idealmente una que sea beneficiosa, podría ser útil para lograr un cambio de conducta duradero. Estas son las principales cuestiones de investigación que se abordan en los dos primeros capítulos: ¿Hasta qué punto son eficaces a largo plazo las intervenciones que reducen una conducta habitual no deseada a la vez que crean una actividad alternativa beneficiosa? ¿Es más eficaz que reducir el comportamiento habitual no deseado de forma aislada?

Para responder a estas preguntas, diseñé una intervención para ayudar a los consumidores a reducir los hábitos no deseados y, al mismo tiempo, fomentar la adopción de nuevas actividades beneficiosas en sus rutinas diarias. Los capítulos 1 y 2 de la presente investigación detallan dos estudios de campo, en los que los consumidores a los que se incentiva monetariamente para que reduzcan un hábito no deseado y, al mismo tiempo, participen más en una actividad beneficiosa, tienen más éxito a la hora de mantener su cambio de comportamiento, que los consumidores a los que sólo se incentiva para que reduzcan el hábito no deseado. Investigo la persistencia del cambio de comportamiento a corto plazo (inmediatamente después de la intervención) y a largo plazo (40 días después de finalizar la intervención).

Elegí el uso de las redes sociales en el Capítulo 1, y el uso general del smartphone en el Capítulo 2, como conductas focales. Estos comportamientos no sólo pueden medirse objetivamente con aplicaciones de tiempo de pantalla, sino que también son muy habituales, lo que los convierte en los principales candidatos para probar esta intervención. Es importante destacar que muchos consumidores están preocupados por su consumo de medios digitales y, lo que es más importante, les gustaría reducirlo (Anderson & Wood, 2021; Deloitte, 2019).

Como actividades alternativas, elegí aprender un nuevo idioma (capítulo 1) y realizar actividad física (capítulo 2). Al igual que el uso del smartphone, ambas actividades pueden medirse objetivamente con aplicaciones de seguimiento. Además, se consideran beneficiosas debido a su impacto positivo en la salud, la función cognitiva y el bienestar emocional (por ejemplo, Lavie et al., 2019; Nieman & Wentz, 2019, Zimmermann & Chakravarti, 2022; De la Rosa et al., 2020; Egger et al., 2019, Chekroud et al., 2018; Kvam et al., 2016). Es importante destacar que ambas son comúnmente reportadas como actividades en las que a los consumidores les gustaría participar más (Hoare et al., 2017; Young, 2014).

Los resultados muestran que integrar una actividad alternativa positiva (por ejemplo, aprender un nuevo idioma o aumentar la actividad física) en la rutina diaria al tiempo que se reduce un comportamiento no deseado (por ejemplo, el uso excesivo del teléfono o las redes sociales) puede reducir el riesgo de recaída. Este hallazgo se aplica a las personas que realizan de forma más constante la actividad positiva al tiempo que reducen el comportamiento no deseado. Al vincular dos conductas distintas a un único incentivo monetario (en lo sucesivo conocido como enfoque de «doble objetivo»), podemos mitigar el riesgo de recaída, ofreciendo una intervención más rentable en comparación con las intervenciones tradicionales que se centran únicamente en una conducta. Aunque pueda parecer contradictorio, ya que alcanzar dos objetivos conductuales con un incentivo es más difícil, demuestro que el mayor esfuerzo merece la pena. Cumplir con éxito ambos objetivos, en lugar de uno solo, mejora las posibilidades de lograr un cambio de comportamiento persistente a largo plazo.

En conjunto, los capítulos 1 y 2 aportan varias contribuciones. En primer lugar, mi enfoque se basa en intervenciones previas de cambio de conducta que abordan una conducta a la vez, y las amplía, así como en la investigación sobre adicciones compensatorias (es decir, la sustitución de un hábito no deseado por otros similares y negativos; Dai y Fishbach, 2014; Dey et al., 2019; Melumad y Pham, 2020). Mi planteamiento es diferente del temptation bundling, en el que los vicios se vinculan a virtudes (por ejemplo, ver tu serie favorita mientras haces ejercicio en el gimnasio) para apoyar la formación de hábitos (Milkman et al., 2014). En su lugar, animo a los consumidores a utilizar el tiempo ganado al reducir un vicio para dedicarse a una alternativa virtuosa. Por lo tanto, ambas actividades no pueden realizarse a la vez.

Además, la intervención de doble objetivo está relacionada con el desbordamiento conductual, pero es distinta de él (por ejemplo, Dolan y Galizzi, 2015; Galizzi y Whitmarsh, 201 9). Un spillover significa que el comportamiento primario «malo» (por ejemplo, el uso de los medios sociales) se reduciría involuntariamente con solo aumentar las actividades alternativas «buenas» (por ejemplo, el uso de Duolingo). Algunos ECA anteriores están directamente relacionados con este fenómeno (p. ej., Brailovskaia et al., 2022), en los que (1) se pedía a los participantes que realizaran los comportamientos alternativos y se medían tanto los comportamientos objetivo como los alternativos, o (2) se sugería a los participantes que realizaran ambos comportamientos, pero no se proporcionaban incentivos. Durante la intervención de doble objetivo, tanto los buenos como los malos comportamientos se persiguen activamente y se incentivan. Así pues, la presente investigación se sitúa en un ámbito relacionado, pero diferente, en el que la reducción de uno y el fomento de otro son mucho más deliberados y, a falta de una palabra mejor, calculados.

Investigaciones de marketing anteriores han demostrado que las empresas están motivadas para mantener a los consumidores constante y puntualmente comprometidos a través de los medios sociales, porque mejora las relaciones consumidor-marca (Hudson et al., 2016;

Dhaoui & Webster, 2021) y recluta a los consumidores como contribuyentes al éxito de la marca (Saboo et al., 2016). Sin embargo, los estudiosos han empezado a destacar recientemente la importancia del consumo sostenible y moderado en el marketing para fomentar la lealtad a la marca a largo plazo, mejorar la reputación corporativa y abordar las preocupaciones sociales (White et al., 2019; Cornil et al., 2022; Somasundaram et al., 2023). Frenar el consumo excesivo puede ser incluso beneficioso para las empresas que ofrecen productos digitales. Como demuestran Nevskaya y Albuquerque (2019), frenar el tiempo de juego excesivo beneficia a los beneficios de las empresas con suscripciones más largas.

Así, el presente trabajo ofrece una intervención práctica para reducir los hábitos de consumo excesivo y presenta nuevas vías de marketing que las empresas pueden capitalizar ofreciendo alternativas beneficiosas. Estas alternativas sirven para llenar el vacío de necesidades que deja la reducción del consumo no deseado. Por ejemplo, al reducir el consumo indeseado de tecnología, los consumidores tienen la necesidad de ocupar su nuevo tiempo (es decir, el vacío de necesidades), por lo que podrían buscar productos y servicios alternativos. Los profesionales del marketing pueden aprovechar esta brecha para introducir nuevas alternativas que no solo se alineen con sus objetivos financieros, sino que también respondan a la creciente demanda de responsabilidad social, en línea con el marketing social (Kotler y Zaltman, 1971; Lee y Kotler, 2015).

Varios estudios recientes han empleado incentivos monetarios para reducir el uso de smartphones (Stanley et al., 2022; Collis & Eggers, 2022; Allcott et al., 2022; Allcott et al., 2020). Del mismo modo, algunos estudios han investigado cómo la recomendación de una actividad alternativa afecta al cambio de comportamiento (Reed et al., 2023; Brailovskaia et al., 2022; Esmaeili & Ahmadi, 2018). La investigación actual se suma a este cuerpo de literatura, ya que combina ambos enfoques para prevenir la recaída (véase el Apéndice A1 para una comparación detallada de esta investigación con la literatura relevante). En lugar de confiar en la iniciativa propia de los consumidores para participar en una actividad alternativa, establezco objetivos concretos y procesables tanto para reducir un hábito no deseado como para aumentar una actividad alternativa, y vinculo ambos objetivos a un único incentivo para fomentar el cambio de comportamiento a diario, lo cual es crucial para la formación de hábitos (Somasundaram et al., 2023; Somasundaram et al., 2024). Por último, supero un inconveniente metodológico al medir todas las conductas de forma objetiva.

Investigaciones recientes han demostrado que existe un vínculo negativo entre el tiempo de pantalla móvil y el rendimiento académico (Giunchiglia et al. 2018), el rendimiento en el trabajo (Liu et al., 2020), así como el bienestar (Twenge & Campbell 2019) y el sueño (Hisler et al., 2020). Los capítulos 1 y 2 proporcionan una respuesta oportuna a la necesidad emergente de una solución eficaz para reducir el uso de teléfonos inteligentes/medios sociales. Sin embargo, el alcance de los capítulos 1 y 2 se limita únicamente a la clase de intervenciones que promueven el cambio de comportamiento mediante incentivos y/o actividades alternativas. Existe una amplia gama de enfoques para abordar el consumo problemático de teléfonos inteligentes y redes sociales. Por lo tanto, me animé a explorar otros tipos de intervención que pretenden paliar este problema en el Capítulo 3, mediante una exhaustiva búsqueda bibliográfica y un metaanálisis. Así pues, el capítulo 3 pretende responder a las siguientes preguntas de investigación: ¿Hasta qué punto son eficaces las intervenciones para reducir el uso de teléfonos inteligentes y redes sociales y mejorar el bienestar? ¿Existen diferencias entre los distintos tipos de intervención? ¿Cómo afectan los factores demográficos (como la edad y el sexo) y las características del estudio (como el diseño del estudio y la duración de la intervención) a la eficacia de estas intervenciones?

En el capítulo 3, identifico qué estrategias han sido evaluadas por su eficacia en la reducción del uso problemático, y cómo se comparan entre sí en dos tipos de resultados: consumo (es decir, tiempo de pantalla del smartphone, tiempo de uso de los medios sociales), y mejora de las consecuencias (es decir, adicción, bienestar, malestar, calidad del sueño, productividad). Utilizando un enfoque fundamentado, clasifico las intervenciones en tres grandes tipos: Limitadas, Limitadas Plus y No Limitadas, para que puedan compararse de forma significativa en modelos metaanalíticos. Amplío el alcance de mis análisis examinando la influencia de las características relevantes de los estudios (p. ej., diseño de la intervención, duración del tratamiento) y las características de la población (p. ej., edad, sexo) en los tamaños del efecto de las intervenciones.

Mi análisis en el capítulo 3 demuestra que, en promedio, las intervenciones tienen un efecto medio en la reducción del consumo de medios sociales y teléfonos inteligentes, y mejoran los resultados de las consecuencias en una cantidad de pequeña a mediana. Además, en promedio, no hay diferencias en el tamaño de los efectos entre los tres tipos de intervenciones. Sin embargo, dependiendo del objetivo principal (ya sea la reducción del consumo o la mejora de las consecuencias) del tipo de intervención, las intervenciones Sin Límite funcionan mejor en los resultados de las consecuencias, mientras que las intervenciones Con Límite y Con Límite Plus ayudan más a reducir el consumo. También obtuve información sobre cómo diversos factores pueden influir en la eficacia de las intervenciones. Por ejemplo, mientras que los más jóvenes se benefician más de las intervenciones que mejoran el bienestar y el malestar, las personas mayores experimentan una mayor reducción del consumo de medios sociales y del uso de teléfonos inteligentes gracias a las intervenciones. Y lo que es más notable, las mujeres parecen beneficiarse menos de las intervenciones que los hombres, como demuestran los tamaños del efecto, entre 2 y 3 veces menores que los de los hombres, tanto para el consumo como para las consecuencias.

Con este conjunto de conclusiones, el capítulo 3 realiza las siguientes aportaciones. En primer lugar, proporciona un marco que caracteriza ampliamente las intervenciones en función de si imponen restricciones directas al consumo de tecnología móvil y cómo lo hacen. Esto permite comparar los tamaños de los efectos entre categorías de intervención ampliamente diferentes, lo que amplía el alcance de la síntesis de la investigación de las pruebas causales, en contraste con estudios anteriores que sólo han investigado un conjunto reducido de intervenciones. En segundo lugar, este marco reconoce la heterogeneidad entre los estudios de intervención, una preocupación planteada por investigaciones anteriores en este ámbito (Valkenburg, 2022), mediante la incorporación de múltiples características de resultados, población y estudio. Estadísticamente, abordo esta cuestión mediante el uso de modelos meta-analíticos de efectos aleatorios multinivel y multivariantes, que tienen en cuenta los posibles factores de confusión que podrían sesgar las estimaciones del tamaño medio del efecto producidas por meta-análisis anteriores que utilizan modelos convencionales de sólo intercepción de dos y tres niveles para la estimación. Esto constituye una importante contribución metodológica del presente trabajo a la literatura. En tercer lugar, mi búsqueda bibliográfica da como resultado un número mucho mayor de estudios en comparación con los metanálisis y las revisiones sistemáticas anteriores que han resumido las pruebas causales en este ámbito. Así pues, mis análisis proporcionan una evaluación más exhaustiva de las pruebas disponibles sobre la eficacia de las intervenciones. Por último, basándome en las conclusiones de este capítulo, doy consejos concretos y aplicables que pueden orientar el diseño de futuras intervenciones, lo que puede interesar a los responsables políticos y a los investigadores. Junto con los meta-análisis, la recopilación de intervenciones que han sido probadas (Apéndice D3) también puede servir como una fuente de referencia útil para otras partes interesadas, como los padres de los jóvenes consumidores que buscan enfoques basados en la evidencia para ayudar a guiar su consumo de los medios sociales y el uso de teléfonos inteligentes. Esto es especialmente relevante a la luz del reciente aumento del número de padres preocupados en el Reino Unido y en otros lugares que se han congregado en grupos de interés especial y han pedido una prohibición general de las redes sociales para los niños menores de 16 años (Banfield-Nwachi, 2024).

This page is intentionally left blank.

Chapter 1: Replacing Social Media with Duolingo

Abstract: *In this chapter, I reviewed the literature on how smartphone and social media consumption may be considered maladaptive, and develop a dual-target intervention that might remedy this issue while ensuring long-term intervention effect. I utilized a longitudinal pretest-posttest design to test the main hypothesis that subjects who chose to simultaneously reduce social media usage and increase language learning time for more days during the treatment period would reduce their post-treatment social media usage more from their baseline usage. Subjects (N = 43; 1,588 observations, 72% female; M_{age} = 21.6, SD_{age} = 1.79) were incentivized to achieve two different targets during the treatment period: Subjects could earn €1 per day if they reduced their social media usage by 30% (Single Target); OR they could earn €2 per day if they reduced their social media usage by 30% AND additionally increased their language learning time by 20 minutes in Duolingo (Dual Target). Subjects were free to choose whether they wanted to achieve both targets, only the single target, or neither. Results show that the main hypothesis is supported. In addition, subjects who achieved both targets for more days also experienced lower social media addiction.*

1. Introduction

Changing habits is notoriously difficult, as many consumers struggle to achieve goals that require altering entrenched behavior patterns (Wood & Neal, 2009). For example, those aiming to lose weight often fail to consistently reduce their calorie intake, and individuals wishing to decrease social media use frequently end up mindlessly scrolling through these platforms. The force of habit makes these behaviors challenging to modify, leading to repeated failures in reaching consumption goals (Wood & Neal, 2016). While previous research has focused on addressing habits in isolation, I propose that a comprehensive approach is necessary, as habits are intertwined with other daily activities (Marlatt & George, 1984; Witkietvitz & Kirouac, 2016). When consumers try to cut down on an undesired habit, particularly one that occupies a significant part of their routine, they often have more idle time. Building on prior studies (Kushlev & Leitao, 2020; Adriannse et al., 2011; Laibson, 2001), I argue that occupying this time with an alternative activity reduces the likelihood of relapse. Engaging in a beneficial alternative activity can support lasting behavior change.

In this chapter, I investigate the effectiveness of this approach in an intervention that aims to reduce social media usage while encouraging language learning. After a baseline period of seven days, subjects proceeded to the treatment period that lasted for 20 days, followed by a post-treatment period of 13 days. At the start of the treatment period, subjects were informed about two different incentive schemes: (1) earning €1 each day they reduce social media usage by 30% from baseline usage (single-target), or (2) earning €2 each day they reduce social media usage by 30% and increase learning time with Duolingo by 20 minutes (dual-target). I chose Duolingo because of its availability on both Android and iOS and built-in habit-fostering features (Tsai, 2016). Each day, subjects could choose to achieve the single target, the dual target, or neither of the two.

I found that, for each additional day of dual-target achievement (instead of single-target achievement) during the treatment (incentivization) period, subjects reduced, on average, 4.3 minutes of post-treatment social media consumption compared to baseline. In addition, subjects with higher dual-target achievement perceived themselves to be less addicted than those with higher single-target achievement. This opens up the possibility that by replacing social media with using Duolingo, these subjects might have self-signalled that social media is not as necessary for them as they used to believe, which might result in the observed sustained social media reduction in the post-treatment period. Furthermore, I utilized individual variation in the ratio of Duolingo target to social media target to show the existence of a behavior replacement effect[2]. Finally, I found that the overall intervention resulted in an improvement in subjects' well-being, regardless of the type of target achievement.

Next, I present the relevant literature that contextualize the current study and its findings within existing research.

1.1 Smartphone and social media usage as maladaptive habits

Habits are behaviors that become automatic and effortless after regular repetition (Becker & Murphy, 1988; Pollak, 1970; Wood & Rünger, 2016). Smartphone usage is considered a highly habitual behavior. The average consumer checks their phone anywhere from 47 to 58 times a day (Deloitte, 2017; RescueTime, 2019), while spending around four hours per day looking at their phone (App Annie, 2021). Many adolescents reportedly reach for their smartphones as the first thing in the morning and the last thing at night (Toh et al., 2019; Walsh et al., 2008). In multiple surveys studying the habitual nature of social media usage (Anderson & Wood, 2021), consumers consistently revealed that their posting activities on platforms such as Twitter and Facebook were largely automatic. Given that social media comprises a large

[2] This is further explained in Section 2.3 and 3.3.

portion of smartphone usage (about 44%; Kemp, 2021), there is strong evidence for the automaticity and habitual nature of smartphone usage.

Habitual smartphone usage can become maladaptive and addictive (Alter, 2017; Allcott et al., 2022; Wang et al., 2015). For instance, smartphone users excessively crave the rewarding feeling or seek the stress-soothing effect of their smartphones (Melumad & Pham, 2020), while neglecting other activities or goals that might hold greater importance to them. Hence, smartphone usage can negatively affect consumers (Turel & Bechara, 2021). For example, research has pointed to negative associations between mobile screen time and academic performance (Lepp et al., 2015; Giunchiglia et al., 2018; Zimmermann, 2021), performance at work (Lanaj et al., 2014; Liu et al., 2020; Chadi et al., 2022) as well as well-being (Lee et al., 2014; Volkmer & Lermer, 2019) and sleep patterns (Hisler et al., 2020; Twenge & Campbell, 2019; Levenson et al., 2017). Scholars have suggested that smartphones can negatively affect well-being by displacing time that should be spent on healthier activities (Kushlev & Leitao, 2020).

I argue that frequent smartphone usage precludes the opportunities for engaging in more beneficial alternative activities (e.g., achieving personal goals) simply because it consumes a significant amount of time that would otherwise be available. In this research, apart from investigating the focal habits and alternative activities, I also measure smartphone addiction as a potential mechanism, and subjective well-being as a secondary outcome of behavior change.

1.2 Monetary incentives for reducing mobile consumption

Monetary incentives can be effective in encouraging consumers to lower their smartphone and social media usage (Stanley et al., 2022; Collis & Eggers, 2022; Allcott et al., 2022; Allcott et al., 2020; Somasundaram et al. 2024). However, this type of incentives can also “crowd out” intrinsic motivation and be counterproductive for long-term habit formation

(Gneezy et al., 2011). As mentioned earlier, a common problem of many habit formation interventions is that after a period of successful behavior change, consumers revert to their previous consumption level, showing the typical "triangular relapse pattern"[3] (Wood & Neal, 2016). To avoid this, I draw on literature from addiction research and make the case for changing an undesired habit sustainably by encouraging consumers to pursue alternative activities.

1.3 Previous research on alternative activities

Viewing mobile usage from an addiction lens allows me to tap into methods proven effective for more tangible addictions, such as smoking. For example, nicotine replacement therapy (NRT) delivers a prescribed amount of nicotine (usually via a skin patch or chewing gum) to ease withdrawal symptoms (Stead et al., 2012). In the current context, similar to a nicotine patch, monetary incentives can compensate for the negative feelings of anxiety and stress from using one's smartphone less (e.g., nomophobia, King et al., 2013). Combining NRT with a substituting activity, such as physical exercise, leads to more success in smoking cessation (Prapavessis et al., 2007; Prapavessis et al., 2016). While the precise mechanism is unclear (Ussher et al., 2014), scholars have suggested that physical exercise provides distraction from withdrawal bouts (Dishman, 2009; Van Rensburg et al., 2009). Similarly, when consumers deliberately occupy the time freed up by reducing their smartphone usage with a beneficial alternative, they should experience less intense withdrawal symptoms and feel less addicted to their smartphone. I empirically explored this idea in Chapter 1.

There are several theoretical explanations related to cues and self-signaling of why engaging in alternative activities might work better than simply trying to reduce a habitual

[3] This refers to a line graph pattern characterized by three points. The first point represents the baseline quantity of a behavior, the second point represents this behavior during the intervention period which deviates considerably from baseline, and the third point represents the behavior in the post-intervention period, which almost always recedes closer to baseline. These three points form three vertices of an imaginary triangle, hence the name.

behavior. First, thinking about *not* executing or abstaining from a habitual behavior (a *negation* implementation intention) ironically strengthens that exact habit one aims to break. Focusing on an alternative activity (a *replacement* implementation intention) weakens the cue-response linkage of the undesired habit by associating the cue with a different response (Adriannse et al., 2011)[4]. According to Laibson's (2001) cue theory of consumption, for a consumer to engage in an activity, the associated cue must push the behavioral tendency beyond a certain 'threshold' of activation. While this threshold is lower for a habit (i.e., even the slightest, internal cue like getting bored might activate the habitual response), it increases if other activities are present, serving as alternatives that compete with the habitual activity. In the current context, an alternative activity such as learning a new language or walking would increase the activation threshold for habitual social media or smartphone usage. I expect that consistent performance of an alternative activity will serve to maintain a higher activation threshold for the undesired habit.

Second, self-signaling processes can stem from performing an alternative activity instead of an undesired habit. Dai and Fishbach (2014) show that the presence of salient alternatives can lower the desire for an unconsumed good when abstaining from it because consumers infer that they have developed new tastes. In their study, consumers abstained from using the social media platform Facebook and instead replaced it by using Twitter. In contrast, the beneficial alternative activities in the current studies are not direct substitutes for the undesired habits. Thus, I prevent consumers from developing compensatory addictions by simply trading one addiction for another (Dey et al., 2019; Melumad & Pham, 2020).

[4] The dual-target interevention in the current research differs from using implementation intention (or specific if-then plans; Adriaanse & Verhoeven, 2018) in that it does not prescribe an explicit cue-action behavior rule (e.g., if I want to use social media, I will use Duolingo instead). Instead, it lets consumers decide for themselves how they approach the behavior change(s), in the face of incentives.

I argue that by engaging in a beneficial alternative activity that is not a direct substitute, consumers self-signal that they are less addicted to their smartphone than they thought they were. Consequently, consumers become more receptive to exploring and adopting various alternative activities, possibly extending beyond the incentivized one. This shift in perspective encourages consumers to view their time as a valuable resource, thereby contributing to a sustained behavior change. In line with this reasoning, Mead and Patrick (2016) show that a replacement activity can help with unspecific postponement ("I can have it some other time") which signals that the tempting activity is of low importance, thus reducing craving in the moment and over time.

Based on this literature and my theorizing, I expect consumers with higher dual-target achievement during the treatment period to maintain a lower social media usage and screen time during the post-treatment period compared to those with more single-target achievement or those who did not achieve any target.

2. Methods

2.1 Study design

The study was pre-registered[5] (1,588 observations across $N = 43$; 31 females; $M_{age} = 21.6$, $SD_{age} = 1.79$) and approved by the IRB at a large international university. At the start, I asked student subjects to install a screen time tracking app and the Duolingo app[6] to obtain objective measures. I collected individual-level panel data on daily social media usage during six-weeks. Subjects reported their daily social media usage over the past seven days by completing an online report at the end of each week. To validate these reports, subjects submitted a screenshot

[5] https://aspredicted.org/blind.php?x=D3H_HPR. Some changes were made to the hypotheses, and these are explained in Appendix B11.

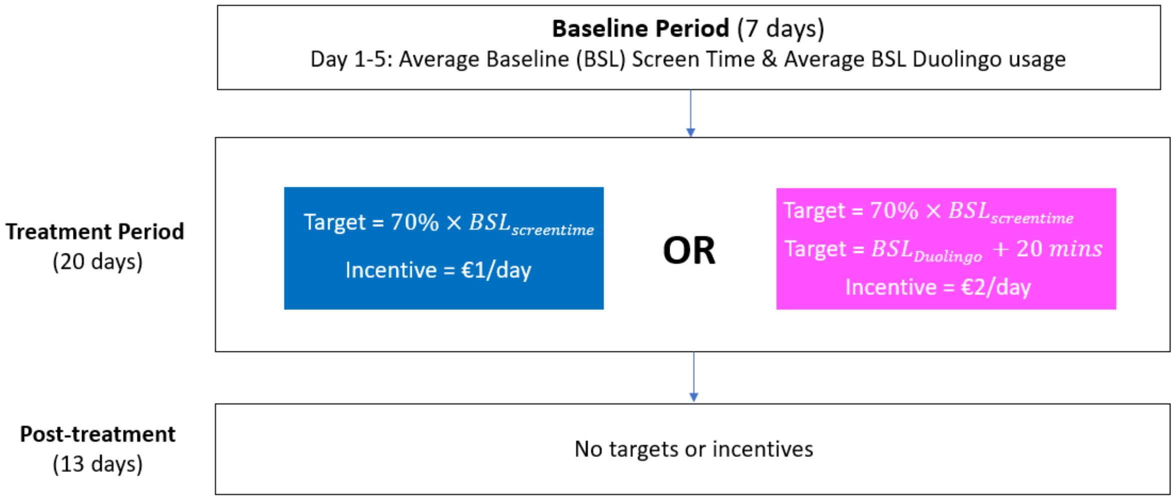
[6] For iPhone users, Apple's inbuilt iOS Screen Time feature was used: https://support.apple.com/en-gb/HT208982. For Android users, an equivalent app, which could be downloaded free of charge, was used: https://play.google.com/store/apps/details?id=com.burockgames.timeclocker&referrer=webpage. Duolingo apps for the respective platforms can be downloaded from: https://www.duolingo.com/.

of their screen time app and the Duolingo usage for the preceding seven days (see Appendix B1 for an example screenshot). The duration of the baseline, treatment and post-treatment period was decided based on constraints of the academic calendar to minimize attrition, as well as to be consistent with past habit formation studies (e.g., Loewenstein et al, 2016). Figure 1 shows the study design and timeline. During the baseline period, subjects were given no targets.

At the start of the treatment period, all subjects were informed about their average baseline social media and Duolingo usage and provided with a handout outlining the benefits of reducing social media and learning a new language. They were informed that they would receive 2€ or 1€ each day during the treatment period for achieving either dual or single

Figure 1

Design of Chapter 1. Subjects can choose on a daily basis during the Treatment Period whether they want to achieve a single-target (ST) or dual-target (DT).



targets, respectively[7]. Note that, rather than being randomized into one target type, subjects could choose on a daily basis the type of target they wanted to achieve. Thus, this study has a quasi-experimental design.

[7] I opted to offer extra incentive for the dual-target achievement. This is because, with equal incentives, there would be no reason to achieve dual-target, and I would have very few dual-target data points to work with. Thus, to incentivize a sufficient number of people to achieve dual-target, I used a higher incentive. I discuss the limitation of this approach in the Discussion section.

To ensure that subjects understood the instructions and to give them time to prepare, there was a gap of three days between the baseline and the treatment period during which they completed a comprehension survey. To strengthen the intervention, at the end of the first and second week of treatment, I sent personalized text messages informing subjects about the number of days they had achieved each target (see Appendix B2 for the SMS template). In the post-treatment period (Period 2), I removed the targets and incentives, but subjects continued to report their social media and Duolingo usage for the next thirteen days.

2.2 Surveys

Subjects completed a baseline, pre-treatment, and post-treatment survey at the start of the corresponding period. In addition to demographic and social media usage characteristics, I measured self-reported addiction to social media and subjective well-being (adapted from Allcott et al, 2022; see Appendix B4). The 6-item addiction questionnaire (Cronbach's $\alpha = .75$) measures addiction to social media over the previous 3 weeks (e.g. "How often have you been worried about missing out on things online when not checking your phone?") on a 5-point Likert scale (1 – Never, 5 – Always). The 7-item well-being questionnaire (Cronbach's $\alpha = .84$) assesses emotional and cognitive states over the previous 3 weeks. Subjects rated statements such as "I was a happy person.", "I was easily distracted." on a 7-point Likert scale (1 = Strongly Disagree, 7 = Strongly Agree)[8].

I also included questions related to language proficiency. Subjects were asked how many languages they spoke, whether they were currently learning a language, and if they utilized any mobile app for language learning. The survey results are reported in Appendix B3. The correlation matrix of survey variables is in Appendix B5.

[8] Survey materials are available for both studies at: https://osf.io/p7jn9/?view_only=d0f66908927d462291d45d47f194ff22

2.3 Hypotheses and analysis plan

I propose that the alternative activity (learning a language with Duolingo) helps consumers reduce and sustain lower social media usage by filling up the freed up time. Therefore, I expect subjects with more dual-target achievement (measured in days) to maintain a lower social media usage during the post-treatment period compared to those with more single-target achievement or those who did not achieve either target type:

H1: Subjects with higher dual-target achievement during the treatment period will reduce their post-treatment social media usage more, as compared to those with higher single-target achievement or no target achievement.

To test H1, I use a panel regression with individual daily social media usage as the outcome variable and the corresponding unit of observation. The predictors included are dummy-coded Period variable (3 levels corresponding to Baseline, Treatment, and Post-treatment Periods; Baseline Period is the base category), dual-target achievement (in days), and their interaction terms. Standard errors are robust and clustered at the subject level. The result of this analysis is presented in Table 2, Model 2.

To quantify the amount of replaced time, I operationalize a 'target replacement ratio' (TRR). TRR indicates the extent of time replaced of one activity by another. TRR is high when a person has a high Duolingo target relative to their social media target. For example, if a subject's social media target is 2 hours and Duolingo target is 0.4 hours, then the TRR is 20%. Thus, subjects with a higher proportion of achieving the language learning target relative to their social media target should be more likely to maintain a lower social media usage after the treatment period than subjects with a lower replacement ratio.

H2: Subjects with higher TRR and high dual-target achievement during the treatment period will have a lower post-treatment social media usage compared to their baseline usage.

To test H2, I use an OLS regression predicting the difference in social media usage between post-treatment and baseline periods with TRR, dual-target achievement (in days), their interaction term, and relevant controls (age, gender, phone's OS). The unit of observation is per subject. The result of this analysis is presented in Table 3, Model 2.

By reducing the use of social media with an alternate activity (Duolingo) in the treatment period, consumers might self-signal that they are less addicted to social media and rather spend their time on other activities (Mead & Patrick, 2016; Dai & Fishbach, 2014). This should help to maintain a reduced social media usage over an extended period. In the current study, I use perceived addiction to social media as a proxy measure for this self-signalling process.

H3: Subjects with higher dual-target achievement during the treatment period will experience a greater reduction in perceived social media addiction compared to those with lower dual-target achievement.

To test H3, I use an OLS regression predicting the difference in perceived addiction to social media between post-treatment and baseline periods with dual-target achievement (in days), and relevant controls (age, gender, phone's OS). The unit of observation is per subject. The result of this analysis is presented in Table 4.

Considering the detrimental impact of excessive social media usage on well-being (Giunchiglia et al., 2018; David & Roberts, 2017; Levenson et al., 2017) and the documented improvements in well-being following reduced social media usage (Allcott et al., 2020; Hunt et al., 2018), I posit the following hypothesis:

H4: Reducing social media during the intervention (regardless of types of target achievement) is associated with improvements in subjective well-being after the intervention.

To test H4, I use a paired-sample t-test comparing the subjective well-being index before and after the treatment period.

3. Analyses and Results

3.1 Target achievement

All subjects except one achieved at least one of the targets for all twenty treatment days (Table 1). This is evident from the highly negative correlation between the two types of target achievement (Pearson's $r = -0.998$, $p < .001$). One subject achieved no target at all. Excluding this subject did not change the results of any analyses. Therefore, all subsequent analyses were performed on the full sample.

Table 1

Number of target achievement days in the sample

	No target achievement	Single-target achievement	Dual-target achievement	Total days in treatment
Number of days	2	687	171	860

3.2 Average change in social media usage and the effect of dual-target achievement

In Figure 2, I plot the average daily change in social media usage from the average usage (3.08 hours) during the entire baseline period. During the treatment period, subjects reduced their baseline social media usage by 0.912 hours (54.7 minutes) on average. In the post-treatment period, this reduction was 0.194 hours (11.6 minutes).[9]

To statistically test this pattern, I regressed daily social media usage on the categorical period dummy variables which capture the change in usage across each period, compared to baseline usage (i.e., baseline is the base category). The results of this regression are presented in Table 2, Model 1. On average, subjects decreased their social media usage during the

[9] Refer to Appendix B6 and B6.2 for timeline graphs of average change in Duolingo usage. Subjects increased their Duolingo usage in the intervention period compared to their baseline, but did not maintain this increase during the post-treatment. I elaborate on this in the Discussion of Chapter 1 and 2.

treatment period (Period 2, β = -1.029, p < .001) compared to their baseline usage, as expected, and maintained a marginally lower usage during the post-treatment period (Period 3, β = -0.235, p = .088).

Figure 2

Social media usage of the full sample vs. subsample with higher dual-target achievement

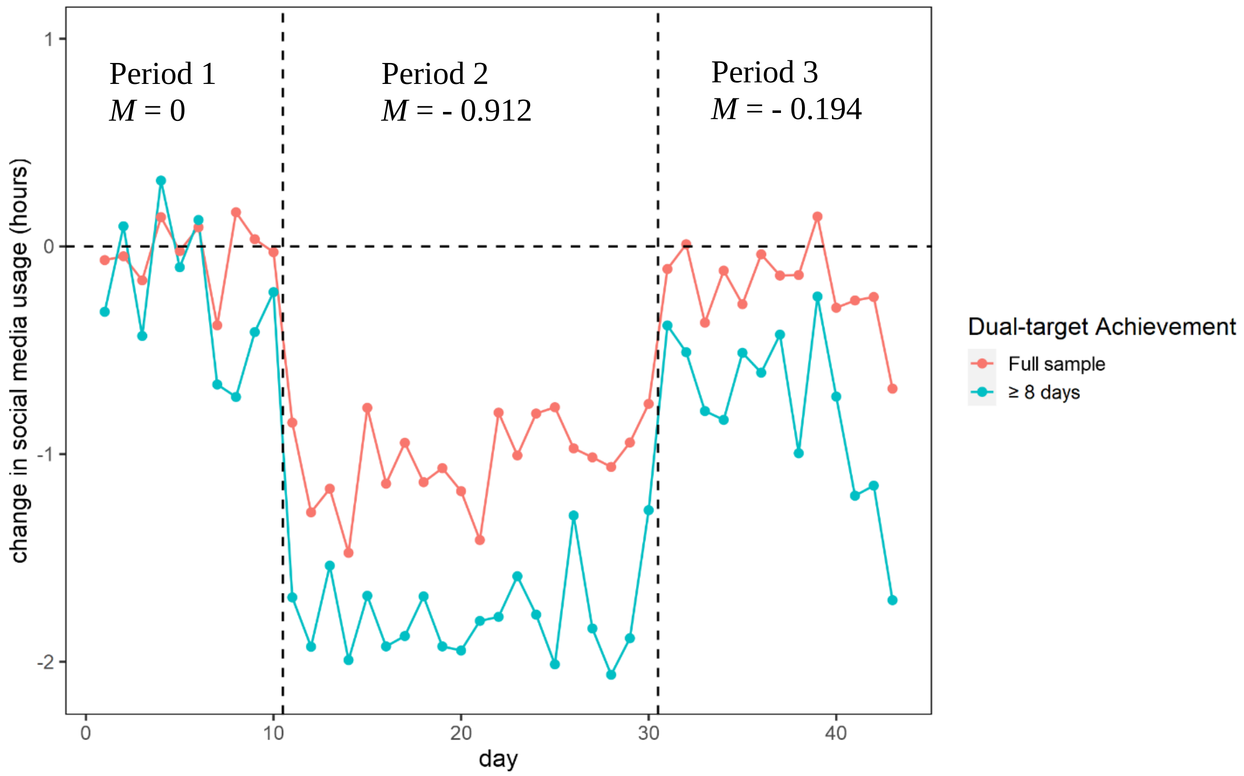


Figure 2 shows that 'high-performing' subjects who achieved at least 8 days of dual-target (i.e., 75th percentile and above) had lower social media usage during the treatment and post-treatment period. I formally test this by regressing daily social media usage on (1) the categorical period dummy variables which capture the change in usage across each period compared to the baseline usage (i.e., baseline is the base category), (2) the number of days of single-target achievement, (3) the number of days of dual-target achievement, and most importantly, (4) the interaction between each target achievement variable and each period dummy to capture the change in usage from the baseline as a function of single vs. dual-target achievement.

Table 2

Panel regression analyses of Social Media usage time (Chapter 1). Dual-target achievement predicts post-treatment social media reduction.

	Dependent variable = Social Media Usage (hour)	
	Model 1	Model 2
Treatment Period (Period 1)	-1.029***	-0.627***
	(0.141)	(0.159)
Post-Treatment Period	-0.235*	0.065
(Period 2)	(0.138)	(0.174)
Dual-target achievement		0.012
(days)		(0.026)
Dual-target achievement ×		-0.101***
Period 1		(0.017)
Dual-target achievement ×		-0.072***
Period 2		(0.024)
Constant	3.081***	3.032***
	(0.165)	(0.208)
Observations	1,588	1,588
R^2	0.085	0.154
Adjusted R^2	0.083	0.151
Residual Std. Error	1.435 (df = 1585)	1.381 (df = 1582)
F Statistic	73.272*** (df = 2; 1585)	57.659*** (df = 5; 1582)

Note: * $p < .10$; ** $p < .05$; *** $p < .01$. The base category is the baseline period. Standard errors (in parentheses) are robust and clustered at the subject level.

Due to the nearly perfect negative correlation between single- and dual-target achievement, I can only choose one of them as a predictor[10]. For ease of interpretation, I opted

[10] Dual and single target achievements are correlated because participants freely chose between single- vs. dual-target during the treatment period. Thus, an increase in a type of target necessarily leads to the decrease in the other type of target. During the intervention, participants happened to adhere to either target for most of the days during the treatment period, which led to these two types of target being almost perfectly negatively correlated.

for dual-target achievement. Using single-target achievement results in nearly identical coefficients with the opposite sign.

Regression estimates are presented in Table 2, Model 2. Standard errors are clustered at the subject-level to account for potential within-subject correlation. The interaction terms indicate that for every additional day of dual-target achievement during the treatment period, subjects decreased their usage by 0.101 hour (or 6.1 minutes, $p < .001$) in the treatment period, and by 0.072 hour (or 4.3 minutes, $p = .003$), on average, during the post-treatment period. Because of the almost perfect negative correlation, the marginal reduction in social media usage should be interpreted as a reduction when subjects decided to achieve both lower social media *and* higher Duolingo usage compared to achieving only lower social media usage. These results support Hypothesis 1[11].

Following the same procedure, I conducted a panel regression for Duolingo usage. Results are presented in Appendix B7, Model 2. For every additional day of dual-target achievement (instead of achieving only the social media target), subjects increased their average daily Duolingo usage by 0.028 hours (or 1.7 minutes, $p < .001$) during the treatment period. This increase was also positive during the post-treatment (0.004 hours, $p = .039$). The increase in Duolingo usage should be interpreted as the increase compared to those who achieved only the social media target for a particular day.

3.3 Target replacement ratio (TRR)

I next investigate Hypothesis 2 that subjects with higher TRR and high dual-target achievement during the treatment period have a lower post-treatment social media usage compared to their baseline usage. Due to the current design, subjects had a fixed 20 minutes

[11] Including Willingness to use Duolingo as a control variable does not change this result (see Appendix Table E1, Model 3). Including Age, Gender, and phone OS as controls does not change this result either (see Appendix Table E1, Model 4).

increase of Duolingo usage as an absolute target, but a percentual reduction of social media as a relative target. Therefore, the TRR differed across subjects. The ratio is computed by dividing a subject's Duolingo target by their social media target. To investigate whether a higher level of the TRR facilitates target achievement, I predict post-treatment social media usage (compared to baseline) using the TRR, dual-target achievement, their interaction term, and relevant controls. Table 3, Model 2 shows these regression coefficients[12]. Because the level of analysis is the participant, the outcome variable (average change in social media usage) is computed by subtracting a participant's average baseline usage from his/her average post-treatment usage.

[12] Model 1 shows the regression coefficients when the interaction term is excluded.

Table 3

OLS regression of Post-Pre social media usage on target replacement ratio (TRR) and dual-target achievement. TRR and dual-target achievement jointly predict post-treatment social media reduction.

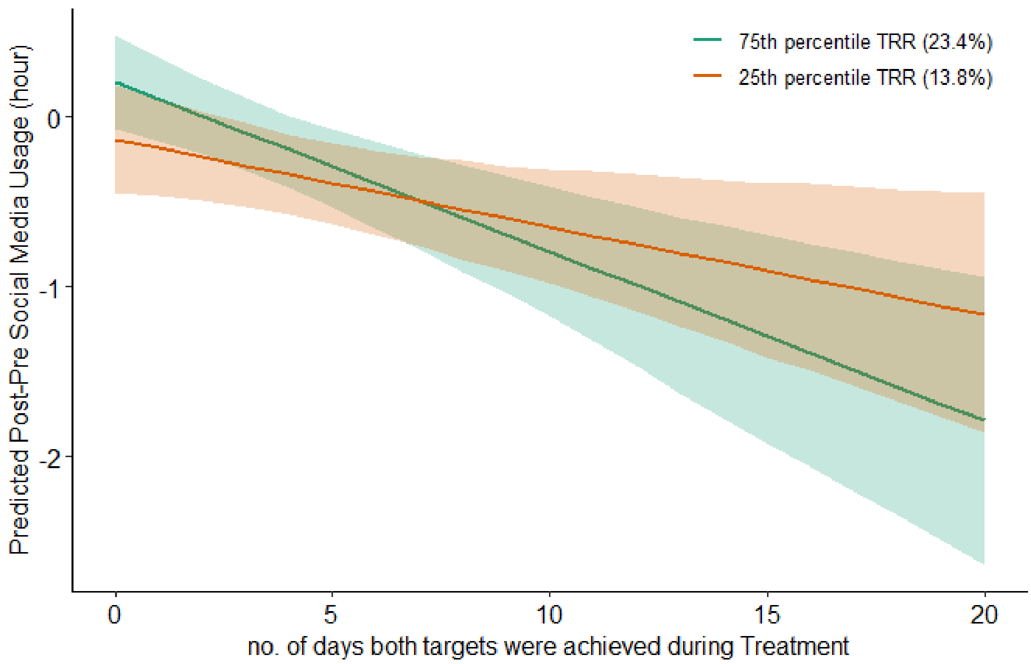
	Dependent variable = Post – Pre social media usage		
	Model 1	Model 2	Model 3
Target replacement ratio	0.020**	0.035***	0.020
(TRR)	(0.009)	(0.011)	(0.014)
Dual-target achievement	-0.072***	0.019	0.023
(days)	(0.021)	(0.047)	(0.046)
TRR × Dual-target		-0.005**	-0.005**
achievement		(0.002)	(0.002)
Baseline social media			-0.386*
usage (hours)			(0.211)
Age	-0.00002	-0.022	-0.049
	(0.067)	(0.064)	(0.064)
Gender (Female)	-0.112	-0.115	-0.120
	(0.269)	(0.257)	(0.249)
OS (iOS)	0.373	0.407	0.454
	(0.377)	(0.361)	(0.350)
Constant	-0.573	-0.419	1.296
	(1.573)	(1.503)	(1.732)
Observations	43	43	43
R^2	0.345	0.419	0.470
Adjusted R^2	0.257	0.322	0.364
Residual Std. Error	0.746 (df = 37)	0.712 (df = 36)	0.690 (df = 35)
F Statistic	3.899*** (df = 5; 37)	4.330*** (df = 6; 36)	4.432*** (df = 7; 35)

Note: * $p < .10$; ** $p < .05$; *** $p < .01$.

The interaction term is negative (β = -.005) and significant, p = .039. To facilitate interpretation, I plotted a post-estimation graph illustrating the effect of dual-target achievement on change in social media usage at the 25th and 75th percentile of TRR (Figure 3). This graph shows that subjects with a higher TRR, relative to those with a lower TRR, have a larger marginal reduction of social media usage with higher dual-target achievement during the treatment period, as evidenced by the steeper negative slope of the 75th percentile TRR. These results confirm Hypothesis 2. Dual-target achievement during the treatment period is particularly effective for lowering post-treatment social media usage when a higher proportion of the undesired habit is replaced by the positive activity (i.e., Duolingo usage).

Figure 3

Post-estimation graph of Model 2 (Table 3) at High TRR (75th percentile) and Low TRR (25th percentile). The predicted post-pre social media usage and its 95% CIs were generated with 1,000 bootstrapping samples. All control variables (age, gender, OS) were held at their means.



Subjects with higher baseline social media usage might respond better to the treatment, and these subjects also tend to have a lower TRR, because the target of social media usage

(70% of baseline usage) is used as the denominator to compute the TRR. It is therefore possible that the joint effect of TRR and dual-target achievement disappear once target social media usage is controlled for. As a robustness check, I added baseline social media usage as a control variable in Model 3; the interaction term between the TRR and dual-target achievement remained significant.[13]

3.4 Dual-target achievement reduced perceived social media addiction

Subjects with higher dual-target achievement had lower post-pre social media usage compared to those with higher single-target achievement. I suspect that dual-target achievement more effectively reduced the level of addiction to social media than single-target achievement. The result of a regression analysis (Table 4) confirms this. Dual-target achievement (instead of single-target achievement) is associated with a significant decrease in perceived social media addiction (p = .037), which supports Hypothesis 3.

[13] As an additional robustness check, I used an alternative operationalization of the interaction between TRR and Dual-target achievement to predict Post – Pre social media usage. The results are presented in Appendix B9.

Table 4

Dual target achievement predicts reduction in social media addiction (Chapter 1)

	Dependent variable
	Addiction (Post-Pre)
Dual-target achievement (days)	-0.038**
	(0.018)
Age	-0.025
	(0.055)
Gender (Female)	-0.221
	(0.217)
OS (iOS)	-0.425
	(0.311)
Constant	1.010
	(1.288)
Observations	43
R^2	0.181
Adjusted R^2	0.095
Residual Std. Error	0.615 (df = 38)
F Statistic	2.104* (df = 4; 38)

Note: * $p < .10$; ** $p < .05$; *** $p < .01$.

3.5 Well-being outcomes

To investigate the effect on well-being, I conduct a paired-sample t-test comparing the subjective well-being index before (*M* = 4.37, *SD* = 1.09) and after (*M* = 4.62, *SD* = 1.09) the Treatment period. I find that reducing social media resulted in improvements of subjective well-being, $t(42) = 2.03$, $p = .049$. Thus, Hypothesis 4 is supported.

4. Discussion

Chapter 1 demonstrates that individuals with more days of dual-target achievement (relative to single-target achievement) were more likely to maintain a lower social media usage in the post-treatment period. Thus, engaging in an alternative activity may help prevent the typical relapse pattern and support consumers to successfully achieve their goals in the long run. However, individuals self-selected into the two incentive schemes on a daily basis. Hence, there might be unobservable factors that drive the difference in post-treatment sustenance of lower social media usage. Since there was almost no day when neither target type was achieved, I cannot compare the effect of no target achievement vis-à-vis dual- or single-target achievement. Chapter 2 addresses these limitations with a randomized design.

The results of chapter 1 also provide initial evidence for a replacement process of the undesired behavior (social media consumption) by the beneficial alternative (Duolingo usage). By using a target replacement ratio (TRR), I found that, at the same level of dual-target achievement (in days), subjects with higher TRR (i.e., higher Duolingo usage target as a percentage of social media consumption target) reduced their social media consumption in the post-treatment period more compared to those with lower TRR. For example, at 15 days of dual-target achievement, a 10% difference in TRR results in an average of 0.40 hour (24 minutes) difference in post-pre social media consumption (Figure 3). However, since this is the first time TRR is used as a predictor of change in social media usage, further research is needed to examine the validity and consistency of this variable in predicting behavior replacement in other contexts. In addition, since subjects did not continute to use Duolingo during the post-treatment period, it is likely that they might have engaged in other types of alternative activities instead. Future studies should ideally collect more fine-grained data that measures all other activities consumers have done so as to capture whether replacement of social media usage persists in the post-treatment period.

While I found that subjects with higher (vs. lower) dual-target achievement had lower perceived addiction to social media after the treatment period, an exploratory mediation test examining the potential mechanism via this variable (i.e., subjects realize that they are not as dependent on social media as they used to believe) was not significant (Appendix E2). The direct effect of dual-target achievement on post-pre social media usage is very strong, and even though dual-target achievement reduced both perceived addiction and social media usage, and reduced addiction to social media was associated with lower post-pre usage, when controlling for dual-target achievement, the indirect effect of perceived addiction on post-pre usage disappears. The sign of the coefficient for change in addiction on post-pre is as expected (positive, meaning reduced addiction results in reduced post-pre social media usage), and the p value is not extremely high (p = 0.27). Thus, I suspect that this might be due to the limited sample size (given that this test was conducted with only 43 observations). Future research with more subjects might be able to test this mechanism with more statistical power.

The target types were differentially incentivized (dual-target achievement for 2€/day, and single-target achievement for 1€/day) to motivate higher participation in dual-target scheme. This makes it difficult to distinguish whether the reduction observed during the post-treatment period is due to subjects being paid more or due to some other factors (e.g., task-switching between browsing social media and using Duolingo). In general, there is a self-selection issue due to the non-randomized nature of the design.

Finally, since multiple inferences are being drawn from the same dataset, it is necessary to apply a correction for multiple-hypothesis testing. Given that I am testing four hypotheses, the Bonferroni correction adjusts the critical value for rejecting the null hypothesis to α = 0.0125. With this more stringent threshold, only Hypothesis 1 ($p = 0.003$) remains statistically significant, while Hypothesis 2 ($p = 0.039$), Hypothesis 3 ($p = 0.037$), and Hypothesis 4 ($p = 0.049$) are not. Thus, the latter three hypotheses should be subjected to further testing with data obtained from a bigger sample .

Chapter 2: Replacing Smartphone Usage with Walking

Abstract: *In this chapter, I address the shortcomings of Chapter 1's study with a randomized controlled trial. I randomly assigned subjects (N = 110; 4,098 observations, 58% female;* $M_{age} = 24.4$, $SD_{age} = 3.53$*) to three different conditions: single-target (ST), dual-target (DT), and control (C). ST subjects were incentivized with €2 per day if they lowered their baseline mobile screentime by 30%. DT subjects were incentivized with €2 per day if they lowered their baseline mobile screentime by 30% AND additionally walked 2,000 steps more than their baseline. I address the problem of adherence to treatment with the instrumental variable regression approach, and found that subjects who achieved more dual-habit target days tended to reduce their screen time more than subjects who achieved more single-habit target in the post-treatment period.*

1. Introduction

A drawback of Chapter 1 is the lack of random assignment. Chapter 2 aims to address this issue with a randomized controlled trial (RCT), and extends the investigation of a dual-target intervention in several aspects. In Chapter 2, I incentivize single-target and dual-target achievement in different conditions and compare it with a control condition, where subjects are not provided targets. In addition to having a control group, separate intervention arms, and a longer post-treatment period, Chapter 2 differs from Chapter 1 in three ways.

First, I chose walking as a positive alternative activity for the following reasons. Walking is a healthy physical activity that is easy to carry out and can be tracked objectively using mobile apps and portable devices. Also, walking as an activity might distance consumers from sensory cues (e.g., phone vibrations, ringtones, lights) that can stimulate phone pickup. Finally, physical exercise helps reduce smartphone addiction (Liu et al., 2019).

Second, I tracked total screen time instead of social media usage because general smartphone usage is considered a 'gateway behavior' through which social media habits are formed (Schnauber-Stockmann & Naab, 2019). In addition, social media apps are not the sole culprit in causing negative impacts on well-being and productivity; other usage, such as gaming, might be equally responsible (Chiang et al., 2019). By measuring total screen time, subjects cannot resort to using substitutes to social media apps on their smartphones such as gaming, while still keeping their usage low and earning the incentive.

Third, I equate the incentives of both the dual target and single target conditions as a more conservative design to test the effectiveness of the dual-target intervention. In both conditions, subjects could earn €2 per day if they would achieve their respective target(s).

Chapter 2 has three conditions: In the dual-target (DT) condition, besides having a target of reducing mobile screen time by 30% from baseline, subjects had to achieve a daily step count target to receive a monetary incentive (€2 per day). In the single-target (ST) condition, subjects had the same screen time target and monetary incentives as DT condition but had no step count target. These treatments were compared to a control condition without targets and incentives. After a baseline period of seven days, the treatment period lasted for 20 days. I tracked subjects' behavior over two separate post-treatment periods: (1) immediately after, and (2) forty days after the treatment period. I refer to these two periods as early post-treatment and late post-treatment.

I found that, when subjects in each treatment condition (ST and DT) adhered to their target(s) for the same number of days, subjects in the DT condition had more sustained screen time reduction in the post-treatment period (relative to baseline period), as compared to their ST condition counterparts. Furthermore, I found that decreased smartphone usage was replaced by step count increase, but this only applied to subjects in the DT condition, who were incentivized to increase their step count and reduce screen time simultaneously during the treatment period.

Next, I present the literature on walking as an appropriate alternative activity, and why smartphone consumption should be a target for interventions.

1.1 Walking as an alternative activity

Walking as an alternative activity might distance consumers from sensory cues (e.g., phone vibrations, ringtones, lights) that can stimulate phone pickup (Laibson, 2001). Unlike the usage of alternative apps (instead of social media apps) that requires consumers to interact with their smartphones (e.g., learning a language with Duolingo), walking does not require such direct interaction. This constitutes a 'shielding' effect from more exposure and encounters with incidental or scheduled stimuli emitted by the smartphone. This is an important feature of a good alternative activity to be incentivized in a dual-target intervention, which is tested in the current chapter.

Thanks to the healthy nature of walking and its ease of execution, walking is well-recognized as an important physical activity for public health (Lee & Buchner, 2008). Increased step count helps reduce blood pressure (Lefferts et al., 2023), improves blood glucose control in people with type-2 diabetes (Moghetti et al., 2020), and has been found to decrease all-cause mortality (Rodríguez-Gutiérrez et al., 2024). Furthermore, higher degree of physical activity is found to improve mental health in young adults (Pascoe et al., 2020; Zhu et al., 2020). Interestingly, a meta-analysis has found that physical exercise intervention may help reduce smartphone addiction (Liu et al., 2019). For these reasons, I chose walking as the beneficial alternative activity in this chapter.

1.2 Smartphone as a gateway behavior to more addictive behavior

Schnauber-Stockmann and Naab (2019) investigate the process by which media habits are formed. They analyzed users' self-reported data of an app associated with the 2016 European soccer championship (the UEFA Euro App 2016) over a month and found that habitual app usage was formed independently of regularity or context stability of using the app (a common prerequisite for habit formation, e.g., Wood & Neal, 2016). They inferred that smartphone usage itself acted as a 'gateway behavior' that served as a stable context for the formation of other media usage habits, such as social media consumption. Because smartphone usage is considered a more stable and less reward-sensitive behavior (Schnauber-Stockmann & Naab, 2019), an implication for choosing smartphone usage as the 'undesired' habit to be reduced in the current chapter is that it is potentially more difficult to be changed by interventions, in general., this should serve as a more robust test of the effectiveness of the dual-target intervention.

While Chapter 1's intervention targeted social media consumption due to its various negative associations with well-being and productivity, social media apps are not the sole culprit in causing negative impacts on well-being and productivity; other usage, such as gaming, might be equally responsible (Chiang et al., 2019). These activities are increasingly packaged in a plethora of apps easily accessible on the smartphones. For example, in 2024, gaming apps alone takes up 12.3% out of more than 2.35 million apps on the Google Play store (42matters, n.d.) Thus, targeting overall screen time might benefit consumers more, because they are allowed some freedom to evaluate and prioritize what smartphone activities to reduce.

In sum, this chapter focuses on reducing overall smartphone usage, a broader behavior that comprises social media consumption and other activities that require the smartphones. In addition, walking is the alternative behavior that is incentivized alongside the reduction of smartphone usage in the dual-target (DT) condition, an activity that is beneficial, distanced from the smartphones (as compared to using Duolingo in Chapter 1), objectively trackable, and easy to do. Thus, the dual-target intervention is examined under a different context and pair of activities in Chapter 2. Next, I describe the design of the RCT and outline the associated hypotheses.

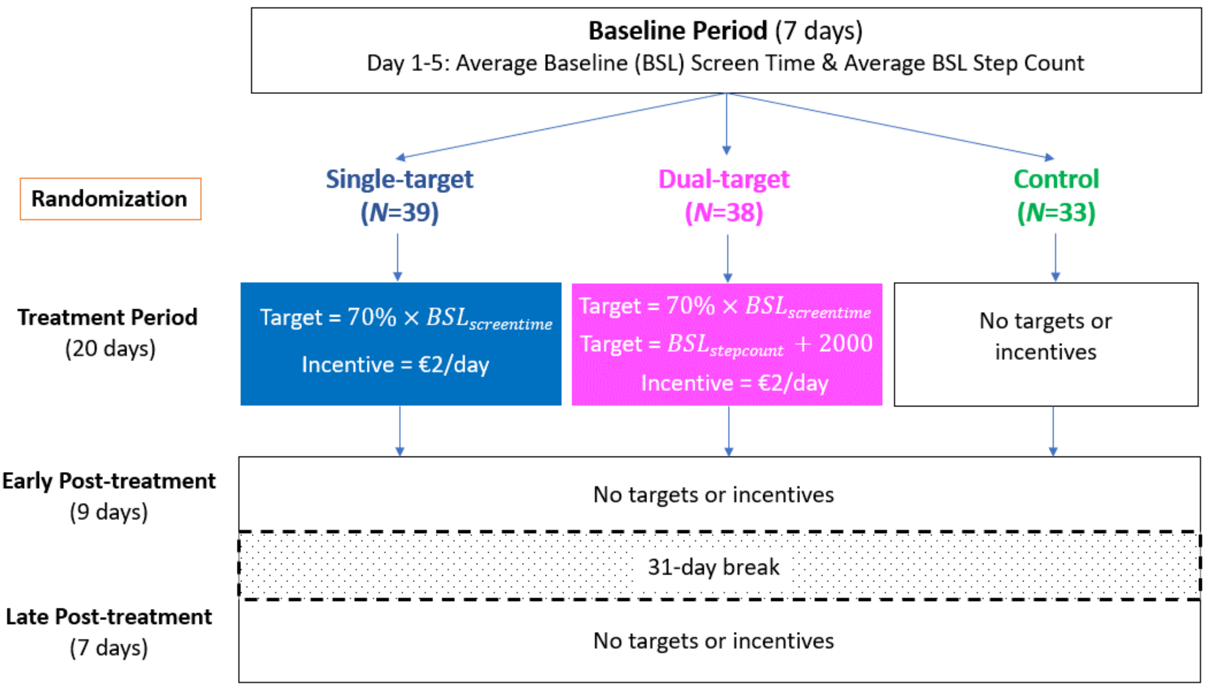
2. Methods

2.1 Study design

Chapter 2's RCT (4,098 observations across $N = 110$; 64 females; $M_{age} = 24.4$, $SD = 3.53$) was approved by the IRB at a large international university. The procedure was similar to Chapter 1, except for the study duration (seven weeks), number of post-treatment periods (early and late), type of behavior (mobile screen time[14] and daily step count[15]; see Appendix C1 for an example of a screenshot), and random assignment. The study was divided into four periods as illustrated in Figure 4.

Figure 4

Design of the RCT (Chapter 2)



[14] For iPhone users, Apple's inbuilt iOS Screen Time feature was used: https://support.apple.com/en-gb/HT208982. For Android users, an equivalent app, which could be downloaded free of costs, was used: https://play.google.com/store/apps/details?id=master.app.screentime.

[15] For iPhone users, Apple's inbuilt iOS Health app was used: https://support.apple.com/en-us/HT203037. For Android users, Google Fit, which could be downloaded free of costs, was used: https://play.google.com/store/apps/details?id=com.google.android.apps.fitness.

After the baseline period, I calculated subjects' screen time and step count, and randomly assigned subjects into three conditions: 1) Control (C), 2) Single-target (ST) treatment, 3) Dual-target (DT) treatment[16]. Subjects in the C condition (*N* = 33) had no targets nor incentives. Subjects in the ST treatment (*N* = 39) were paid €2 each day their screen time was at least 30% lower than their baseline usage during the treatment period. Subjects in the DT condition (*N* = 38) were paid €2 each day their screen time was at least 30% lower than their daily baseline usage *AND* their step count exceeded their baseline step count by 2,000 steps. I chose an increment of 2,000 steps because it is a manageable increase (Tudor-Locke et al., 2011) with positive effects on weight loss and blood pressure (Bravata et al., 2007).[17]

At the start of the treatment period, subjects were informed about their baseline average screen time and step count and provided with a handout outlining the benefits of reducing screen time and increasing step count. In the treatment conditions, I used the average screen time and step count over the first five days (out of seven) of the baseline period to determine subjects' targets. I informed subjects in both treatment conditions about their screen time targets. Only subjects in the DT condition were additionally informed about their step count target. In the early post-treatment period (Period 2), I removed the targets and incentives, but subjects continued reporting their screen time and step count for the next nine days. Thirty-one days later (or forty days after the end of the Treatment period),[18] I asked subjects for another screen time report which constitutes the late post-treatment (seven days).

[16] Balance checks for randomization can be found in Table C2.
[17] I did not choose a relative target for step count (e.g., 30% of baseline). This is because there was large variation in daily step count (it ranged from as low as 363 steps to as high as 21,428 steps). A relative target would have been too challenging for subjects with high baseline step count, and vice versa.
[18] Compared to Somasundaram et al. (2024)'s late post-treatment period which started 3 weeks after incentives and targets were removed, the late post-treatment period occurred even later, which let me examine the robustness of the post-treatment effects in a more stringent manner.

2.2 Surveys

Subjects completed a baseline and a post-treatment survey at the start of the corresponding periods. I measured demographic and mobile usage characteristics, as well as self-reported productivity and anxiety (from being without a smartphone). Based on past research (Cappelen et al., 2018), I included health-related measures such as how physically active subjects were, height and weight, sleeping pattern, number of days subjects felt tired, overall satisfaction with life and health (Appendix C7). I also measured beliefs about target achievement in the ST and DT condition before the incentive (treatment) period, and the perceived difficulty of meeting the target measured before and after the incentive period (Appendix C2). The baseline survey required subjects to confirm they had activated the screen time and step count apps by uploading a screenshot of each app. Additionally, after sending out instructions for each period, I confirmed if subjects had understood these instructions by asking them comprehension questions.

2.3 Hypotheses and analysis plan

I expected subjects in both treatments to reduce screen time during the treatment period. In addition, I expected subjects in the DT condition to increase step count in the treatment period. Importantly, I expected subjects in the DT condition who successfully achieved both targets to be more successful in sustaining a lower screen time in the early and late post-treatment period when compared to the ST condition. Subjects in the DT condition must fulfill two goals (rather than one) to receive the same incentive as subjects in the ST condition who only need to achieve one goal. The combination of two goals should be harder to achieve, and some subjects in the DT condition might give up on both targets altogether. However, I was encouraged by the result from Chapter 1 that subjects with higher dual-target achievement during the intervention period outperformed those with equivalent single-target achievement. This leads to the following hypotheses[19]:

H5: In the post-treatment periods, subjects with higher dual-target achievement during the treatment period will have lower screen time compared to those with higher single-target achievement.

To test H5, I use an instrumental variable approach with a 2-stage least squared regression (Wooldridge, 2013) to infer the causal effect of the endogenous target achievement (i.e. number of days that single-target, or dual-target is achieved) variable on the change in screen time between post-treatment and baseline periods (i.e., the outcome variable). An instrumental variable isolates the exogenous variation in the endogenous predictor (i.e., the part not correlated with the error term) that significantly correlates with the outcome variable. I take advantage of the random assignment into conditions as the instrument variable to estimate the unbiased coefficents of single/dual target achievement on the outcome variable. The relevant equations are presented in section 3.1. The analyses' results are presented in Table 5.

H6: In the post-treatment periods, subjects in the DT condition who achieved higher step count during the treatment period will have lower screen time compared to those in the control condition.

Similar to H5, to test H6, I use an instrumental variable approach to infer the causal effect of the endogenous step count during the Treatment period variable on the change in screen time between post-treatment and baseline periods (i.e. the outcome variable). I take advantage of the random assignment into conditions as the instrument variable to estimate the

[19] Unlike Chapter 1's study, this study was not pre-registered. A reason for this is that I was not sure about target adherence, and the intervention's effect was likely to depend on it, as was the case for Chapter 1.

unbiased coefficents of step count during the Treatment period on the outcome variable. The relevant equations are presented in section 3.2. The analyses' results are presented in Table 6[20].

3. Analyses and Results

The average screen time during the baseline, intervention, and post-treatment periods of all subjects are reported in Table C3.1 in the Appendix. I observe that ST and DT conditions had a lower usage than the control group in the intervention and post-treatment periods. However, there was no significant difference between the treatments.[21] Below I test the stated hypotheses, conditional on target achievement.

3.1 Adherence to dual-target intervention predicts post-treatment outcomes better than adherence to single-target intervention

An activity must be carried out repeatedly to form a habit (Harris & Kessler, 2019). Therefore, *adherence* to the target(s) is necessary for the reduction in screen time from pre- to post-treatment. I operationalize adherence as the number of days that target(s) was(were) achieved during the treatment period. Based on Figure 5, I observe that adherence to the

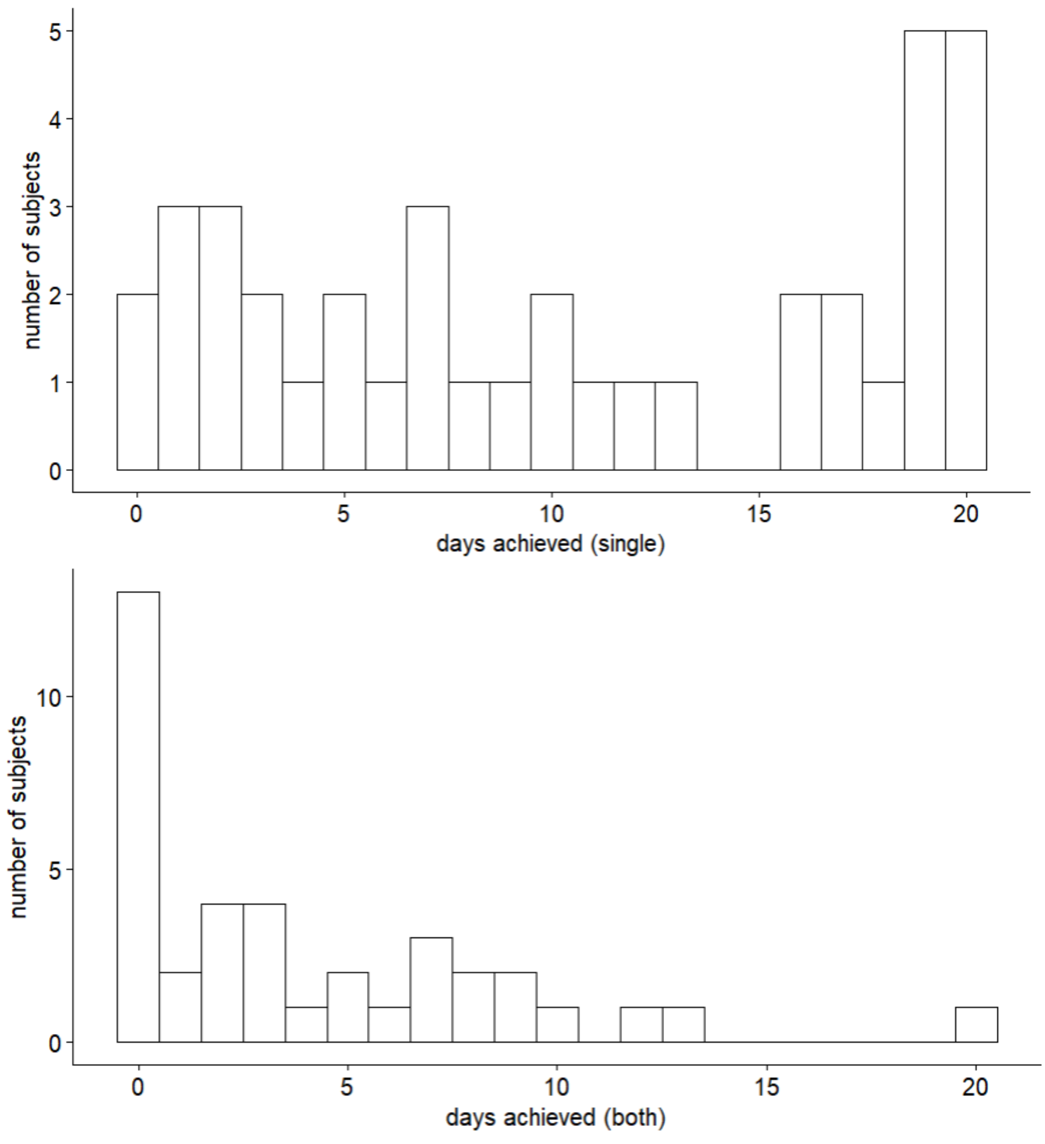
[20] I estimated four models which correspond to four subsets of data, so as to compare ST vs. Control conditions, and DT vs. Control conditions, separately for the Early Post-treatment and Late Post-treatment periods.
[21] I showed in Appendix C4 that this is the conclusion a simple dif-in-dif OLS regression would draw.

treatment was lower in the DT than in the ST condition, possibly because subjects in the DT condition found targets more difficult to achieve than their ST counterparts.

To test hypothesis H5, I use target achievement to predict post-treatment effects.

Figure 5

Target achievement is higher in ST condition (top) than DT condition (bottom).



However, as target achievement is endogenous, estimates of the OLS regression might be biased (Angrist & Imbens, 1995). Therefore, I utilize an instrumental variable (IV) approach using 2-SLS regression (Wooldridge, 2013). Previous RCTs have utilized this method to address the issue of treatment compliance (Penrod et al., 2009; Somasundaram et al., 2024). I chose *Condition* as the IV. Note that *Condition* is a good IV as it satisfies two criteria: it predicts the days of target achievement in the treatment period P1 (relevance), and it is not correlated with the error term ϵ_i due to random assignment and because subjects in all conditions do not have a task in the post-treatment (exclusion restriction). I first ran the following two regressions:

(1)

Second stage: $$PostTx\ usage\ change_{it} = \alpha + \beta_{single}\left(No.days\ achieved(single)in\ P1\hat{}_i\right) + \sum_{j=2} \beta_j\ (Other\ controls) + \epsilon_i$$

First stage: $$No.days\ achieved(single)\ in\ P1_i = \alpha_1 + \beta'(Condition_i) + \sum_{j=2} \beta_j\ (Other\ controls) + \epsilon'_i$$

(2)

Second stage: $$PostTx\ usage\ change_{it} = \alpha + \beta_{dual}\left(No.days\ achieved(dual)in\ P1\hat{}_i\right) + \sum_{j=2} \beta_j\ (Other\ controls) + \epsilon_i$$

First stage: $$No.days\ achieved(dual)\ in\ P1_i = \alpha_1 + \beta'(Condition_i) + \sum_{j=2} \beta_j\ (Other\ controls) + \epsilon'_i$$

where $PostTx\ usage\ change_{it} = PostTx\ usage_{it} - baseline_i$. Controls include gender, age, and operating system (OS). The results of the second stages of each regression are shown in Table 5.

Table 5

Instrumental Variable Regressions (2nd stage). Dual-target achievement predicts post-treatment screen time reduction above single-target achievement.

	Dependent variable = Post-treatment usage - Baseline	
	Model 1	Model 2
Predicted number of days achieved in the Treatment Period (P1) - single target (β_{single})	-0.12*** (0.02)	
Predicted number of days achieved in the Treatment Period (P1) - dual target (β_{dual})		-0.21*** (0.04)
Gender (1 = Male, 0 = Female)	-0.06 (0.13)	0.05 (0.12)
Age	-0.03** (0.02)	-0.04** (0.02)
OS (1 = Android; 0 = iOS)	0.24* (0.14)	0.09 (0.13)
Observations	1326	1326
R^2	0.08	0.02
Adjusted R^2	0.08	0.02

Note 1: * $p < .1$; ** $p < .05$; *** $p < .01$. Baseline corresponds to average mobile usage during the baseline period. Standard errors (in parentheses) are robust and clustered at the subject level.

Note 2: In the 1st stage (Appendix C5), Condition significantly predicts target achievement; Dual-target condition had significantly more dual-target achievement days than either Single-target or Control condition.

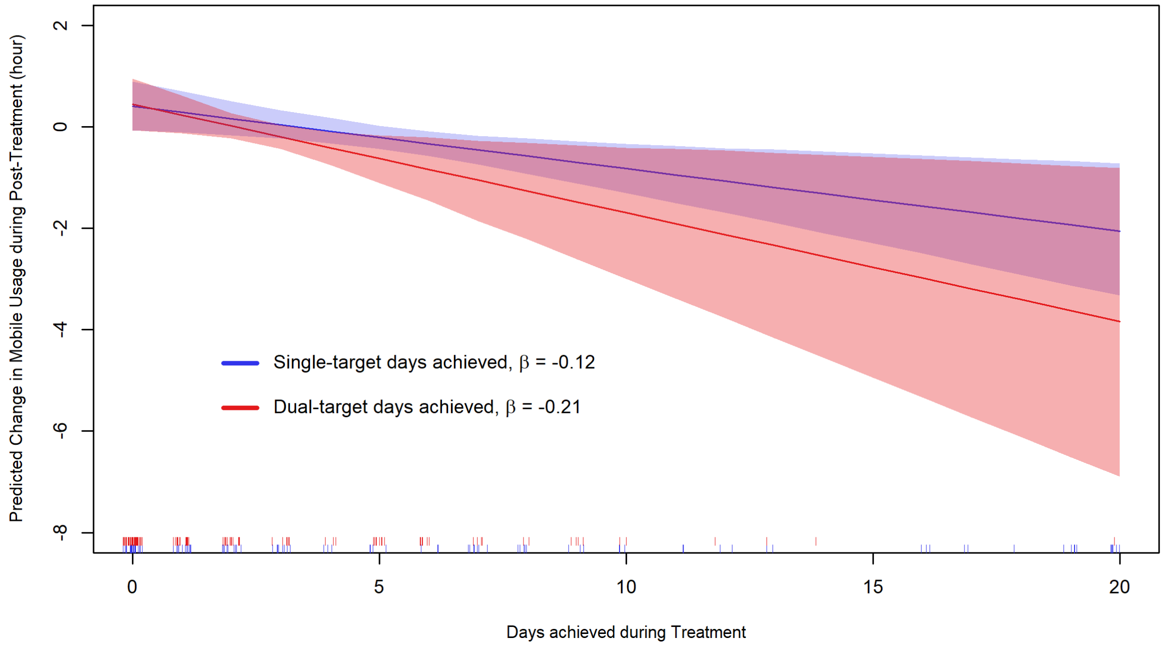
Higher adherence to either single-target or dual-target results in lower screen time during the post-treatment period relative to the pre-treatment period. However, the coefficient of dual-target adherence is larger than of single-target adherence (β_{dual} = -0.21 vs. β_{single} = -0.12).

To compare the two coefficients, I conducted a Wald Chi-square test of a linear hypothesis that the system of model 1 and 2 is different from a system which assumes that the two coefficients were equal. The result of this test is significant, *Wald* χ^2 (1, 2722) = 5.73, *p* = .017. Relative to single-target adherence, dual-target adherence helps maintaining a lower screen time after a habit formation intervention has ended (Hypothesis 5)[22]. To illustrate this finding, a post-estimation graph of the difference between the two coefficients is shown in Figure 6. The slopes represent marginal reduction in smartphone usage in the post-treatment period compared to the baseline. The slope for dual-target achievement is steeper than that of single-target achievement.

[22] When willingness to increase step count (measured at baseline) was included as an additional control, the coefficients β_{dual} and β_{single} and their associated significance levels did not change (Appendix E3). This shows that even when intrinsic motivation to walk more is accounted for, the effects of the treatment conditions on post-treatment change remains stable.

Figure 6

Post-estimation of 2SLS instrumental regression (Chapter 2)



3.2 Step count achievement in the treatment period predicts screen time in the post-treatment periods

To investigate if engaging in a desired activity (e.g., walking more) helps reduce an undesired habit (e.g., looking at the phone), I conducted an instrumental variable regression. Change in post-treatment mobile usage with respect to the baseline was the dependent variable. I regressed it on the average step count in the treatment period (Period 1 – P1) as the independent variable. In the first regression, I compare the ST condition with the control condition, and in the second regression, I compare the DT condition with the control condition. The formulae of these regressions are as follows:

Second stage:

$$PostTx\ usage\ change_{it} = \alpha + \beta_1 (\ Avg.\ step\ count\ in\ P1\hat{}_i) + \sum_{j=2} \beta_j\ (Other\ controls) + \epsilon_i$$

First stage: $Avg.\ step\ count\ in\ P1\hat{}\ _i$

$$= \alpha_1 + \beta'(Condition_i) + \sum_{j=2} \beta_j\ (Other\ controls) + \epsilon'_i$$

where $PostTx\ usage\ change_{it} = PostTx\ usage_{it} - baseline_i$. Controls include gender, age, and operating system (OS) type.

The second-stage results of the regressions are presented in Table 6. Subjects in the DT condition had a higher average step count than control subjects (Appendix C6). Further, an increase in step count predicted lowered screen time in both post-treatment periods (*p*s < .01). In contrast, subjects in the ST condition did not differ from control subjects on average step count during the treatment period (Appendix C6), and their step count did not significantly predict mobile usage in the post-treatment periods (*p* = 0.13 and *p* = 0.08 for early and late post-treatment period, respectively). Thus, I find evidence to support Hypothesis 6.I

Table 6

Instrumental Variable Regression (2nd stage). Higher step count during Treatment period associates with higher reduction only for the DT condition, but not the ST condition.

	Dependent variable = Post-treatment usage - Baseline			
Sub-sample	ST condition vs Control		DT condition vs Control	
	Early post-Tx	Late post-Tx	Early post-Tx	Late post-Tx
Step count in Treatment period	-0.001	-0.001	-0.001**	-0.0003*
	(0.001)	(0.001)	(0.0002)	(0.0001)
Gender	1.27	0.78	-0.94**	-0.51
	(0.78)	(0.55)	(0.30)	(0.26)
Age	-0.03	0.17	-0.05	-0.001
	(0.06)	(0.08)	(0.03)	(0.04)
OS	-1.12	-1.36	-0.11	0.39*
(1 = iPhone; 0 = Android)	(0.92)	(0.81)	(0.24)	(0.28)
Observations	555	341	538	318
R^2	-6.04	-2.67	-1.30	-0.19
Adjusted R^2	-6.09	-2.72	-1.32	-0.20

Note: * *p* < .05; ** *p* < .01; *** *p* < .001. Baseline corresponds to average mobile usage during the baseline period. Standard errors (in parentheses) are robust and clustered at the subject level.

Finally, I examine the effect of each condition on well-being and productivity. The detailed list of measures (and their descriptive statistics) can be found in Appendix C2. Overall, I did not find any difference among the three conditions on well-being and productivity.

4. Discussion

Chapter 2 is an RCT that has improved on the non-experimental design and thereby the inferences of Chapter 1. Subjects in each treatment condition (ST or DT) could only earn incentive for their assigned treatment, unlike in Chapter 1 where subjects were free to choose either ST or DT to achieve for each day. Importantly, I equated the daily incentive amount for both treatment conditions, resulting in a more conservative test to pitch the effectiveness of the DT condition against the ST condition. I also tracked subjects' consumption over a longer post-treatment period because sustaining intervention effect in the long-term is the focus of the dual-target approach. Finally, Chapter 2 extended Chapter 1's findings to a different pair of behaviors: reducing smartphone consumption and increasing step count.

Chapter 2 shows that subjects with more days of adherence to their assigned target(s), be it single-target or dual-target, had more sustained reduction in their screen time in the post-treatment period relative to their baseline usage. Importantly, this reduction was higher in DT subjects than in ST subjects. Finally, I found that DT subjects had higher step count during the treatment period compared to control subjects (this difference is not significant between ST subjects and control subjects), and using instrumental variable regressions, I found that higher step count during the treatment period results in more sustained reduction of screen time in the post-treatment period relative to baseline screen time.

Thus, while engaging in an alternative activity like walking might require more effort, it can help sustain the effect of behavior change interventions even after incentives are discontinued. This should provide consumers with a valuable tool to adjust their consumption goals in the long run.

Chapter 2 did not find an improvement in well-being and productivity with the reduction of smartphone screen time. This might be because the smartphone as a medium might not be the root of problematic usage, but the specific content that is conveyed via smartphones is (Lowe-Calverley & Pontes, 2020). This further reinforces the idea that smartphone usage is a ‘gateway behavior’ to the formation of other media habits (Schnauber-Stockmann & Naab, 2019), for example, social media app usage. Because I did not record usage of individual apps, I was not able to tell on a finer level the nature of the reduced screen time usage. Future studies should track this information as well so that a definitive answer could be reached on whether reduction of overall smartphone screen time or particular app usage is beneficial for mental health.

While Chapter 2’s RCT resolves a major design problem of Chapter 1’s longitudinal study, it is still difficult to tell whether the behavior change is caused by just receiving the incentive or by faithfully adhering to the dual-target intervention. Future studies may consider incorporating a placebo group with only incentive in order to rule out this explanation (similar to Study 2 in Charness & Gneezy, 2009, where a separate control group that received the full amount of monetary reward without having any kind of target was included).

Finally, Chapter 2’s RCT tested the effectiveness of a dual-target intervention with an alternative activity (i.e., walking) that is more distanced from the habitual behavior to be reduced (i.e. smartphone usage) compared to the pair of activities examined in Chapter 1 (i.e., social media usage and learning a language with Duolingo). Although this is likely to decrease cue exposure to initiate the undesired behavior, this effect of ‘cue distance’ between two activities has not been operationalized and tested in the current research. This is a promising and novel direction that future research may undertake, where only cue distance should be varied between two treatment conditions. For example, scholars could keep the undesired behavior constant between the two treatments, and vary only the alternative activity in terms of its dependency on the smartphone to be executed.

General Discussion of Chapter 1 and 2

While consumers often express a desire to cut down on mobile screen time, habitual smartphone usage makes this goal challenging (Anderson & Wood, 2021). The habitual nature of smartphone usage is a result of both the positive reinforcement it offers, and the way it has become deeply embedded in daily life. This makes efforts to reduce screen time not just a matter of willpower but also a challenge to undo well-established patterns of behavior. Following periods of reduced usage, consumers typically relapse into previous consumption patterns—a phenomenon known as the triangular relapse pattern (Wood & Neal, 2016). The current research addresses this by substituting the time spent on the undesired habit with an alternative, beneficial activity. For sustained target achievement, an incentive scheme that both reduces excessive smartphone usage and promotes the adoption of a desirable activity (e.g., language learning in Chapter 1 or physical activity in Chapter 2) proves more effective in the long run than one solely focused on reducing smartphone usage.

The current intervention was more effective if the alternative activity occupied a higher proportion of time of the undesired habit i.e., if the activity had a higher "target replacement ratio" (TRR). I found that TRR and dual-target achievement jointly predicted lower social media usage. The more consumers filled the time freed up by reducing an undesired habit with an alternative activity (i.e., learning a new language) the more successful they were in maintaining lower social media consumption in the post-treatment period, compared to those who only reduced their mobile usage. Furthermore, the alternative activity may serve as a reminder that consumers gain valuable time as a resource to be spent on other, perhaps more fruitful, endeavors.

Consumer heterogeneity plays an important role. The current intervention was more effective for individuals with higher target adherence during the treatment period. Since adherence is crucial for habit formation, in Chapter 2, I evaluated the relative effectiveness of the dual-target intervention compared to the single-target intervention contingent on adherence to the treatment. With an instrumental variable regression, I showed that consumers with more dual-target achievement reduced their screen time in the post-treatment period more than those with more single-target achievement.[23]

Subjects did not continue to engage in the alternative activities in the post-treatment period. I believe this is due to the following reasons. Dual-target adherence was lower compared to single-target adherence in both studies: in Chapter 1, single-target achievement days were almost four times higher than dual-target achievement days; in Chapter 2, dual-target achievement was lower (Figure 5) in the DT condition. Unless subjects perform the alternative activity repeatedly, they are unlikely to develop a stable habit which extends in the post-treatment period. Second, the alternative activities, though beneficial, might not be behaviors that subjects wanted to engage in permanently due to individual preferences and tastes.[24] Nevertheless, having all subjects perform the same alternative activity in each study allowed for an objective and comparable measurement. Instead of relying on self-report, I was able to objectively measure both mobile usage and the alternative activity in a longitudinal fashion via mobile tracking applications. This empirical control is valuable as it complements previous

---

[23] This result is consistent with Chapter 1, which also found that subjects with more dual-target achievement (instead of single-target) tended to have a lower social media usage during the post-treatment period relative to their baseline.

[24] In Chapter 1, I also asked subjects at the end of the intervention, whether they continued to use Duolingo after they no longer received incentives. About a third of subjects affirmed this (Appendix Table B10.1). Subjects provided various reasons as to why they thought Duolingo might (or might not) have helped them to reduce social media usage (Appendix Table B10.2).

research on replacement interventions which allowed subjects to freely choose the alternative activity (e.g., Esmaeili & Ahmadi, 2018) without obtaining objective measurements.

Subjects who reduced their social media consumption in Chapter 1, showed improvements in subjective well-being. However, in Chapter 2, I found no difference in well-being between subjects incentivized to reduce their overall smartphone usage and control subjects. The smartphone as a medium might not be the root of problematic usage, but the specific content that is conveyed via smartphones (Lowe-Calverley & Pontes, 2020). Chapter 1 and 2's findings are consistent with this. Consumers' well-being may benefit more from reducing social media than from reducing overall mobile screen time.

In line with my reasoning that consumers self-signal that they are less addicted to social media when engaging in an alternative activity, individuals with high dual-target (but not single-target) achievement showed a significant reduction in perceived social media addiction (path a), and lower addiction resulted in lower post-treatment consumption relative to baseline. However, the corresponding mediation test (with perceived social media addiction as the mediator) did not yield a significant indirect effect (Appendix E2). It is reasonable to speculate that the self-signalling process via perceived addiction to social media might be further contingent on other factors. Thus, the proposed mechanism needs further research to confirm.

Practical implications

The approach to problematic smartphone and social media usage of the first 2 chapters can be of interest to policymakers, consumers, social marketers, and companies. For policymakers, dual-target interventions might be more cost-effective. These types of interventions not only encourage consumers to reduce an undesired habit, but also motivate them to engage in new beneficial activities such as exercising. This "2-for-1" approach can deliver added long-term advantages over single-target interventions.

For consumers, focusing on alternative activities might be a useful strategy when trying to break an undesired habit, such as excessive smartphone usage, for good. However, target adherence is crucial for success. Therefore, consumers must carefully consider setting realistic goals, creating achievable schedules, and finding ways to stay motivated and accountable. Dual-target intervention might also prevent compensatory addictions. Often consumers reduce one addictive behavior by simply substituting it with another. For example, Melumad & Pham (2020) show that recent smoking-quitters grew more attached to their smartphones out of a need for stress relief. Dey et al. (2019) found that problematic smartphone usage in young men was negatively associated with both frequent cannabis use, and cigarette smoking. Dual-target interventions may break such compensatory cycles by replacing the unhealthy habit with a healthy alternative which heightens the time barrier to entry for compensatory addictions.

Besides public policy programs and individual efforts to change behavior, the current studies reveal a valuable market gap for social marketers to promote beneficial alternatives to consumer segments striving to reduce their consumption habits (Lee & Kotler, 2015). This marketing opportunity arises from the need for consumers to fill the void left by reducing undesired habits. Chapter 1 and 2 demonstrate how this void can be addressed, utilizing alternative productive apps (e.g., Duolingo) and personal enrichment programs (e.g., step count challenges).

The current findings are also relevant to technology companies. Even though reducing consumption of social media might hurt ad revenue of technology firms, it is perhaps in their best interest to consider addressing the backlash from the public generated by over-consumption of social media. For example, this matter has garnered enough attention to be presented to the U.S. congress as a bill that proposes several measures to reduce social media usage such as imposing time limits or banning of features like auto-playing videos (SMART Act, 2019). There has also been evidence of "social media fatigue", a phenomenon where consumers are tired out by excessive interactions on social media and quitting altogether (Ravindran et al., 2014; Fu et al., 2020; Sheng et al., 2023). Supporting consumers in reducing their usage might help mitigate these problems, benefiting both companies and customers in the long run. In addition, from a business ethics perspective, it may be beneficial for technology companies to harness social media as tools to empower consumers. Companies might benefit from being viewed as consistent with their purported prosocial mission statements.[25] This is especially important in light of recent revelations that social media sites might be manipulating consumers to make them spend more time on the sites.[26]

Furthermore, some worries may arise from the abuse of the dual-target intervention by profit-driven businesses, who might entice consumers to increase 'bad' consumption and simultaneously make them disengage from other beneficial activities. The current two studies differ from this scenario in a subtle but important way: because they are explicitly making an 'offer' to consumers to reduce an undesired behavior and increase a beneficial behavior, consumers are fully informed about what they must do to obtain the incentive. Thus, large-scale businesses that want to offer products/services in a similar manner must also inform their consumers up front about the details of these offers or promotions, or risk running into trouble with consumer protection laws or even antitrust laws, which prohibit companies from using unfair practices to boost market power and erecting barriers to entry not stemmed from natural competitive advantages, such as making products addictive and creating consumer dependency (Day, 2021; Rosenquist et al., 2021).

[25] Meta's (formerly Facebook) mission statement (Meta, 2022): "… to give people the power to build community and bring the world closer together. Our products empower more than 3 billion people around the world to share ideas, offer support and make a difference."
YouTube's (a subsidiary of Alphabet Inc.) mission statement (YouTube, n.d.): "… to give everyone a voice and show them the world. We believe that everyone deserves to have a voice, and that the world is a better place when we listen, share and build community through our stories."
[26] For example, a whistleblower revealed that Facebook promoted certain harmful political content that increases engagement and the advertisements shown on the site (Dent, 2021).

Limitations and future research

While both chapters demonstrate the potential for dual-target intervention to maintain an intervention's long-term effect, evidence supporting the potential mechanism is still lacking. While Chapter 1 shows that a consumer who performs more days of the alternative activity (Duolingo usage in this case) may self-signal that they are less addicted to social media, results of the mediation test (Appendix E2) indicated a non-significant indirect effect via perceived addiction to social media, and a strong direct effect. Regarding the former, as previously discussed, the mediation test should be confirmed by a future study with a bigger sample size.

Beside, a strong direct effect from dual-target achievement indicates that there could still be other possible psychological mechanisms that are at play. One possible explanation that might interest future research is in line with the ego depletion theory (Baumeister & Vohs, 2007; Baumeister et al., 2008): consumers engaging in the alternate activity have less time (and therefore, less cognitive bandwidth) to accommodate the undesired habit. Thus, relapse would be prevented as their cognitive capacity is already depleted. Another mechanism could be mental accounting (Thaler, 1985; Heath & Soll, 1996; Rajagopal & Rha, 2009), where the alternative activity may serve as a reminder that consumers gain valuable time as a resource to be spent on other, perhaps more fruitful, endeavors.

Target adherence is key to habit formation, whether it is for the focal activity (e.g., reducing social media usage) or the alternative activity (e.g., language learning). The effectiveness of the dual-target approach might strongly depend on how it is implemented to encourage target adherence. Future research should explore different incentive schemes to maximize adherence. For instance, gamification approaches (e.g., streaks) might be particularly useful (Eisingerich et al., 2019). Consumers may also use commitment devices which require them to stake their money to modify a behavior (Ashraf et al., 2006). Future research should investigate how effective different incentive structures that consumers set for themselves or for others are (e.g., parents incentivizing children with pocket money for reducing smartphone usage). Moreover, longer treatment periods might allow consumers to adjust to the demands of the intervention, thereby increasing the frequency of target achievement and adherence (Harris & Kessler, 2019; Liu et al., 2019).

Identifying consumer segments that benefit most from dual-target treatments in different contexts would be interesting and consistent with the framework of social marketing (Lee & Kotler, 2015). For example, teenagers and adolescents are the biggest and most vulnerable demographic group to engage with mobile technology. Numerous research has linked problematic smartphone and social media usage with lower well-being and productivity in this population (e.g., McNamee et al., 2021; Hisler et al., 2020; Levenson et al., 2017), and there has been a persistent call for interventions to prevent these negative outcomes. Future research should continue to investigate the effectiveness of dual-target interventions in such vulnerable populations.

It is crucial to identify the characteristics of alternative activities that can effectively substitute for the undesired habit. The choice of beneficial activity needs to consider not only the difficulty, but also, how well it can distract the consumer from the cues that activate the undesired habit (Laibson, 2001; Adriannse et al., 2011). For example, brisk walking (Chapter 2) distances consumers from the cue (their smartphones), thereby shielding them from stimuli such as push notifications that ignite their desire to use the phones. In contrast, using Duolingo (Chapter 1) on the same mobile device that hosts frequent social media usage arguably does not have the same distancing effect. While the two studies were motivated by this cue activation principle, they did not directly test the variation of cue strength based on how separated they are from a consumer. This could be an interesting direction for future research.

Future research should explore alternative activities that closely match the consumption pattern of the undesired habit. Smartphone usage occupies many small "pockets" of time throughout the day, and 50% of usage sessions start within 3 minutes of the previous one (MacKay, 2019; Oulasvirta et al., 2012). An ideal alternative activity should prevent consumers just in time from using their smartphones on a whim. Thus, alternative activities that can be divided into smaller spontaneous sessions might be more beneficial.

During the dual-target intervention, both good and bad behaviors are targeted, and not just either behavior alone. Thus, as previously mentioned, this approach does not rely on a spillover effect to work. However, a 'tertiary' spillover is still possible. This spillover could manifest as the change in activities other than the two targeted behaviors. For example, a person who decreases social media consumption and uses more Duolingo during the incentivization period might also engage in additional language-learning activities such as going to more language exchange sessions, or substitute his/her craving for social media with other social activities. Depending on the nature of such activities, they can either buttress or undermine the intervention. In any case, since this is likely to introduce confounds to the intervention, future research should track other activities of each subject, for example, by means of a daily activity diary.

Currently, both interventions in Chapters 1 and 2 set a relatively short treatment period. Recent research in habit formation found that the duration for habit formation varies across both behavior types and individual differences (Buyalskaya et al., 2023). This might have contributed to the non-signficant change in well-being measures in Chapter 2's RCT. Future research needs to allot more treatment days and, at the same time, keeps intervention cost low so that interventions are scalable.

Finally, the sample consists primarily of students, which may limit the external validity of the findings. Additionally, the sample is predominantly female, raising concerns about representativeness. Future research might tap on a more diverse, non-student sample to examine the generalizability of the dual-target intervention's effect to other demographics.

This page is intentionally left blank.

Chapter 3: A Meta-Analysis of Interventions for Problematic Smartphone and Social Media Usage

Abstract: *Many people would like to reduce their smartphone or social media usage in the hope of improving productivity and wellbeing but fail to do so due to the addictive nature of many smartphone applications. Understanding how to encourage consumers to change a habitual behavior is important for policy makers. In this chapter, I investigate strategies to reduce smartphone and social media addiction with a meta-analysis examining what types of strategy work best. Specifically, I compare the effectiveness of interventions -- ranging from simply setting usage restriction goals to choice architecture based digital nudges (such as design friction tools like the grayscale mode to restrictive time limits). Within this framework, I also address the critical question of consumer heterogeneity (i.e., which interventions work best for which users and why). With this chapter, I aspire to guide future research of digital addiction interventions, and provide actionable recommendations for policy-makers, advocacy groups and consumers to instigate cost-effective and lasting behavior change in the domain of smartphone and social media usage.*

1. Introduction

In 2007, the first iPhone was released, and quickly popularized the touchscreen smartphone format that is the paradigm of today's mobile devices. Just a year later, Facebook launched its first mobile application on the iPhone, marking the inexorable trend of social media consumption on smartphones. While the burgeoning of smartphones and social media in the past 15 years has benefited important aspects of life such as work and leisure, problems related to health, privacy, and social relationships have emerged and sparked a debate among researchers: one side insists that social media and smartphones are harmful (e.g., Twenge et al., 2018; Twenge & Campbell, 2019), while the other says that it is not (e.g., Coyne et al., 2020), or there is very weak evidence against it (Orben & Przybylski , 2019), and interventions are therefore not warranted.

Policymakers, however, did not wait for the debate to settle. They have conducted inquiries into giant technology companies' operations of the biggest social media brands (United States Senate Committee on the Judiciary, n.d.). Some have even initiated the legislation process to counteract the negative effects of social media consumption. For example, a bill presented to the U.S. congress proposed several measures to reduce social media usage such as imposing time limits or banning a host of addictive features implemented in social media apps (SMART Act, 2019). Smartphone use, as an inextricable part of social media consumption, has also been targeted. For instance, the government of France has hinted at the possibility of imposing a minimum-age for smartphone use, in light of a state-commissioned report's recommendations (Chrisafis, 2024).

State-actors will likely benefit from evidence-based advice on how on how to best regulate and to what extent they should curb social media and smartphone usage. In this aspect, many research articles have synthesized evidence based on correlational studies to try to make sense of the relationship between problematic smartphone/social media use and well-being and performance measures (Valkenburg, 2022). However, research synthesis on interventions is lacking. In addition, there is no research that has attempted to compare across broad categories of interventions. The two meta-analyses (Malinauskas & Malinauskiene, 2019; Liu et al., 2019) and one systematic review (Radtke et al., 2022) that examine causal evidence (interventions) on this topic suffer from both scope and methodological problems, which limit their overall impacts on research and practice.

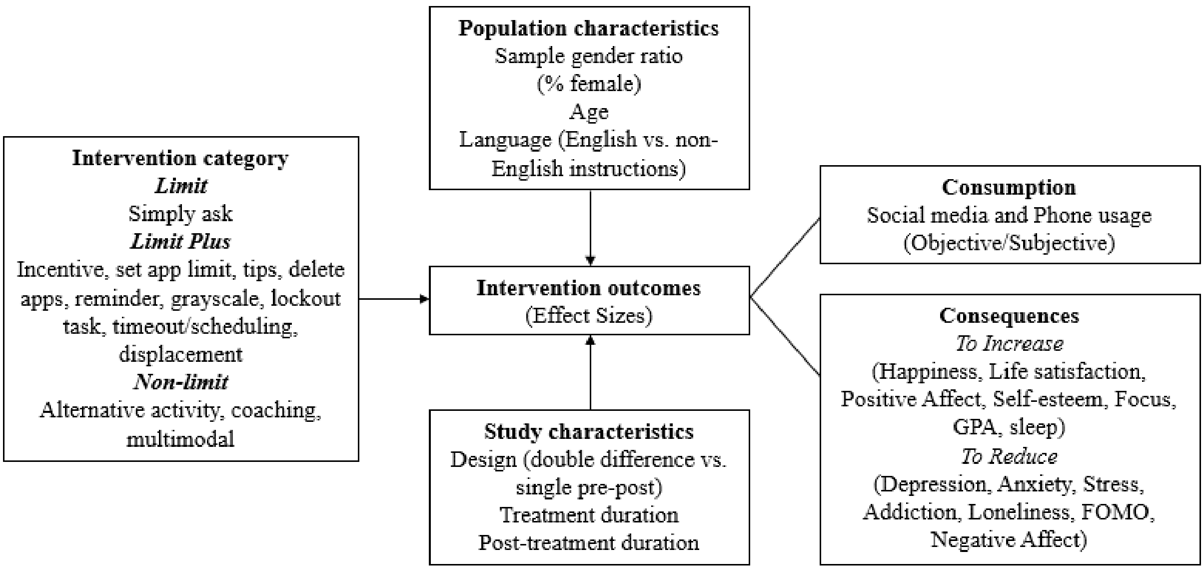
Thus, Chapter 3 aims to provide a holistic overview and synthesis of research with a focus on interventions for problematic smartphone and social media usage, by means of an exhaustive literature search and meta-analytical model. I identify what strategies have been evaluated for their effectiveness in reducing problematic usage, and how they measure up against one another in two types of outcomes: consumption (i.e., smartphone screen time, social media usage time), and consequence (i.e., addiction, well-being, ill-being, sleep quality, productivity) improvement. Using a grounded approach, I classified interventions into three broad types: *Limit*, *Limit Plus*, and *Non-Limit*, so that they can be meaningfully compared in meta-analytical models. I expanded the scope of my analyses by examining the influence of relevant study characteristics (e.g., intervention design, treatment duration) and population characteristics (e.g., age, gender) on interventions' effect sizes.

The remainder of Chapter 3 is organized as follows: I describe my grounded approach, data collection, and coding procedures in section 2. Sections 3 and 4 describe the meta-analyses and their results for the consequence and consumption outcomes, respectively. Finally, I discuss Chapter 3's findings and give recommendations for future research.

2. Conceptual Framework through a Grounded Approach

During my initial literature scoping search, I found three attempts that have been made to summarize interventions for reducing social media and smartphone usage (Radtke et al., 2022; Malinauskas & Malinauskiene, 2019; Liu et al., 2019). While these seminal works are informative, they only focused on one specific class of interventions (e.g., cognitive-behavioral treatment; digital detox). In addition, other systematic reviews using correlational research merely gave suggestions on what strategies should be evaluated (e.g., Busch & McCarthy, 2021). None has attempted to create a broad framework with which one may compare strategies that are based on different theoretical backgrounds. I address this gap by deriving a conceptual framework, using a grounded approach (Strauss & Corbin, 1990), that allows for comparison of interventions. Figure 7 shows my conceptual framework. Below, I describe my grounded methodology.

Figure 7
Conceptual framework



2.1 Data collection

2.1.1 Data sources and search terms

I anchored my literature search around two related research questions: What strategies for reducing problematic smartphone and social media usage have been evaluated? Which one works best? To what extent do study and demographic factors influence affect their effectiveness?

I used the SPICE framework (Booth, 2016) to help me come up with relevant search terms regarding the setting, population, intervention, comparison, and evaluation of the research topic. The search terms were agreed upon with my co-authors (Appendix D1). With this set of terms, I queried three different databases (PubMed, Web of Science, Scopus) for journal articles as far back as 2008, which was the year that the Facebook mobile app was launched on iPhones, until the end of 2022. I also examined the references from two relevant meta-analyses (Liu et al., 2022; Malinauskas & Malinauskiene, 2019) and one systematic review (Radtke et al., 2022) that exclusively looked for causal design studies.

I added manually to my final pool of selected papers those that were referred to me by colleagues. In addition, while reading the selected papers, I actively looked for citations that fit my inclusion criteria.

2.1.2 Inclusion criteria

To be included, a study must be an intervention with a treatment period lasting at least one full day and with pre-treatment and post-treatment data collection rounds. The following study types are excluded: (1) cross-sectional studies or longitudinal designs that occur within the same day; (2) longitudinal studies without any form of intervention; (3) interventions that aim to prevent (instead of reducing) smartphone addiction. If the intervention specifically targets social media, it has to be social media usage via smartphones in order to be included[27].

In addition, studies must measure phone screen time/social media time independently from other types of screen time. If they did not do this, their interventions must specifically target these two usage types without targeting other screen time types (e.g., personal desktop/laptop computers, televisions).

Appendix D2 shows the PRISMA diagram which describes the literature search process. Overall, the final dataset consists of 307 effect sizes from 72 studies published in 59 articles (Total number of participants $N = 15{,}124$). Appendix D3 summarizes these 59 articles.

2.2 Coding

2.2.1 Intervention categories

First, I coded each intervention based on how the corresponding authors labelled or described the treatment condition. For example, if the intervention was to “reward $2.5 per hour of social media time reduction” (Allcott et al., 2022), it was coded as “bonus” (which is the same as how the authors called it); if the intervention was to “ask to limit social media use to 60 minutes a day” (Thai et al., 2021), it was coded as “limit”. In addition, I attempted to code all interventions as ‘nudges’ based on the nudge categorization by Sunstein (2014) in order to have a better overview of the interventions.

After the first round of coding, I discussed with my co-authors on how to sort the initial codes into more general categories. Taking into consideration the balance between intervention specificity and the number of effect sizes available in each category to conduct a meta-

[27] About 44% of smartphone usage time is spent on social media (Kemp, 2021).

analysis[28], we agreed on three broad categories of interventions: *Limit*, *Limit Plus*, *Non-Limit*[29]. Within each of these categories, where possible, I created sub-categories that better describe the interventions[30].

The *Limit* category consists of interventions in which participants were asked to simply restrict their smartphone/social media usage, or abstain from it altogether, without any other special instruction (e.g., Mosquera et al., 2019; Vanman et al., 2018).

Limit Plus interventions are more sophisticated. In addition to telling participants to reduce their smartphone/social media usage, they employed other methods to make it more likely for participants to follow that request. These methods might be using monetary incentives (e.g., Allcott et al., 2022), implementing design friction on smartphones (e.g., grayscale screen - Zimmermann & Sobolev, 2022; deleting social media apps – Hanley et al., 2019; lockout – Kim et al., 2019), timeout/scheduling usage (e.g., Hughes & Burke, 2018), or displacement by an alternative activity (e.g., Brailovskaia et al., 2022).

Unlike the previous two categories, interventions in the *Non-Limit* category did not explicitly ask participants to limit their smartphone or social media usage. Instead, they might (1) conduct sessions of skills training (coaching) in, for example, time management, relationship management, communication skills, or psychological counseling, with a focus on smartphone and social media usage (e.g., Chen et al., 2022; Pearson et al., 2020); (2) ask participants to engage regularly in a new activity (e.g., meditation – Choi et al., 2020; basketball – Xiao et al., 2021); (3) implement design friction on the smartphones (e.g., batching

[28] See Appendix D4 for an enumeration of the subtypes of outcome variables (effect sizes) across Intervention types.
[29] See Appendix D5 for descriptive statistics of study-level characteristics across Intervention types.
[30] Note that some of these sub-categories might not be represented by enough effect sizes (at least 2; Ryan & Hill, 2016; Valentine et al., 2010) in order to enable a meta-analysis.

notifications – Fitz et al., 2019; moving non-productivity apps further away from the Home screen – Ochs & Sauer, 2021); or (4) a mix of all of the former (e.g., Throuvala et al., 2020).

I do not hypothesize about the differences in effect sizes among the three proposed intervention types, due to two reasons. First, to the best of my knowledge, this is the first attempt to systematize causal evidence and categorize interventions to reduce problematic smartphone/social media use, and previous interventions have shown mixed results. Thus, there is no ground to claim that one intervention type is better than the others. Second, while it may seem that the Limit Plus and Non-Limit interventions might be more effective than the Limit interventions, purely because they demand more involvement from participants, I take the view that doing more is not necessarily better. In certain situations, simple interventions might work better than complex ones (Sunstein, 2017; Sitzia et al., 2012). Rather than having specific hypotheses, I let the results of the meta-analytical models inform me about the effectiveness of different intervention types.

2.2.2 Outcome variable categories: 'Consumption' and 'Consequence'

Generally, interventions included in my meta-analysis aimed to reduce smartphone screen time and social media usage time, to boost well-being (e.g., happiness, positive affect, self-esteem), performance (e.g., GPA, work performance, cognitive performance), sleep quality, and to alleviate ill-being (e.g., anxiety, depression, stress, addiction to smartphone/social media). Each effect size corresponds to one of these outcome measures.

I separately analyzed smartphone and social media screen time (the 'consumption' outcomes) as I am interested in how effective interventions are in reducing such usage. I distinguish between objective and subjective screen time measures, as past research has shown that there is a difference between the two types (Jones-Jang et al., 2020; Parry et al., 2021). However, past research is mixed about the direction of this difference (Kaye et al., 2020).

A problem that arises with the 'consequence' outcomes is that their distributions across the three intervention types are not even (see Appendix D4). For example, regarding the measure of depression, Limit and Non-Limit interventions have nine effect sizes each, while Limit Plus has 15 effect sizes. To solve this imbalance and the issue of small number of effect sizes per outcome variable, I grouped the outcome variables into "To Increase" outcomes and "To Reduce" outcomes. Because I am interested in how interventions shift the magnitude of the effect sizes, I reverse-coded all "To Reduce" outcomes to make them comparable with "To Increase" outcomes. Because positive and negative outcomes are usually negatively correlated, I do not expect a difference in effect sizes between these two types of outcomes.

2.2.3 The role of population characteristics

Because prior research has shown that problematic smartphone/social media usage affects male and female differently (Twenge & Martin, 2020), I expect that there is a gender difference. I operationalize this with the gender ratio of each sample used in a study, in percentage of females in the sample.

Older people tend to face fewer psychological and productivity problems related to digital addiction (Anshari et al., 2016; van Deursen et al., 2015), therefore I expect the effect sizes of interventions on 'consequence' outcomes to be smaller for older than younger people. However, because older people tend to rely less on digital technologies (Ahn & Jung, 2016), I predict that interventions will result in higher reduction of social media and phone screen time consumption for them. I use mean age of the studies' samples to test these predictions.

I suspect that the instruction language (English vs. non-English) for each intervention might affect effect sizes, for four reasons: (1) some self-reported scales were translated from English to local languages without validation, which may result in biased measurements (Cheng et al., 2021), (2) samples which are proficient enough to understand English instructions might also be higher in education and income levels[31], (3) it serves as a proxy for cultural differences, as there is evidence for social media addiction variation across cultures (Cheng et al., 2021; Olson et al., 2022; Hancock et al., 2022), and (4) it might be predictive of a publication bias in which studies using English instructions might be more likely to enter more reputable journals[32]. Thus, while I do not have a prediction on how instruction language affects effect sizes, it still acts as a good control variable.

During the coding process, I contacted authors of any article that did not contain sufficient information to derive the three aforementioned population characteristics.

2.2.4 The role of study characteristics

I include three study characteristics as control variables. First, I distinguish between studies using a single pre-post (SPP) design without a control condition, and those using a double-difference (DD) design (control vs. treatment groups measured before and after the treatment period). Although the latter is considered more rigorous (i.e., lower statistical bias in the estimation of the effect size), there is no basis to predict effect sizes based on the type of design (i.e., whether the effect size is bigger or smaller in DD designs than in SPP designs).

Second, the duration of the intervention varies from as little as one day (e.g., Przybylski et al., 2021) to as long as 3 months (e.g., Reed et al., 2023)[33]. Treatment effects tend to be stronger for longer treatment duration (Loewenstein et al., 2016).

Third, for field experiments, people tend to revert to their old behaviors over time (Brandon et al., 2017; Wood & Neal, 2016). I expect that longer follow-up post-treatment

---

[31] Most articles do not provide statistics on education and income levels of their samples, therefore, I cannot capture these variables directly.
[32] I included Impact Factor as an additional control variable in the full multivariate model. This does not change the coefficient of the instruction language by much, and it remains significant. See Appendix D6 for more details.
[33] Non-limit, skill-training interventions (e.g., Lan et al., 2018) can take place over months, but I only count the number of treatment sessions as actual treatment duration.

effects are weaker than more immediate post-treatment effects. This control variable is coded as the number of days that has elapsed after the end of the intervention period.

2.3 Effect size calculations

I follow the recommendations by Lipsey and Wilson (2001) to compute effect sizes[34]. All outcome variables in the pool of selected studies are continuous variables, therefore I compute their effect sizes as standardized mean difference (Cohen's *d*), except in the few cases when they were already reported by the studies. I compute this *d* value as the mean difference of an outcome variable between the treatment and control conditions, divided by the pooled standard deviation.

When a study tested more than one intervention, I calculated separate effect sizes per intervention and accounted for their nested nature in the statistical analysis. When a study had multiple post-treatment measurement time points, I computed separate effect sizes at two time points: the closest and the farthest from the end of the intervention period.

Next, I presented the meta-analytical results with 'consequence' outcomes first, as they consist of a much larger number of effect sizes (k = 247), and are generally the more important intervention outcomes, than 'consumption' outcomes (k = 60).

3. Analyses and Results: 'Consequence' Outcomes

Both existing meta-analyses (Liu et al., 2019; Malinauskas & Malinauskiene, 2019) used a standard two-level meta-analytical model (Borenstein et al., 2009). To account for the nested structure of my dataset (Fernández-Castilla et al., 2020), in which one study may measure multiple outcomes, oftentimes with more than one post-test measurement (hence multiple

[34] I used the online effect size calculator available at: https://www.campbellcollaboration.org/research-resources/effect-size-calculator.html (Wilson, 2023).

effect sizes), and one article may contain more than one study, I used a four-level, random-effects meta-analytic model implemented by the '*metafor*' R package (Viechtbauer, 2010).

3.1 Average meta-analytical effect: intercept-only models

These are the benchmark models in standard meta-analyses. They are used to examine the overall effect of all interventions and to check which nested structure would be best to fit the effect size dataset.

In addition to computing the pooled effect size through meta-analysis, it is important to evaluate how trustworthy this metric is, by quantifying the degree of heterogeneity between all the effect sizes included in the analysis. Effect size heterogeneity is measured with the I^2 statistic (Higgins & Thompson, 2002), which ranges from 0 to 100%. The threshold I^2 values of 25%, 50%, and 75% denote low, medium, and high degrees of heterogeneity, respectively. In addition, following Cheung (2014) and using the computation method in Cadario and Chandon (2020), I decomposed this overall heterogeneity further into between-article heterogeneity, within-article heterogeneity, and within-study heterogeneity.

The standard two-level model yields a statistically significant average effect size (d = 0.31, z = 9.42, $p < .001$) with a very large amount of heterogeneity (I^2 = 93.39%). The three-level model fits the data significantly better than the two-level model (χ^2 (1) = 151.50, $p <$ .001) and yields a slightly smaller estimate of the average effect size (d = 0.29, z = 5.50, $p <$ .001). Yet, the four-level model fits the data significantly better than the three-level model (χ^2 (1) = 4.85, p = 0.028) and yields a slightly larger average effect size (d = 0.31, z = 5.09, $p <$ .001). This effect size is considered small-to-medium as per Cohen's (1988) definition. The four-level, random-effects model shows that the total heterogeneity is higher within studies ($I^2_{(2)}$ = 18.5%) than within article ($I^2_{(3)}$ = 7.0%), and higher between articles ($I^2_{(4)}$ = 66.8%) than within article. Additional analyses reported in detail in Section 3.4 (trim and fill, sensitivity analyses, p-curve analyses) suggest that publication bias is very unlikely (Andrews & Kasy, 2019; McShane et al., 2016).

3.2 Influence of predictors: univariate vs. multivariate model

Next, I examine the effects of outcome, study and population characteristics by including them as predictors in univariate and multivariate meta-analytic models.

Univariate analyses are models that include only one predictor. The existing two meta-analyses (Malinauskas & Malinauskiene, 2019; Liu et al., 2019) used univariate meta-analyses. Because univariate models exclude covariates, they increase the possibility that significant differences are due to confounds. Multivariate models would alleviate this concern (Jackson et al., 2011) by accounting for multiple predictors. I performed both univariate and multivariate analyses and confirm that the latter yields a better model fit.

3.2.1 Univariate Models

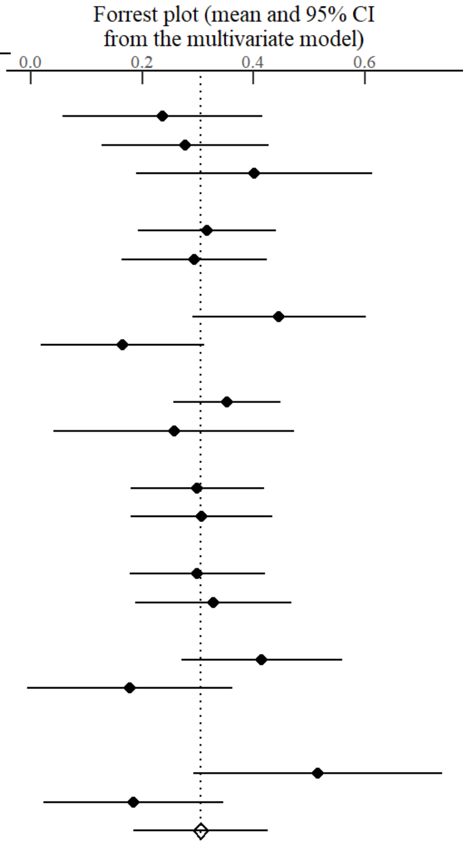
I estimated one univariate meta-regression for each predictor. These univariate analyses provide benchmark values, which can be compared with the estimates obtained in the full multivariate model. Because I have eight predictors, eight univariate models are estimated. When the predictor is categorical (for example, study design), the univariate model estimates the average effect size at each level of that categorical predictor. Continuous predictors, such as treatment duration, are first mean-centered, and have their parameter and intercept estimated in the respective univariate models. I illustrate the influence of each continuous predictor by computing point estimates of effect size at the 10^{th} percentile and 90^{th} percentile of the predictor[35]. The mean effect sizes and standard errors for all predictors are shown in Figure 8.

[35] With the exception of Treatment duration (with point estimates at the 25^{th} and 75^{th} percentiles), all continuous predictors have their point estimates at the 10^{th} and 90^{th} percentiles. This is done because of two reasons: first, for some predictors (e.g., age), the 25^{th} and 75^{th} percentiles correspond to values that do not qualitatively and meaningfully differ (e.g., 20 years old vs. 26 years old). Second, some predictors have a highly right-skewed

Figure 8

Effect Sizes ('Consequence') in the Univariate and Multivariate Models

Predictor	k	d (Univariate)	SE (Univariate)	d (Multivariate)	SE (Multivariate)
Intervention type					
Limit	60	0.17+	0.10	0.24**	0.09
Limit Plus	101	0.20**	0.07	0.28***	0.08
Non-Limit	86	0.60***	0.09	0.40***	0.11
DV category					
To-reduce	165	0.32***	0.06	0.32***	0.06
To-increase	82	0.28***	0.06	0.29***	0.07
Instruction Language					
Non-English	121	0.51***	0.08	0.45***	0.08
English	126	0.14*	0.07	0.16*	0.07
Design					
Double-difference	211	0.33***	0.06	0.35***	0.05
Single Pre-Post	36	0.21	0.13	0.26*	0.11
Treatment duration					
Short (7 days)	247	0.30***	0.07	0.30***	0.06
Long (21 days)	247	0.31***	0.06	0.31***	0.06
Post-treatment duration					
Immediate (0 day)	247	0.30***	0.06	0.30***	0.06
Long (84 days)	247	0.34***	0.07	0.33***	0.07
Age					
Younger (18 y.o.)	247	0.45***	0.08	0.41***	0.07
Older (30 y.o.)	247	0.11	0.09	0.18+	0.09
Gender (proportion of sampled female)					
Fewer female (26%)	247	0.57***	0.14	0.52***	0.11
More female (81%)	247	0.18*	0.08	0.18*	0.08
Average effect	247	0.31***	0.08	0.30***	0.06



Note: +p < .1; *p < .05; **p < .01; ***p < .001

Univariate analyses suggest that the effectiveness of interventions on improving consequences of problematic smartphone and social media use varies by intervention type (R^2 = 26%, $\chi^2(2) = 11.84$, $p = 0.003$). Figure 8 shows that the estimated effect sizes in the univariate analysis of intervention type vary between $d = 0.17$ for Limit interventions to $d = 0.60$ for Non-Limit interventions, which are all at least marginally different from zero. Univariate analyses found no difference in effect sizes between DV categories (to-increase or to-reduce; $R^2 = 1\%$, $\chi^2(1) = 0.73$, $p = 0.394$), but significant effect depending on the language with which the study's instructions were administered to subjects (Non-English vs. English: R^2 = 25%, $\chi^2(1) = 11.04$, $p = .001$). They found no difference between study design ($R^2 = 1\%$,

distribution, which makes values of the 25th and 75th percentiles equal (e.g., both correspond to 0 day for post-treatment duration).

$\chi^2(1) = 0.72$, $p = 0.397$), treatment duration ($R^2 = 0\%$, $\chi^2(1) = 0.003$, $p = 0.957$), or post-treatment duration ($R^2 = 1\%$, $\chi^2(1) = 1.02$, $p = 0.312$). However, they found that age variation significantly predicts differences in effect size ($R^2 = 12\%$, $\chi^2(1) = 6.53$, $p = 0.011$), and between studies with more vs. fewer female subjects ($R^2 = 9\%$, $\chi^2(1) = 4.30$, $p = 0.038$).

3.2.2 Multivariate Model

I estimated a full model with all the predictors entered simultaneously. The multivariate model explained 44% of the variance, a significant improvement over the intercept-only model ($\chi^2(9) = 27.36$, $p = 0.001$). It is also a significant improvement over the best univariate model with intervention type ($\chi^2(7) = 15.52$, $p = 0.030$), which only explains 26% of the variance. This suggests that my overall conceptual framework captured a substantial variation in the effect sizes – much more than any separate univariate model.

Figure 8 shows the multivariate effect sizes estimated for each level of a given predictor when all the other predictors are at their mean value. The average effect size computed across all 247 observations remains at $d = 0.31$ and is significantly different from zero ($z = 4.92$, $p < .001$).

Within each predictor, a level or point estimate that has relatively higher effect size than other levels or point estimates in the univariate models tends to have reduced effect size in the multivariate model, and vice versa. This makes comparisons between levels of a predictor in the multivariate model more conservative than those in the univariate models. For example, Non-Limit interventions have an effect size of $d = 0.40$ in the multivariate model, a reduction of 0.20 from the corresponding univariate model, while Limit Plus interventions have an effect size of $d = 0.28$ in the multivariate model, an increase of 0.08 from the respective univariate model. I detail these comparisons with planned contrasts in the next section.

3.3 Planned contrasts

I compare the reference level to other level(s) within each predictor using simple effect coding (e.g., double-difference design = 0.5, single pre-post design = -0.5 (reference level)) in the multivariate model with all predictors included (intervention type, DV category, instruction language, study design, treatment duration, post-treatment duration, age, and sample gender ratio). Results are shown in Table 7.

Table 7
Parameter Estimates of the Multivariate Model ('Consequence' Effect Sizes)

	β	*SE*	*z*
Intercept	0.30***	0.06	4.92
Intervention type			
Limit	(ref)		
Limit Plus	0.04	0.10	0.40
Non-Limit	0.16	0.14	1.22
DV category			
To-reduce	(ref)		
To-increase	-0.02	0.04	-0.60
Instruction Language			
Non-English	(ref)		
English	-0.28**	0.09	-3.05
Design			
Double-difference	(ref)		
Single Pre-Post	-0.10	0.12	-0.81
Treatment duration (day)	0.0005	0.003	0.19
Post-treatment duration (day)	0.0003	0.0005	0.72
Age	-0.02*	0.009	-2.09
Gender (% of female)	-0.006*	0.003	-2.21
k (no. of effect sizes)	247		
S (no. of studies)	64		
A (no. of articles)	51		
R^2	0.44		
LR test versus intercept-only model	27.36**		

Note: *p < .05; **p < .01; ***p < .001

Effect sizes do not vary significantly between the three intervention types. Although Limit interventions have smaller effect size than both Limit Plus and Non-Limit ones, these differences are non-significant ($\beta_{vs.\ Limit\ Plus}$ = 0.04, z = 0.40, p = 0.690; $\beta_{vs.\ Non\text{-}Limit}$ = 0.16, z = 1.22, p = 0.224). This is in contrast to what the corresponding univariate model suggested. Effect sizes do not vary between the two categories of outcome variables either (To-reduce vs. To-increase, β = -0.02, z = -0.60, p = 0.549).

Interestingly, and consistent with the univariate analyses, interventions whose instructions were delivered in English have significantly smaller effect sizes than those with non-English instructions (β = -0.28, z = -3.05, p = 0.002). Interventions assessed with a single Pre-Post design do not vary in effect sizes compared to those evaluated with a double-difference design (β = -0.10, z = -0.81, p = 0.415). Contrary to my predictions, longer treatment duration does not associate with more effective intervention (β = 0.0005, z = 0.19, p = 0.846), and longer post-treatment follow-up duration does not associate with smaller effect sizes (β = 0.0003, z = 0.72, p = 0.472).

As expected, effect sizes differ across age. Specifically, older people might experience smaller effect sizes from interventions than younger people (β = -0.02, z = -2.09, p = 0.036). Also as expected, there is difference in intervention effectiveness across the two genders: effect size decreases when the proportion of females increases (β = -0.006, z = -2.21, p = 0.027). Notably, these results are consistent with those of the corresponding univariate analyses.

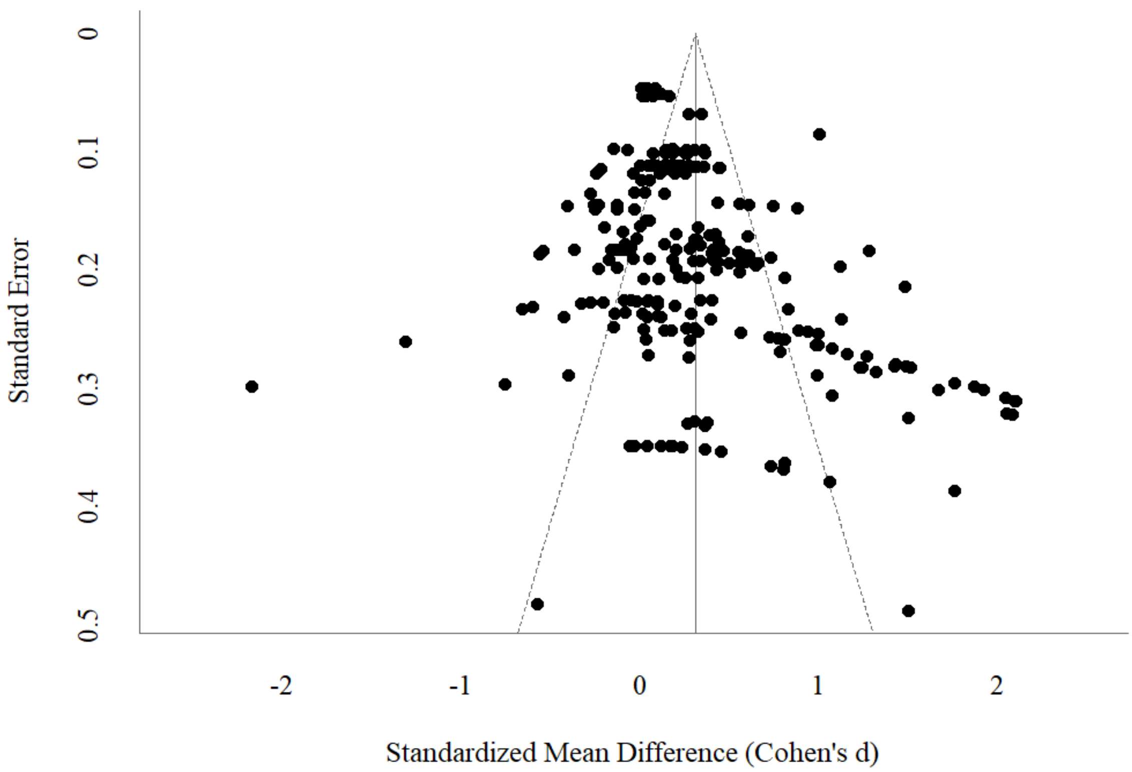
3.4 Publication bias

In Figure 9, the funnel plot displays observations as a function of standardized mean difference (i.e. the effect size) and the standard error (i.e. the sampling variance). Observations outside the "funnel" (demarcated by the two dotted lines) may present potential publication bias, if there are many such observations on one side, and not enough observations on the other side to counterbalance them. The current funnel plot does not show this pattern.

First, using the trim and fill method (Duval & Tweedie, 2000), I found no missing studies

Figure 9

Funnel plot illustrating the trim and fill method. Missing studies on one side are not detected.



on the left side of the funnel plot, which suggests little evidence for publication bias.

Second, I performed sensitivity analyses by removing 29 (23) observations with leverage (Cook's distance) higher than twice the average leverage (Cook's distance). The adjusted effect size from leaving out high-leverage observations is only slightly higher than the original 4-

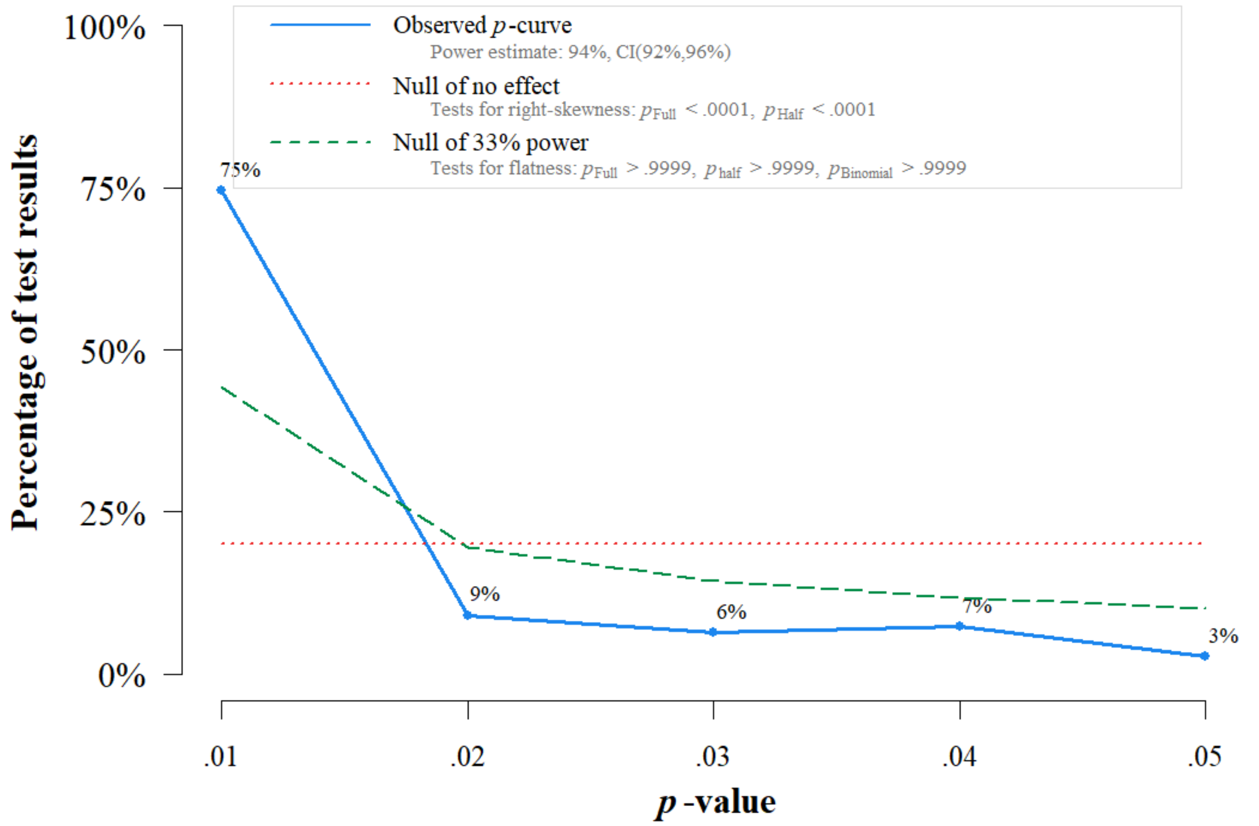
level, intercept-only model (*d* = 0.31, *SE* = 0.08, $p < .001$), and this adjusted effect size from omitting high-Cook's distance observations does not differ from the original (*d* = 0.30, *SE* = 0.06, $p < .001$).

Finally, I plotted the p-curve (Simonsohn et al., 2014b) for all effect sizes included in my sample. As shown in Figure 10, the p-curve is strongly rightly-skewed, indicating that it is likely that a true effect exists in my data. To quantitatively qualify this observation, I conducted two tests: the right-skewness tests and the flatness tests (Simonsohn et al., 2015). My sample contains 247 effect sizes, of which 110 (45%) are significant ($p < .05$), and 137 (55%) are non-significant. Three tests (the binomial test, the full p-curve pp-value test, and the half p-curve test) were conducted to examine the right-skewness and flatness of the p-curve (Harrer et al., 2021). The tests for right-skewness were all significant (*p*s < .001), and the tests for flatness were all non-significant (*p*s > .05)[36].

[36] There is a concern with high heterogeneity of the meta-analytic model (84.2%) that may undermine the validity of the p-curve analyses (van Aert et al., 2016). A remedy is to exclude outliers. However, after excluding outlier observations based on Cook's distance, the model still has a high heterogeneity (80.9%). This subset of sample still yielded results consistent with the full-sample p-curve analyses.

Figure 10

P-curve



Note: The observed *p*-curve includes 110 statistically significant ($p < .05$) results, of which 95 are $p < .025$. There were 137 additional results entered but excluded from *p*-curve because they were $p > .05$.

4. Analyses and Results: 'Consumption' Outcomes

I apply similar analyses to the subset of effect sizes corresponding with 'consumption' outcomes, which consist of social media and overall smartphone usage, with several key differences. First, three-level random-effects models are used instead of the four-level models used by the meta-analytic models for the 'consequence' outcomes. Second, because I am interested in the reduction of social media and smartphone usage, I do not reverse-code the effect sizes. Third, I distinguish between objective and subjective measurement methods, and usage types (social media vs. phone screen time), which are both entered as predictors for the multivariate and univariate models.

4.1 Average meta-analytical effect: intercept-only models

The standard two-level model yields a statistically significant average effect size (d = -0.46, z = -7.36, p < .001) with very high heterogeneity (I^2 = 89.5%). The three-level model does not fit the data significantly better than the two-level model (χ^2 (1) = 1.12, p = 0.289) and yields a slightly higher estimate of the average effect size (d = -0.49, z = -6.83, p < .001). The four-level model fits the data almost identically as the three-level model and yields the same average effect size. This effect size is considered medium (Cohen, 1988).

While statistical tests for model fitness are inconclusive, I choose the three-level model as the benchmark meta-analytical model in this case, because it is the most suitable to account for the nested structure of the effect size data set (k = 60, in 30 articles) out of the three models[37] [38]. The three-level, random-effects model shows that the total heterogeneity is higher between articles ($I^2_{(3)}$ = 67.7%) than within articles ($I^2_{(2)}$ = 22.1%).

4.2 Influence of predictors: univariate vs. multivariate model

4.2.1 Univariate Models

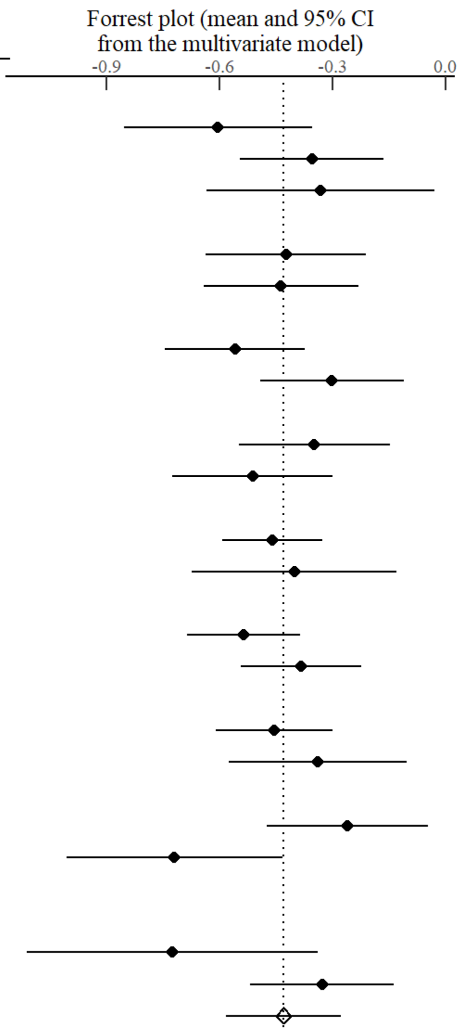
I estimated one univariate meta-regression for each of the nine predictors (similar to section 3.2.1). The mean effect sizes and standard errors for all predictors are shown in Figure 11.

[37] The within-article heterogeneity of the four-level model is 0%, indicating that there is no effect size being nested within a study of a multi-study article. This eliminates the need for a four-level model. In addition, the three-level model yields a more conservative standard error (i.e. bigger) than the two-level model does.
[38] I also conducted similar checks for publication bias (trim and fill, sensitivity analyses, p-curve) as I did for the meta-analysis of 'consequence' outcomes, and found no evidence of publication bias. See Section 4.4 for details.

Figure 11

Effect Sizes ('Consumption') in the Univariate and Multivariate Models

Predictor	k	d (Univariate)	SE (Univariate)	d (Multivariate)	SE (Multivariate)
Intervention type					
Limit	10	-0.74***	0.14	-0.60***	0.13
Limit Plus	36	-0.46***	0.08	-0.36***	0.10
Non-Limit	14	-0.26$^+$	0.14	-0.33*	0.15
Measurement method					
Subjective	26	-0.52***	0.11	-0.42***	0.11
Objective	34	-0.46***	0.09	-0.44***	0.11
Usage type					
Social media	30	-0.62***	0.09	-0.56***	0.09
Phone screen time	30	-0.33***	0.09	-0.30**	0.10
Instruction Language					
Non-English	22	-0.46***	0.12	-0.35***	0.10
English	38	-0.50***	0.09	-0.51***	0.11
Design					
Double-difference	49	-0.47***	0.08	-0.46***	0.07
Single Pre-Post	11	-0.56***	0.16	-0.40**	0.14
Treatment duration					
Short (7 days)	60	-0.53***	0.08	-0.54***	0.08
Long (21 days)	60	-0.46***	0.07	-0.38***	0.08
Post-treatment duration					
Immediate (0 day)	60	-0.50***	0.07	-0.46***	0.08
Long (84 days)	60	-0.42***	0.12	-0.34**	0.12
Age					
Younger (18 y.o.)	60	-0.29**	0.10	-0.26*	0.11
Older (30 y.o.)	60	-0.75***	0.13	-0.72***	0.15
Gender (proportion of sampled female)					
Fewer female (26%)	60	-0.62**	0.23	-0.73***	0.20
More female (81%)	60	-0.44***	0.10	-0.33***	0.10
Average effect	60	-0.49***	0.11	-0.43***	0.08



Note: $^+$p < .1; *p < .05; **p < .01; ***p < .001

Univariate analyses suggest that the effectiveness of interventions for reducing consumption of smartphone screen time and social media use slightly varies across intervention types (R^2 = 12%, $\chi^2(2)$ = 5.21, p = 0.074), between d = -0.26 for Non-Limit interventions to d = -0.74 for Limit interventions. Notably, the effect sizes of the three intervention types are at least marginally different from zero. Univariate analyses found no difference in effect sizes between consumption measurement methods (objective or subjective; R^2 = 1%, $\chi^2(1)$ = 0.14, p = 0.706), but found that interventions are significantly more effective at reducing social media usage than overall phone screen time (R^2 = 13%, $\chi^2(1)$ = 5.05, p = 0.025). They found that effect sizes do not vary between instruction languages (Non-English vs. English: R^2 = 0%, $\chi^2(1)$ = 0.08, p = .773), between designs (R^2 = 1%, $\chi^2(1)$ = 0.24, p = 0.625), treatment duration (R^2 = 6%, $\chi^2(1)$ = 2.19, p = 0.139), post-treatment duration (R^2 = 1%, $\chi^2(1)$ = 0.43, p = 0.510), or gender (R^2 = 1%, $\chi^2(1)$ = 0.42, p = 0.519). However, they found that age significantly predicts effect size (R^2 = 13%, $\chi^2(1)$ = 5.45, p = 0.020).

4.2.2 Multivariate Model

The multivariate model, with all nine predictors included, explains 42% of the variance, much more than the intercept-only model ($\chi^2(10)$ = 24.75, p = 0.006). It also explains significantly more variance than the best univariate model with age as the predictor ($\chi^2(9)$ = 19.29, p = 0.023). Along with the same result obtained with the 'consequence' outcomes, this result further bolsters the validity of my grounded conceptual framework, which captures a substantial variation in the effect sizes, and more than other univariate models do.

Figure 11 shows the multivariate effect sizes estimated for each level of a given predictor when all the other predictors are at their mean value. The average effect size computed across 60 observations is d = -0.43 and is significantly different from zero (z = -5.57, p < .001). This effect size estimate is conservative, and is approximately 12% smaller than the estimate yielded by the intercept-only model and the average estimate of all univariate models (both ds = -0.49).

4.3 Planned contrasts

For categorical predictors, I compare the reference level to other level(s) within each predictor using simple effect coding (e.g., double-difference design = 0.5, single pre-post design = -0.5) in the multivariate model. Results are shown in Table 8.

Table 8

Parameter Estimates of the Multivariate Model ('Consumption' Effect Sizes)

	β	*SE*	*z*
Intercept	-0.43***	0.08	-5.57
Intervention type			
Limit	(ref)		
Limit Plus	0.25^{+}	0.15	1.65
Non-Limit	0.27	0.19	1.40
Measurement method			
Subjective	(ref)		
Objective	-0.01	0.15	-0.09
Usage type			
Social media	(ref)		
Phone screen time	0.26*	0.11	2.25
Instruction Language			
Non-English	(ref)		
English	-0.16	0.14	-1.15
Design			
Double-difference	(ref)		
Single Pre-Post	0.06	0.15	0.37
Treatment duration (day)	0.01**	0.003	3.26
Post-treatment duration (day)	0.001	0.001	1.03
Age	-0.04*	0.02	-2.28
Gender (% of female)	0.007^{+}	0.005	1.70
k (no. of effect sizes)	60		
A (no. of articles)	30		
R^2	0.42		
LR test versus intercept-only model	24.75**		

Note: *$p < .05$; **$p < .01$; ***$p < .001$

Limit interventions reduce consumption marginally more than Limit Plus interventions ($\beta_{vs.\ Limit\ Plus} = 0.25$, $z = 1.65$, $p = 0.100$), but not more than Non-Limit ones ($\beta_{vs.\ Non\text{-}Limit} = 0.27$, $z = 1.40$, $p = 0.162$). This is consistent with what the corresponding univariate model indicated. Like the univariate models, effect sizes do not vary between the two measurement methods (subjective vs. objective, $\beta = -0.01$, $z = -0.09$, $p = 0.931$), but are lower for phone screen time than social media usage ($\beta = 0.26$, $z = 2.25$, $p = 0.024$).

Consistent with the univariate analyses, interventions whose instructions were delivered in English do not differ in effect sizes from those with non-English instructions (β = -0.16, z = -1.15, p = 0.249). Interventions assessed with a single pre-post design do not vary in effect sizes compared to those evaluated with a double-difference design (β = 0.06, z = 0.37, p = 0.709). Surprisingly, contrary to my predictions, longer treatment duration does not associate with reduced consumption (β = 0.01, z = 3.26, p = 0.001), and longer post-treatment follow-up duration does not associate with increased consumption (β = 0.001, z = 1.03, p = 0.303).

As expected, and in line with univariate analysis, effect sizes differ across age. Specifically, older people tend to reduce consumption with interventions than younger people (β = -0.04, z = -2.28, p = 0.022). Finally, in contrast to univariate analysis, the difference in intervention effectiveness across the two genders is marginally significant: consumption is reduced less when the proportion of females in the sample increases (β = 0.007, z = 1.70, p = 0.097).

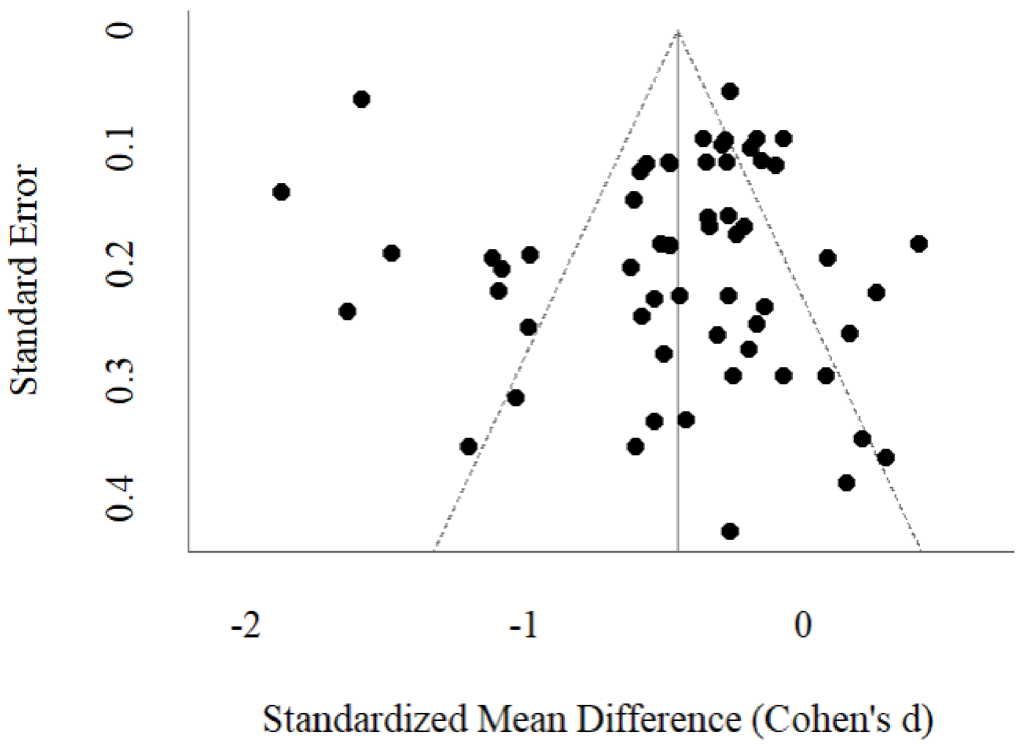
4.4 Publication bias

In Figure 12, the funnel plot displays observations as a function of standardized mean difference (i.e., the effect size) and the standard error (i.e., the sampling variance). Observations outside the “funnel” (demarcated by the two dotted lines) may present potential publication bias, if there are many such observations on one side, and not enough observations on the other side to counterbalance them. The current funnel plot does not show this pattern.

First, using the trim and fill method (Duval & Tweedie, 2000), I found no missing studies on the left side of the funnel plot, which suggests little evidence for publication bias.

Figure 12

Funnel plot illustrating the trim and fill method. Missing studies on one side are not detected.

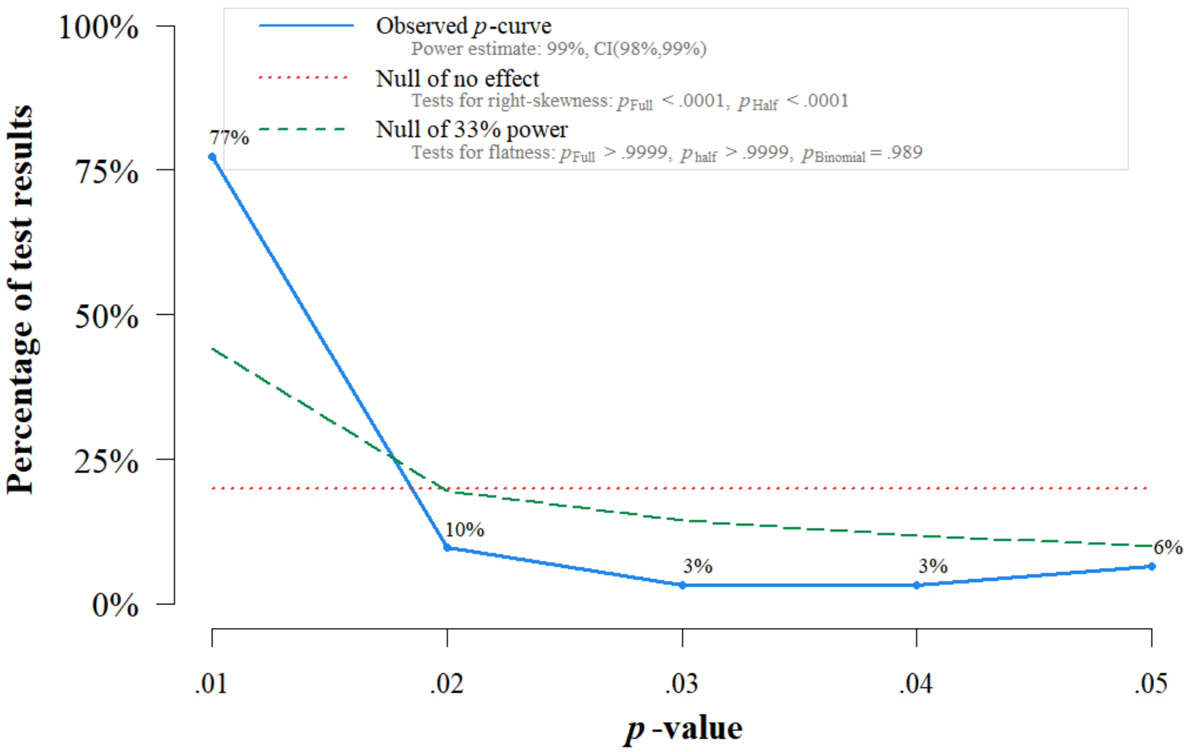


Second, I performed sensitivity analyses by removing 6 observations with Cook's distance higher than twice the average Cook's distance. The adjusted effect size from leaving out high-Cook's distance observations does not differ from the original (d = -0.550, SE = 0.09, p < .001). There was no observation with abnormally high leverage.

Finally, I plotted the p-curve (Simonsohn et al., 2014b) for all effect sizes included in my sample. As shown in Figure 13, the p-curve is strongly rightly-skewed, indicating that it is likely that a true effect exists in the data. To quantitatively qualify this observation, I conducted two tests: the right-skewness tests and the flatness tests (Simonsohn et al., 2015). The sample contains 60 effect sizes, of which 31 (52%) are significant (p < .05), and 29 (48%) are non-significant. Three tests (the binomial test, the full p-curve pp-value test, and the half p-curve test) were conducted to examine the right-skewness and flatness of the p-curve (Harrer et al., 2021). The tests for right-skewness were all significant (*p*s < .001), and the tests for flatness

Figure 13

P-curve



Note: The observed *p*-curve includes 31 statistically significant ($p < .05$) results, of which 27 are $p < .025$. There were 29 additional results entered but excluded from *p*-curve because they were $p > .05$.

were all non-significant (*p*s > .05)[39]. This indicates that the current body of evidence, from which I infer the effects of interventions to reduce social media and smartphone usage, is likely to be based on true effects.

[39] There is a concern with high heterogeneity of the meta-analytic model (91.9%) that may undermine the validity of the p-curve analyses (van Aert et al., 2016). A remedy is to exclude outliers. However, after excluding outlier observations based on Cook's distance, the model still has a high heterogeneity (91.6%). This subset of sample still yielded results consistent with the full-sample p-curve analyses.

Discussion of Chapter 3

How effective are interventions?

The four-level multivariate analysis of 247 ‘consequence’ effect sizes yields an average effect size of $d = 0.30$, 95% CI = [0.18, 0.42]. Meanwhile, the three-level multivariate analysis of 60 ‘consumption’ effect sizes shows that the average effect size of interventions is $d = -0.43$, 95% CI = [-0.59, -0.27]. Both of these estimates are considered small-to-medium (Cohen, 1988).

Are there differences between intervention types?

Interventions generally aim to improve the ‘consequence’ outcomes, and reduce the ‘consumption’ outcomes. However, each type of interventions may focus on a different goal. For example, Limit and Limit Plus interventions tend to focus on the latter, while Non-Limit interventions tend to focus on the former. The results of the two multivariate models demonstrate this pattern: Limit interventions have the smallest effect size for ‘consequence’ outcomes ($d = 0.24$), but have the biggest effect size for reducing consumption ($d = -0.60$), out of the three types. Conversely, Non-Limit interventions improve ‘consequence’ outcomes the most ($d = 0.40$), yet affect consumption the least ($d = -0.33$). I note that while this pattern is supported by the set of univariate analyses with intervention type as the sole predictor, it is not supported by the more conservative pair of multivariate models, which found no difference between the three intervention types.

However, if this pattern were true, it would mean that policymakers should not design interventions to directly reduce social media and smartphone use. Instead, interventions that focus on other aspects such as time management or skill training might be more efficient, because the end goal of most interventions is to improve mental health, productivity, and quality of life. This would also be consistent with the finding by Orben and Przybylski (2019) that there is only a weak negative association between digital technology use and well-being in adolescents.

How do other factors affect the effectiveness of interventions?

The consequence and consumption multivariate models account for 44% and 42% of the variance among the effect sizes, respectively, which indicates that some of the characteristics of the outcome, study, and population uniquely influence the effectiveness of interventions. Table 9 shows a summary of the findings regarding each of these predictors, and indicates whether corresponding hypotheses, when applicable, are supported or refuted. Several hypotheses which do not predict direction of the effects are also elaborated with this information from the data.

Table 9

Summary of multivariate model's findings regarding effect sizes

Predictor	*A priori* hypothesis	Outcome type (Consumption / Consequence)	Findings
Intervention type (Limit / Limit Plus / Non-Limit)	No	Both	No difference
DV category (to-increase / to-reduce)	No	Consequence	No difference
Measurement method (objective / subjective)	Yes; different	Consumption	Hypothesis not supported
Usage type (social media / screen time)	No	Consumption	Social media > screen time
Instruction language (English / non-English)	No	Both	English < non-English (consequence); No difference (consumption)
Design (double difference / single pre-post)	No	Both	No difference

Treatment duration	Yes; longer > shorter	Both	Hypotheses not supported; longer < shorter (consumption)
Post-treatment duration	Yes; longer < shorter	Both	Hypotheses not supported
Age	Yes; older < younger (consequence); older > younger (consumption)	Both	Hypotheses supported
Gender	Yes; different	Both	Hypotheses supported; female < male

First, I found that interventions are equally effective at increasing positive consequences ($d = 0.29$) and decreasing negative consequences ($d = 0.32$). To put these numbers into perspective, I compare them with those obtained by the largest meta-analysis using correlational evidence, that also aggregated consequence outcomes into well-being and ill-being (Huang, 2020). They found that the associations between problematic social media use with (1) well-being and (2) ill-being, respectively, are $r = 0.16$ (which approximates $d = 0.32$[40]), and $r = 0.27$ (which approximates $d = 0.56$). Given this stronger association with ill-being than well-being, I should expect a bigger effect size of interventions on negative consequences in the meta-analysis; however, this is not the case. This implies that interventions might not be as efficient at reducing ill-being as they are at recovering well-being outcomes.

Previous correlational studies (e.g., Parry et al., 2021; Kaye et al., 2020) found a mismatch between objective and subjective measurements of social media and phone screen time. I did not find this difference in terms of interventions' effect size. This does not mean that there is no difference between the two methods. Rather, this might suggest that the underestimation/overestimation of subjective reporting of screen time or social media use is uniform across the pre- and post- treatment periods. I recommend that researchers should not

[40] I used Ruscio's (2008) formula to make the conversion from Pearson's *r* to standardized mean difference *d*.

be quick to dismiss subjective reporting of phone/social media usage, because in cases where objective methods are not possible (perhaps due to implementation difficulties), subjective reporting may still provide good-enough measurements to track intervention's effectiveness, similar to the sentiment expressed by other scholars (Jones-Jang et al., 2020).

To the best of my knowledge, no previous study has investigated whether interventions are more effective at reducing social media use than phone screen time. Scholars have made the distinction between social media addiction and smartphone addiction. Some have argued that social media addiction should be targeted more than smartphone use reduction (e.g., Panova & Carbonell, 2018), while others have argued that smartphone, being the "gateway" to more addictive behaviors, should be curbed instead (Schnauber-Stockmann & Naab, 2019; Chen et al., 2020). This distinction is important, because it is directly related to how future interventions should focus on, so as to maximize their benefits, while keeping costs low. I found that currently existing interventions are 1.9 times more effective at reducing social media usage (d = -0.56) than general smartphone usage (d = -0.30).

Interestingly, I found that consequence outcomes' effect sizes are affected by instruction languages, such that interventions conducted in English reported significantly smaller effects (2.9 times) than those conducted in non-English languages. I ruled out the possibility that this was due to interventions using English-speaking samples were more likely to be published in more reputable journals, using journal impact factor as an additional control in the multivariate model (Appendix D6). However, I cannot pinpoint whether this remarkable difference is due to (a) cultural differences, (b) income / education level, or (c) measurement errors (i.e., some self-reported scales originally developed in English were not properly validated when they were translated to a non-English language, an issue raised by Cheng et al., 2021). I would like to highlight the relevance of this finding in light of an emerging research thread that examines the association between cultural factors and social media addiction (e.g., Cheng et al., 2021; Olson et al., 2022; Hancock et al., 2022). On the other hand, I also note that the literature is silent on the relationship between problematic social media / smartphone use and income / education level. I did not find the same difference for the consumption outcomes.

I did not find any difference in effect sizes between double-difference and single pre-post designs. While the more rigorous double-difference design should be used as frequently as possible, I believe that single pre-post design is still valuable in pilot tests of novel intervention ideas, without the disproportionate risk of getting biased results.

Surprisingly, the analysis shows that longer treatment duration does not significantly improve interventions' effectiveness. This implies that treatment duration alone may not guarantee a more effective intervention (Angus et al., 2014); rather, other variables, such as treatment intensity, quality, or adherence to treatment, might play a more significant role. For example, it has been shown that the longer an intervention goes on, the lower the rate of participation (lower adherence; van der Bij et al., 2002). Treatment duration, however, is important for interventions based on habit formation principles, and researchers have found that the sufficient time needed to form a habit (e.g. using social media less) varies across habit types and individuals (Gardner et al., 2012; Buyalskaya et al., 2023).

Furthermore, I did not find that the effectiveness of interventions decay over longer post-treatment periods. This might be because of a lack of sufficient studies with a non-zero follow-up post-treatment duration. Specifically, only 20 out of 72 studies (28%) in my data set followed up with their participants after the treatment period has ended. Since policymakers are concerned about the long-term effects of interventions, it is critical that future interventions include a post-treatment period to evaluate how persistent intervention effects are.

The multivariate meta-analysis for consequence outcomes found that as people age, intervention effect sizes get smaller. This is consistent with the finding of Busch and McCarthy

(2021)'s systematic review, that younger users' wellbeing is more affected by problematic smartphone and social media usage than older people's. Conversely, older people tend to reduce their consumption more than younger people during interventions. A possible explanation for this result is that younger generations are more likely to have early exposure to smartphones and social media, and they rely more on these technologies than older generations (Ahn & Jung, 2016).

Finally, the finding that genders respond differently to interventions clarifies the uncertainty in the literature that was raised by Busch and McCarthy (2021), that "… many studies were not able to identify any correlation between gender and PSU [problematic smartphone usage]". Specifically, the current meta-analyses found that samples with more females experienced about 2.9 times less improvement in consequence outcomes, and reduced about 2.2 times less than samples with more males in terms of smartphone / social media consumption. In addition, I respond to Chen et al. (2017)'s observation that gender differences in smartphone addiction are usually related to culture, by pointing to the multivariate meta-analytical model which shows that gender and culture (as expressed through the instruction language variable) factors introduce unique heterogeneous influence on interventions' effect sizes.

Limitations and future research

My grounded framework and meta-analysis of causal evidence for a link between problematic smartphone / social media usage and well-being elucidate where more research is needed. First, using Appendix D4, future research may find it helpful to investigate the less populated cells, which might represent gaps in knowledge about intervention effects on certain consequence / consumption outcomes. A clear example of this is the lack in interventions that focus on performance / productivity, which is indicated by a dearth of effect sizes for the ‘distraction’ and ‘GPA’ variables. Furthermore, one can also spot the imbalance in research among the three intervention types. For instance, some research has investigated the effect of Limit and Non-Limit interventions on self-esteem, but none has been done for Limit Plus interventions.

An apparent limitation of the current framework is the categorization of consequence outcomes into To-Increase and To-Reduce groups. This is a common approach that has been used by past meta-analyses of correlational evidence in this domain, and it has drawn criticisms (Valkenburg, 2022). Understandably, the lack of research regarding some outcomes (Appendix D4) equates the lack of data, and this necessitates aggregating outcomes into larger categories that make sense so that research syntheses such as the current meta-analysis can inform policymakers, researchers, and other stakeholders.

An intriguing finding in my analysis is the stark difference in effect sizes between interventions that used English as the instruction language and those that used non-English instructions. Because language generally influences many domains, I do not have a theoretical explanation for this finding. While I managed to rule out publication bias as a confound, future research should look into the various factors which are closely related to language, such as culture, income, or education level, as these factors are also likely to have implications for interventions.

The finding that current interventions are generally less effective for females than males might warrant gender-specific interventions. This proposal is further supported by past research which found that women tend to be more categorized as addicted in smartphones and social media than men (Vaghefi et al., 2017), and that the genders differ in how they are affected by these addictions (e.g., Anshari et al., 2016; Qudah et al., 2019; Wei et al., 2023).

Finally, some of my findings may help future research improve the efficiency of interventions. First, social media consumption tends to have larger relative reduction from interventions than overall phone screen time. If problematic social media use is more harmful than problematic smartphone use, perhaps it is more worthwhile to focus on devising interventions that specifically cater to reducing social media consumption instead of smartphone usage. Second, because improvement gains from interventions are significantly higher for younger people, and older people tend to rely less on smartphones, future interventions should focus on younger age groups. Given that many governments are paying attention to adolescents' use of social media, my synthesis of causal evidence should further corroborate this position.

Conclusion *(English version)*

This dissertation has examined the broad theme of interventions to reduce problematic digital consumption, focusing specifically on smartphone and social media usage, and how to improve the effectiveness of such interventions. Chapter 1 and 2 proposed and tested a novel dual-target intervention that incentivized consumers to simultaneously reduce their digital consumption, while increase a beneficial alternative activity. Chapter 3 provided a bird's-eye view of a broader range of interventions that aim to reduce problematic consumption; this third chapter organized, evaluated, and synthesized the causal evidence with a grounded framework.

Previous research (Wood & Neal, 2016) found that the effect of behavior change interventions using incentivization often diminished in the post-treatment period when incentives were stopped. Chapter 1 and 2 aim to address this problem by investigating a dual-target intervention to improve the persistence of an intervention effect once a monetary reward is no longer provided. This intervention capitalizes on the displacement of an undesired consumption pattern by incentivizing it along with inculcating beneficial alternative activities. Thus, it bridges the gap between two distinct types of past intervention studies: those that employed monetary incentives to reduce smartphone usage (e.g., Stanley et al., 2022; Collis & Eggers, 2022; Allcott et al., 2022), and those that simply introduced an alternative activity to a person's schedule (e.g., Reed et al., 2023; Brailovskaia et al., 2022; Esmaeili & Ahmadi, 2018).

Two field studies, designed to test the proposed dual-target intervention, in the context of reducing social media / smartphone consumption, showed that it could prolong the intervention effect for people who diligently performed the set targets during the treatment period, compared to those who only had a single reduction target. This finding offers a tool for policymakers interested in boosting the long-term effect of public interventions, not only in reducing digital addiction, but potentially in other domains of behavior change where habit formation plays a prominent role. In addition, social marketers might also find in this tool an opportunity to introduce new products that help consumers deal with addictive behaviors.

A clear direction for future research is to elaborate on the psychological mechanism(s) of dual-target interventions. Chapter 1 proposed that subjects with higher dual-target achievement self-signalled that they were less addicted to social media than they once believed, thereby maintaining a lower post-treatment social media consumption. The test for this indirect effect was not conclusive, but indicative of a strong direct effect from dual-target achievement, which might hint at other possible psychological mechanisms at play. Given that the sample size of Chapter 1's longitudinal study was small (N = 43), the mediation test should be confirmed by a future study with a bigger sample size.

In addition, another interesting future research direction deals with examining how cue distance between the pair of activities might affect a dual-target intervention's effectiveness. Walking (Chapter 2) creates a physical distance between consumers and their smartphones, effectively shielding them from stimuli such as push notifications that trigger the desire to use their phones. In contrast, using Duolingo (Chapter 1) on the same mobile device where social media is frequently accessed does not provide the same distancing effect. Although the contexts of both studies differ along the dimension of inter-activity cue distance, they did not directly examine how the effectiveness of dual-target intervention varies as a function of the cue strength based on its proximity to the consumer.

Past literature is relatively silent on the overall effectiveness of interventions for reducing smartphone and social media consumption, and the causal effect of this consumption on various well-being measures is still not clear. Thus, Chapter 3 seeks to compare the effectiveness of existing interventions to reduce problematic smartphone / social media usage, and examines whether this reduction is also linked to improved well-being. By organizing causal evidence found from intervention studies around a grounded framework, the chapter has managed to synthesize interesting insights using multivariate, multilevel meta-analytical models. Among these, interventions were found to have small-to-medium effect sizes on consumption reduction and consequence improvement. Importantly, heterogeneities in effect sizes are shown to be significant, and some of these (e.g., age and gender) might have particularly meaningful and potentially impactful influence on future intervention designs.

Apart from the findings on intervention effectiveness in Chapter 3, it is crucial to note that the meta-analyses used only *causal evidence*. Thus, it is one of the first research / research syntheses that provides a definitive answer to the debate on whether there is a causal link between problematic smartphone / social media usage and important outcomes such as mental health and productivity. The fact that interventions *do* improve various well-being measures lends strong causal support for this relationship. Apart from settling the academic debate, this chapter has carved a path forward to design more effective *and* efficient interventions.

Chapter 3 also found that interventions of which the main target is reduction of social media / smartphone consumption (i.e. Limit and Limit Plus interventions) tend to be better at reducing their consumption, but worse than interventions (i.e. Non-Limit) which primarily target consequence improvements. This might explain why dual-target intervention, which belongs to the Limit Plus family, reduced social media (Chapter 1) and smartphone (Chapter 2) consumption, but did not consistently improve well-being outcomes. This could be an interesting direction for future research.

Overall, Chapter 3's grounded framework and meta-analysis of causal evidence for a link between problematic smartphone / social media usage and well-being reveal three prominent areas where more research is needed. First, the literature is lacking interventions that focus on performance / productivity improvement. Second, there is an imbalance in research among the three intervention types identified by the grounded approach with regard to specific well-being measures. Third, there is a dearth of intervention studies with a non-zero follow-up post-treatment duration. Given policymakers' concern about the long-term effects of interventions, it is crucial for future interventions to incorporate longer post-treatment periods to assess the persistence of their effects.

Conclusión *(versión en español)*

Esta tesis ha examinado el amplio tema de las intervenciones para reducir el consumo digital problemático, centrándose específicamente en el uso de teléfonos inteligentes y redes sociales, y cómo mejorar la eficacia de dichas intervenciones. Los capítulos 1 y 2 propusieron y probaron una novedosa intervención de doble objetivo que incentivaba a los consumidores a reducir simultáneamente su consumo digital y a aumentar una actividad alternativa beneficiosa. El capítulo 3 proporcionó una visión general de una gama más amplia de intervenciones destinadas a reducir el consumo problemático; este tercer capítulo organizó, evaluó y sintetizó las pruebas causales con un marco fundamentado.

Investigaciones anteriores (Wood y Neal, 2016) descubrieron que el efecto de las intervenciones de cambio de comportamiento que utilizan incentivación a menudo disminuía en el período posterior al tratamiento cuando se interrumpían los incentivos. Los capítulos 1 y 2 pretenden abordar este problema investigando una intervención de doble objetivo para mejorar la persistencia del efecto de una intervención una vez que ya no se proporciona una recompensa monetaria. Esta intervención aprovecha el desplazamiento de un patrón de consumo no deseado incentivándolo junto con la inculcación de actividades alternativas beneficiosas. Por lo tanto, tiende un puente entre dos tipos distintos de estudios de intervención anteriores: los que emplearon incentivos monetarios para reducir el uso de teléfonos inteligentes (por ejemplo, Stanley et al., 2022; Collis & Eggers, 2022; Allcott et al., 2022), y los que simplemente introdujeron una actividad alternativa en la agenda de una persona (por ejemplo, Reed et al., 2023; Brailovskaia et al., 2022; Esmaeili & Ahmadi, 2018).

Dos estudios de campo, diseñados para probar la intervención de doble objetivo propuesta, en el contexto de la reducción del consumo de medios sociales / teléfonos inteligentes, mostraron que podría prolongar el efecto de la intervención para las personas que realizaron diligentemente los objetivos establecidos durante el período de tratamiento, en comparación con aquellos que solo tenían un único objetivo de reducción. Este hallazgo ofrece una herramienta para los responsables políticos interesados en potenciar el efecto a largo plazo de las intervenciones públicas, no solo en la reducción de la adicción digital, sino potencialmente en otros ámbitos del cambio de comportamiento en los que la formación de hábitos desempeña un papel destacado. Además, los profesionales del marketing social también podrían encontrar en esta herramienta una oportunidad para introducir nuevos productos que ayuden a los consumidores a hacer frente a los comportamientos adictivos.

Una dirección clara para futuras investigaciones es profundizar en el mecanismo o mecanismos psicológicos de las intervenciones con doble objetivo. En el capítulo 1 se propuso que los sujetos con mayores logros en el objetivo dual señalaban que eran menos adictos a las redes sociales de lo que creían, manteniendo así un menor consumo de redes sociales tras el tratamiento. La prueba de este efecto indirecto no fue concluyente, pero sí indicativa de un fuerte efecto directo del logro del objetivo dual, que podría apuntar a otros posibles mecanismos psicológicos en juego. Dado que el tamaño de la muestra del estudio longitudinal del Capítulo 1 era pequeño ($N = 43$), la prueba de mediación debería confirmarse en un estudio futuro con una muestra mayor.

Además, otra dirección de investigación interesante para el futuro consiste en examinar cómo la distancia entre el par de actividades puede afectar a la eficacia de una intervención de doble objetivo. Caminar (capítulo 2) crea una distancia física entre los consumidores y sus teléfonos inteligentes, protegiéndolos eficazmente de estímulos como las notificaciones push que desencadenan el deseo de utilizar sus teléfonos. En cambio, utilizar Duolingo (capítulo 1) en el mismo dispositivo móvil en el que se accede con frecuencia a las redes sociales no proporciona el mismo efecto de distanciamiento. Aunque los contextos de ambos estudios difieren a lo largo de la dimensión de la distancia entre las señales de actividad, no examinaron directamente cómo varía la eficacia de la intervención de doble objetivo en función de la fuerza de la señal basada en su proximidad al consumidor.

La literatura anterior es relativamente silenciosa en cuanto a la eficacia global de las intervenciones para reducir el consumo de teléfonos inteligentes y medios sociales, y el efecto causal de este consumo sobre diversas medidas de bienestar aún no está claro. Por lo tanto, el capítulo 3 trata de comparar la eficacia de las intervenciones existentes para reducir el uso problemático de teléfonos inteligentes y redes sociales, y examina si esta reducción también está vinculada a la mejora del bienestar. Mediante la organización de las pruebas causales halladas en los estudios de intervención en torno a un marco fundamentado, el capítulo ha logrado sintetizar ideas interesantes utilizando modelos metaanalíticos multinivel y multivariantes. Entre ellos, se observó que las intervenciones tenían tamaños de efecto de pequeños a medianos sobre la reducción del consumo y la mejora de las consecuencias. Es importante destacar que las heterogeneidades en los tamaños de los efectos son significativas, y algunas de ellas (por ejemplo, la edad y el sexo) podrían tener una influencia especialmente significativa y potencialmente impactante en los diseños de futuras intervenciones.

Aparte de las conclusiones sobre la eficacia de las intervenciones del capítulo 3, es crucial señalar que los metanálisis utilizaron únicamente pruebas causales. Por lo tanto, es una de las primeras investigaciones / síntesis de investigación que proporciona una respuesta definitiva al debate sobre si existe una relación causal entre el uso problemático de teléfonos inteligentes / medios sociales y resultados importantes como la salud mental y la productividad. El hecho de que las intervenciones mejoren varias medidas de bienestar proporciona un fuerte apoyo causal a esta relación. Además de zanjar el debate académico, este capítulo ha abierto un camino para diseñar intervenciones más eficaces y eficientes.

En el capítulo 3 también se observó que las intervenciones cuyo objetivo principal es la reducción del consumo de medios sociales/teléfonos inteligentes (es decir, las intervenciones Límite y Límite Plus) tienden a ser mejores en la reducción de su consumo, pero peores que las intervenciones (es decir, No Límite) cuyo objetivo principal es la mejora de las consecuencias. Esto podría explicar por qué la intervención de doble objetivo, que pertenece a la familia Límite Plus, redujo el consumo de medios sociales (capítulo 1) y teléfonos inteligentes (capítulo 2), pero no mejoró de forma consistente los resultados de bienestar. Esta podría ser una dirección interesante para futuras investigaciones.

En general, el marco fundamentado del capítulo 3 y el metaanálisis de las pruebas causales de un vínculo entre el uso problemático de teléfonos inteligentes y redes sociales y el bienestar revelan tres áreas destacadas en las que se necesita más investigación. En primer lugar, la bibliografía carece de intervenciones centradas en la mejora del rendimiento y la productividad. En segundo lugar, hay un desequilibrio en la investigación entre los tres tipos de intervención identificados por el enfoque basado en medidas específicas de bienestar. En tercer lugar, escasean los estudios de intervención con una duración del seguimiento posterior al tratamiento distinta de cero. Dada la preocupación de los responsables políticos por los efectos a largo plazo de las intervenciones, es crucial que las futuras intervenciones incorporen períodos de seguimiento más largos para evaluar la persistencia de sus efectos.

Appendix A. Comparison with related past interventions

Table A1.

Review of past interventions to reduce social media or smartphone usage

Article	Intervention	Incentive structure	Alternative activity	Treatment length	Long-term follow up	Measures of screen time/social media usage	Measures of alt. activity
Esmaeili & Ahmadi (2018)	Subjects set their own limit of social media usage. An app suggests alternative activities to replace social media use.	N/A	Yes; activity type suggestions only	14 days	No	Objective	Subjective
Brailovskaia et al. (2022)	Subjects were asked to reduce social media usage by 30 minutes per day, and asked to increase physical activity (PA) by 30 minutes per day.	N/A	Yes; suggested increase in daily physical activity by 30 minutes	14 days	6 months	Subjective	Subjective
Reed et al. (2023)	Subjects were asked to decrease social media by 15 minutes per day, and participate in another leisure activity (suggestions provided, such as reading, exercise, non-computing related)	N/A	Yes; activity type suggestions only	84 days	No	Objective	Subjective
Allcott et al. (2020)	Subjects were asked to deactivate Facebook (FB)	Deactivate FB and receive $102	N/A	28 days	28 days	Objective	N/A
Allcott et al. (2022)	Subjects: - set their own limit of social media usage (no incentive) Vs. - decrease by an amount and get incentivized for it (see Incentive structure).	After a period of 21 days of incentivization announcement, incentive of $2.5 per hour of social media time reduction for 21	N/A	42 days	42 days	Objective	N/A

		days, for a maximum of $150.					
Collis & Eggers (2022)	Subjects were asked to limit social media to max 10 minutes per day.	€ 1,000 (1-in-100 chance) lottery for those students staying under 10-minute limit	N/A	70 days	77 days	Objective	N/A
Stanley et al. (2022)	Subjects were asked to reduce smartphone usage by 30%/40%/50% (for 1, 1, 5 days) of which reduction of social media is 26%/33% (for 1, then 6 days)	$10 for each day reduction is achieved	N/A	7 days	21 days	Objective	N/A
This paper	- Chapter 1: subjects chose between (a) daily decrease of 30% of social media (single-target) OR (b) single target PLUS 20 minutes of language learning on Duolingo (dual-target) - Chapter 2: daily decrease of 30% of smartphone screen time (single-target) vs. single-target plus daily 2,000 steps (dual-target)	- Chapter 1: €1/day of single-target achievement OR €2/day of dual-target achievement - Chapter 2: €2/day of target achievement in both intervention arms	- Chapter 1: Yes; use Duolingo, time-tracked - Chapter 2: Yes; Walking, step-count-tracked	- Chapter 1: 20 days - Chapter 2: 20 days	- Chapter 1: 13 days - Chapter 2: 47 days	Objective	Objective

Appendix B. Chapter 1

Figure B1

Examples of screenshots submitted by subjects (Left: Social media; Right: Duolingo)

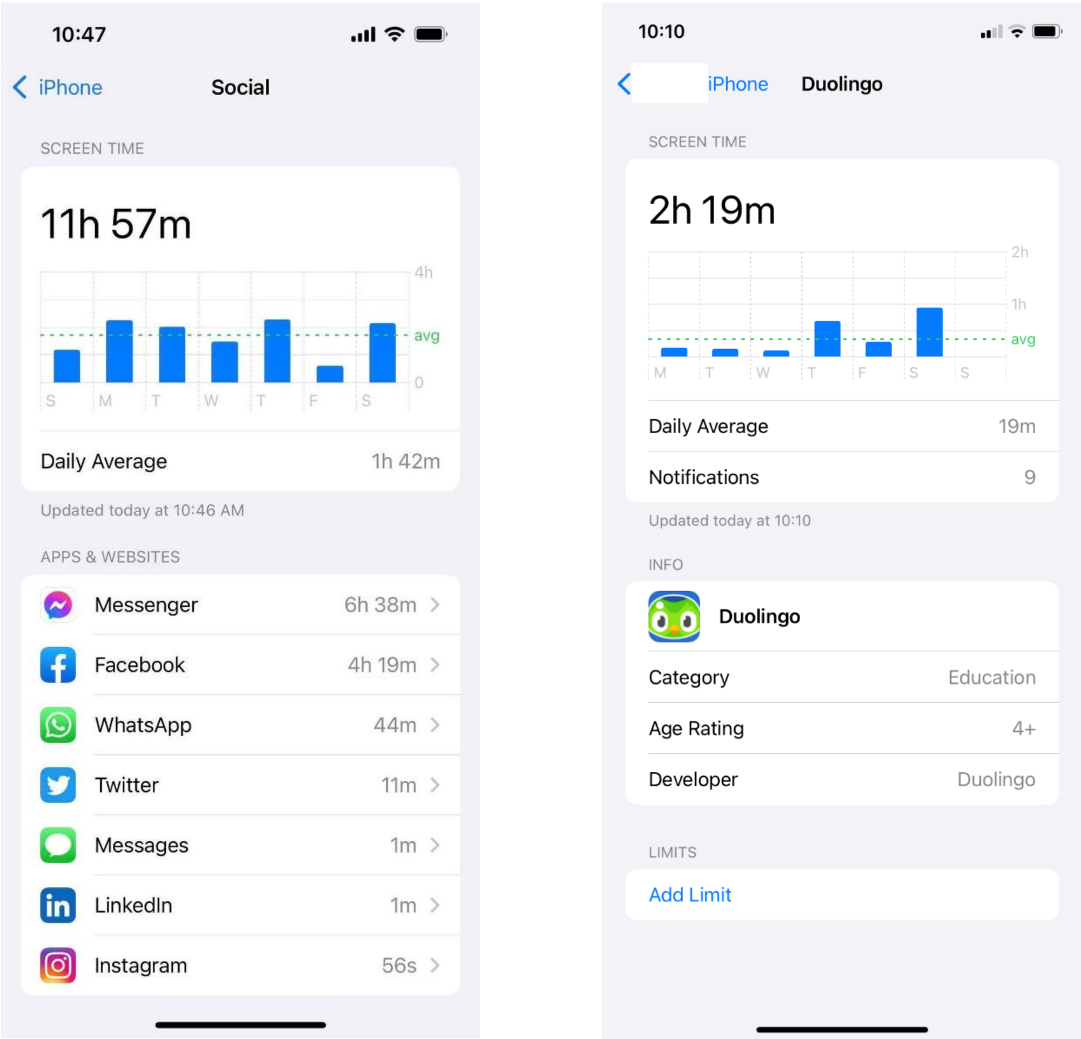


Figure B2

SMS template

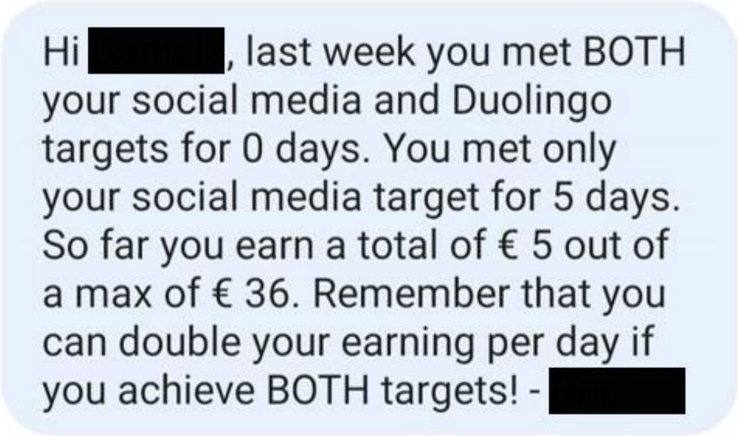


Table B3

Baseline characteristics of subjects (Chapter 1)

	Mean (SD)
Demographics	
Age	21.6 (1.79)
Female	72.1%
Bachelor	62.8%
Smartphone variables	
iOS (Apple iPhone)	88.4%
Owner of two phones	9.3%
Monthly data allowance	
• 5-10Gb	11.6%
• 10-20Gb	11.6%
• 20-30Gb	27.9%
• >30Gb	41.9%
Proportion of social media time using mobile	73.5% (18.98)
Proportion of mobile usage time spent on social networking	61.6% (18.48)
Willingness to reduce social media usage	4.8 (1.51)
Has previously attempted to reduce social media usage	79.1%
Frequency of notifications (1 = Not at all; 7 = Very much)	5.12 (1.58)
Language profile variables	
Number of languages with high proficiency	2.47 (0.74)
Proportion of English as a second language speakers	72.1%
Proportion of subjects not learning any new language	25.6%
Proportion of subjects already using Duolingo	41.9%
Willingness to use Duolingo to learn a language	4.77 (1.94)

B4

Survey questions

Addiction scale (1 = Never; 5 = Always)

Over the past 3 weeks, how often have you...

Been worried about missing out on things online when not checking your phone?

Used social media longer than intended?

Preferred to use your phone rather than interacting with your partner, friends, or family?

Felt your performance in school or at work suffers because of the amount of time you use social media?

Lost sleep due to using social media late at night?

Felt anxious when you don't have your phone?

Well-being scale (1 = Strongly Disagree; 7 = Strongly Agree)

Over the past 3 weeks,…

I was a happy person.

I was satisfied with my life.

I felt anxious.

I felt depressed.

I could concentrate on what I was doing.

I was easily distracted.

I slept well.

Table B5

Correlation Matrix (Chapter 1)

	1	2	3	4	5	6	7	8	9	10
1. Age	-									
2. Gender (Female = 1)	-0.12	-								
3. OS (Android = 1)	0.25	-0.26	-							
4. Baseline social media use (hour)	-0.14	-0.14	-0.04	-						
5. Baseline Duolingo usage (hour)	-0.13	0.24	-0.10	-0.19	-					
6. Change in social media use (hour)†	-0.06	0.07	-0.15	-0.38*	0.13	-				
7. Dual-target achievement (days)	0.05	-0.04	0.00	0.07	-0.14	-0.49***	-			
8. Single-target achievement (days)	-0.05	0.04	0.00	-0.08	0.14	0.48**	-1.00***	-		
9. Target replacement ratio (TRR)	-0.05	0.23	-0.07	-0.72***	0.71***	0.36*	-0.13	0.14	-	
10. Change in perceived addiction	-0.01	-0.19	0.24	0.11	-0.09	0.30	-0.32*	0.31*	-0.03	-
11. Change in perceived well-being	-0.10	0.24	-0.13	0.06	-0.12	-0.09	-0.13	0.12	-0.01	-0.10

Note: *$p < .05$, ** $p < .01$, *** $p < .001$; †Post-treatment - Baseline

Table B6

Descriptive statistics (standard deviation in parentheses)

Period	Pre-treatment	Treatment	Post-treatment
Social media usage (hour)	3.08 (1.33)	2.05 (1.44)	2.85 (1.47)
Duolingo usage (hour)	0.064 (0.16)	0.216 (0.29)	0.041 (0.15)

Figure B6.1

Timeline of Social Media usage change. Error bars are standard errors. Three days were allocated between Period 1 and Period 2 to complete a comprehension check.

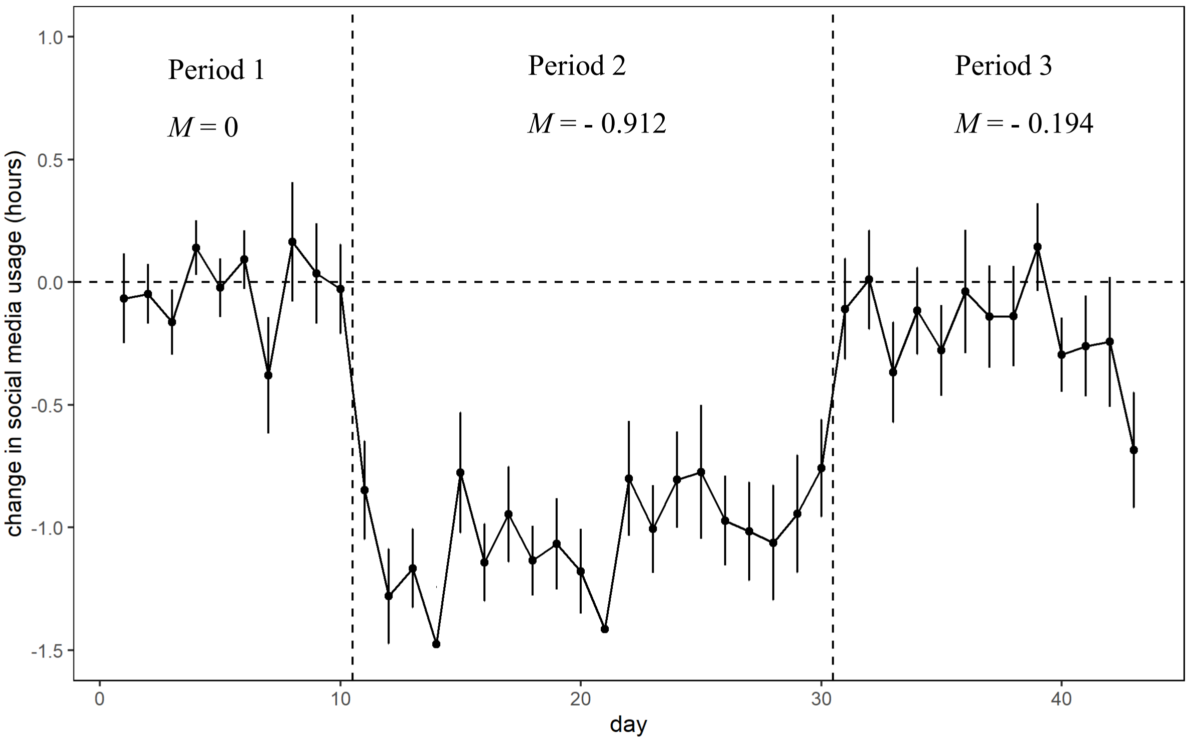


Figure B6.2

Timeline of Duolingo usage change. Error bars are standard errors. Three days were allocated between Period 1 and Period 2 to complete a comprehension check.

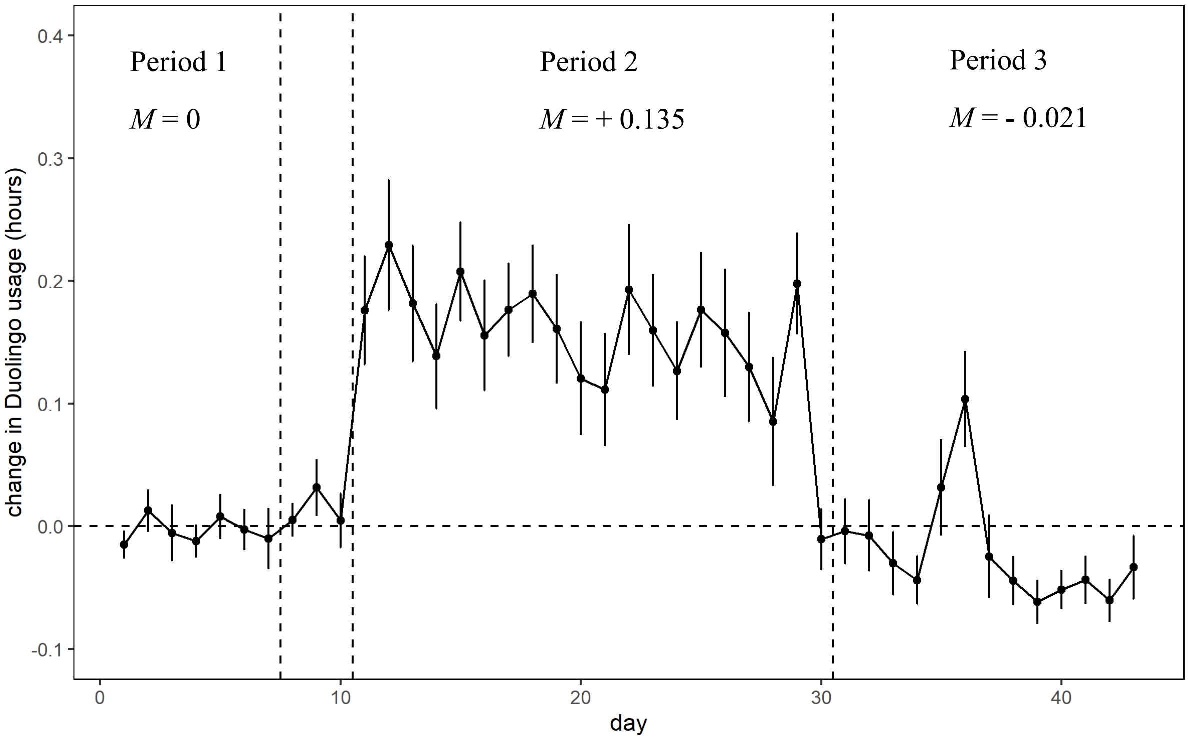


Table B7

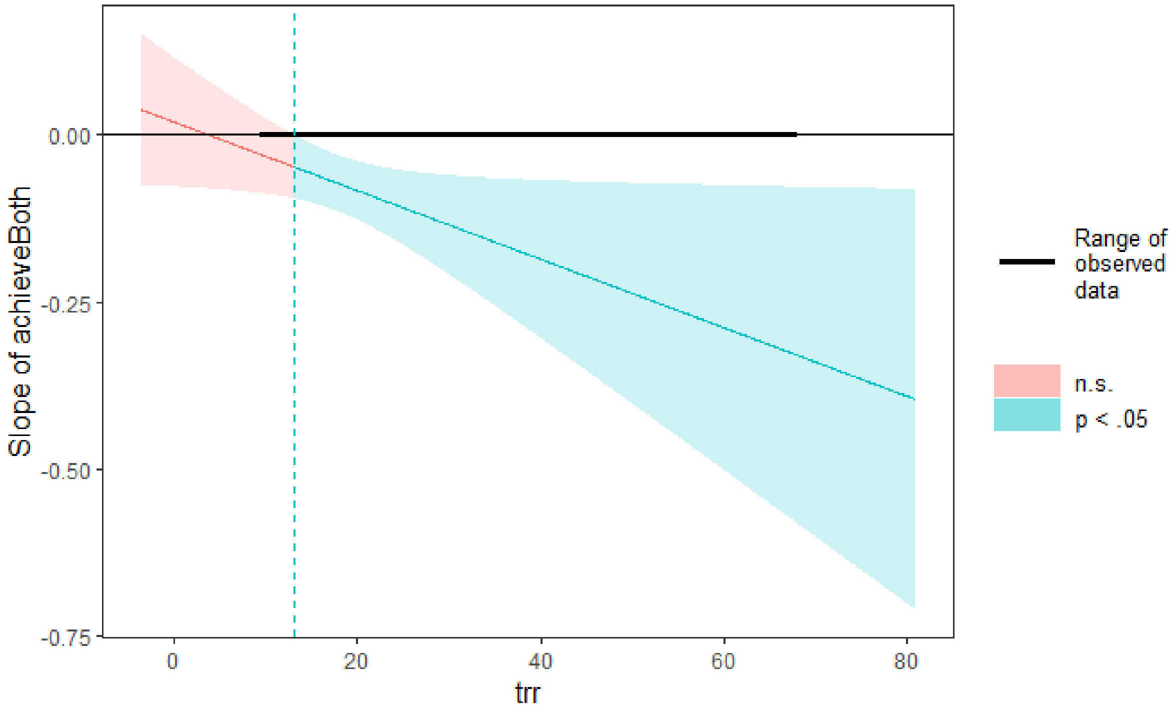
Panel regression analyses of Duolingo usage time (Chapter 1).

	Dependent variable = Duolingo Usage (hour)	
	Model 1	Model 2
Treatment Period (Period 1)	0.150***	0.037
	(0.032)	(0.030)
Post-Treatment Period	-0.025	-0.042
(Period 2)	(0.021)	(0.025)
AchieveBoth (days)		-0.003
		(0.002)
AchieveBoth × Period 1		0.028***
		(0.004)
AchieveBoth × Period 2		0.004**
		(0.002)
Constant	0.066***	0.078***
	(0.018)	(0.022)
Observations	1,520	1,520
R^2	0.114	0.268
Adjusted R^2	0.113	0.266
Residual Std. Error	0.234 (df = 1517)	0.213 (df = 1514)
F Statistic	97.628*** (df = 2; 1517)	110.984*** (df = 5; 1514)

Note: * $p < .10$; ** $p < .05$; *** $p < .01$. The base category is the baseline period. Standard errors (in parentheses) are robust and clustered at the subject level.

Figure B8

Johnson-Neyman plot. Generated by R package interactions (Long, 2019).



B9

Alternative operationalization of Target Replacement Ratio (TRR) x Dual-target achievement

I can more directly examine the effect of the interaction between TRR and Dual-target achievement on Post – Pre (post-treatment minus baseline) social media usage, by substituting this interaction with the change in the ratio of Duolingo-to-social-media usage from baseline to treatment period. This change in ratio can be computed for each individual:

$$\Delta \frac{Duolingo\ usage}{Social\ media\ usage} = \frac{Duolingo\ usage_{Tx}}{Social\ media\ usage_{Tx}} - \frac{Duolingo\ usage_{bsl}}{Social\ media\ usage_{bsl}}$$

The ratio can be interpreted as how different each person is replacing their social media usage with Duolingo usage during the Treatment period relative to their baseline. A positive sign of this ratio signifies that, compared to baseline, there is a higher proportion of social media replacement using Duolingo by an individual, and vice versa. Thus, this measure resembles the concept of TRR x Dual-target achievement as seen in section 3.3.3 of the manuscript.

Next, I run a linear regression predicting Post – Pre social media usage with the ratio change. Table B8.1 shows the results of this regression. The coefficient for the ratio change is negative, which means that an increase in the Ratio (i.e. higher proportion of replacement) results in lower post-treatment social media usage relative to baseline. This result corroborates the main finding of section 3.3.3, that people with higher TRR and achieved more dual-target days would have lower post – pre social media usage.

Table B9.1

OLS regression of Post-Pre social media usage on change in Ratio of Duolingo-to-social-media-usage. Higher Ratio change predicts lower post-treatment usage.

	Dependent variable = Post – Pre social media usage	
	Model 1	Model 2
Duolingo-to-social-media-	-0.85**	-0.90**
usage Ratio change	(0.40)	(0.41)
Age		-0.02
		(0.08)
Gender (Female)		0.07
		(0.30)
OS (iOS)		0.43
		(0.43)
Constant	-0.03	-0.07
	(0.14)	(1.77)
Observations	43	43

R^2	0.10	0.13
Adjusted R^2	0.08	0.04
Residual Std. Error	0.83 (df = 41)	0.85 (df = 38)
F Statistic	4.54** (df = 1; 41)	1.47 (df = 4; 38)

Note: * *p* < .10; ** *p* < .05; *** *p* < .01.

B10. At-payment optional survey

Table B10.1

After the incentive period ended (when you no longer have any target), did you continue using Duolingo? (N = 23)

Response	N	%
Yes	8	34.8
No, but I used a different language learning app.	2	8.7
No, and I was not using any language learning app.	13	56.5

Table B10.2

Why do you think that using a language learning app might (or might not) help you reduce social media usage? (N = 23)

Reason	Example quotes	N	%
Yes, replacement	Because for me clicking on the social media apps is a reflex that I have when I'm free/when I want to free my mind a bit. Replacing some of those times with the language apps takes away from those times.	7	35
Yes, distraction	I used it to distract myself from social media.	1	5
No, forced use	I had to force myself using the app and when the incentive was over I completely stopped. Social media on the other hand is always a source of entertainment and isn't forced.	1	5
No, app attached to phone	I think that the problem is actually using the phone (independently of what app) because anyways you'll end up checking social media.. if i want to avoid this, i'll just avoid using the phone all in all	3	15
No, not as engaging	- It is not as fun as going to social media where you do not have to concentrate - No. It's a boring alternative	5	25
No, not enough to matter	I don't think so, not by the same amount that was imposed during the experiment	1	5
No, no use	For me, I learn Spanish from a tutor. So, App (Duolingo) didn't do anything for me	1	5
No, social media more functional	I had to use WhatsApp the majority of time simply because my peers and I communicate about homework, class questions, important exams coming, so it was hard to restrict myself…	1	5

B11.

Deviations from the Pre-analysis Plan

Pre-analysis Plan: https://aspredicted.org/blind.php?x=D3H_HPR

H1 is a rephrased version of hypotheses 1 and 2 in the PAP. However, the PAP's hypothesis 1 and 2 were phrased such that only dual-target is mentioned, but the possibility existed that days with only single-target and no-target achievement might exist as well, so I combined them into H1 in the chapter to clarify these other two types of targets that exist alongside dual-target

H2 corresponds to hypothesis 3 in the PAP. However, in hindsight, that hypothesis does not make sense because even if participants had high TRR but did not actually comply to the treatment, then one should not expect them to have better post-treatment outcome.

So I altered this into H2 of the thesis.

H3 corresponds to hypothesis 4 in the PAP. It is rephrased to be consistent in terminology usage with previous hypotheses.

H4 corresponds to hypothesis 5 in the PAP. Admittedly they are not the same. But I found that well-being measurement improved after the intervention, regardless of type of target achievement, so I deviated from the PAP to highlight this result.

Appendix C. Chapter 2

Figure C1

Examples of screenshots (Left: Screen time; Right: Step count)

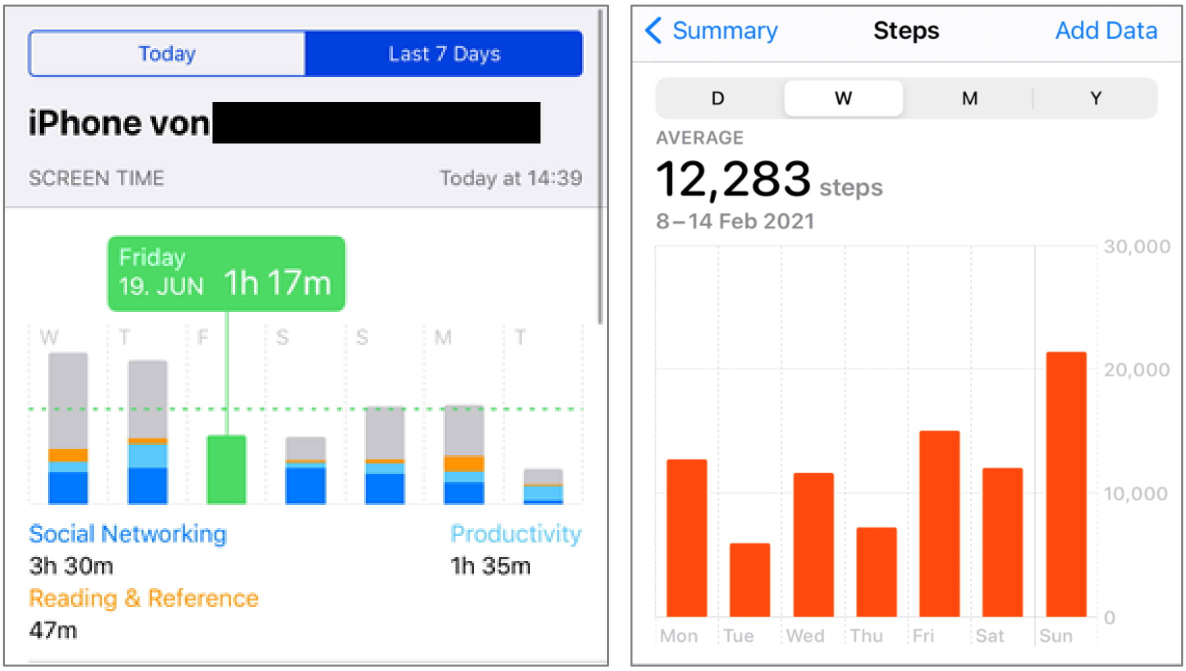


Table C2

Baseline characteristics of subjects across conditions (Chapter 2)

	Mean (SD)			
	Control (N = 33)	**Dual-Target (N = 38)**	**Single-Target (N = 39)**	**Balance check**
Demographics				
Age	25.0 (4.19)	24.3 (3.28)	24.0 (3.18)	F(2,107) = 0.68 (p = 0.509)
Female	45%	58%	69%	χ^2(2) = 1.22 (p = 0.544)
Smartphone variables				
Baseline screen time	5.15 (1.80)	4.92 (1.72)	4.78 (2.00)	F(2, 107) = 0.36, p = 0.698
iOS (Apple iPhone)	85%	82%	85%	χ^2(2) = 0.07 (p = 0.965)
Owner of >= 1 phone	9%	11%	5%	χ^2(2) = 2.24 (p = 0.326)
Proportion of mobile usage time spent on social networking	48.1%	60.5%	52.8%	χ^2(2) = 1.46 (p = 0.483)
Willingness to smartphone usage	4.97 (1.59)	5.05 (1.39)	5.10 (1.37)	F(2, 107) = 0.08, p = 0.927
Has previously attempted to reduce smartphone usage	73%	87%	64%	χ^2(2) = 3.60 (p = 0.165)
Health and performance variables				
Physically active	4.18 (1.88)	4.76 (1.51)	4.51 (1.34)	F(2, 107) = 1.21, p = 0.304
Weight	68.9 (13.94)	69.8 (15.82)	64.7 (13.16)	F(2, 107) = 1.34, p = 0.267
Willingness to walk more	5.27 (1.48)	5.87 (1.17)	5.59 (1.41)	F(2, 107) = 1.71, p = 0.186
Weekly sleeping late (days)	4.70 (2.48)	5.24 (1.73)	4.67 (2.07)	F(2, 107) = 0.88, p = 0.419
Weekly tiredness (days)	4.12 (2.42)	4.39 (1.70)	4.44 (2.01)	F(2, 107) = 0.24, p = 0.785
Satisfaction with life	4.82 (1.49)	4.76 (1.15)	4.92 (1.26)	F(2, 107) = 0.15, p = 0.860
Satisfaction with health	4.00 (1.41)	3.82 (1.52)	4.05 (1.12)	F(2, 107) = 0.32, p = 0.730
Productivity	4.39 (1.37)	4.21 (1.42)	4.28 (1.17)	F(2, 107) = 0.17, p = 0.842
Anxiety (without phone)	4.52 (1.84)	5.21 (1.56)	5.03 (1.81)	F(2, 107) = 1.49, p = 0.230
Delayed Gratification tendency	4.22 (1.38)	3.93 (0.92)	4.26 (0.96)	F(2, 107) = 1.06, p = 0.350
Negative Affect (PANAS)	2.26 (0.72)	2.18 (0.61)	2.13 (0.60)	F(2, 107) = 0..40, p = 0.670
Positive Affect (PANAS)	3.48 (0.71)	3.46 (0.57)	3.44 (0.73)	F(2, 107) = 0.02, p = 0.976
Belief about target achievement (before intervention; days)	-	12.3 (4.49)	12.4 (4.69)	F(1, 73) = 0.03, p = 0.874
Perceived target difficulty (before intervention)	-	4.87 (1.30)	4.92 (1.26)	F(1, 73) = 0.03, p = 0.865

C3. Average change in mobile usage across conditions

Table C3.1

Average (standard deviation) screen time across conditions and periods

	Control	Single-target	Dual-target
Baseline	5.15 (1.80)	4.92 (1.72)	4.78 (2.00)
Period 1 (Treatment)	5.23 (1.93)	3.59 (1.56)	3.93 (1.78)
Period 2 + 3 (Early + Late Post-Treatment)	5.36 (1.68)	4.52 (2.07)	4.50 (1.60)

There was no difference in baseline screen time across conditions, $F(2, 107) = 0.36$, $p = 0.698$.

Figure C3.2 compares the average change in mobile usage (from baseline) between DT and C conditions. In period 1 and 2, the average reduction in the DT condition was 0.98 hours (59 minutes) and 0.70 hours (42 minutes), respectively, compared to the C condition. In period 3, subjects in the DT condition lowered their usage by 0.09 hours (5 minutes) from baseline, while the C condition increased its usage by 0.46 hours (28 minutes).

Figure C3.3 compares the average change in mobile usage (from baseline) between ST and C conditions. In period 1 and 2, the average reduction in the ST condition was 1.28 hours (77 minutes) and 0.51 hours (or 31 minutes), respectively, compared to the C condition. In period 3, subjects in the DT condition lowered their usage by 0.34 hours (20 minutes) from baseline, while the C condition increased its usage by 0.46 hours (28 minutes).

Figure C3.2

Average daily change in mobile usage from baseline (Control vs. Dual-target). Standard error bars are shown.

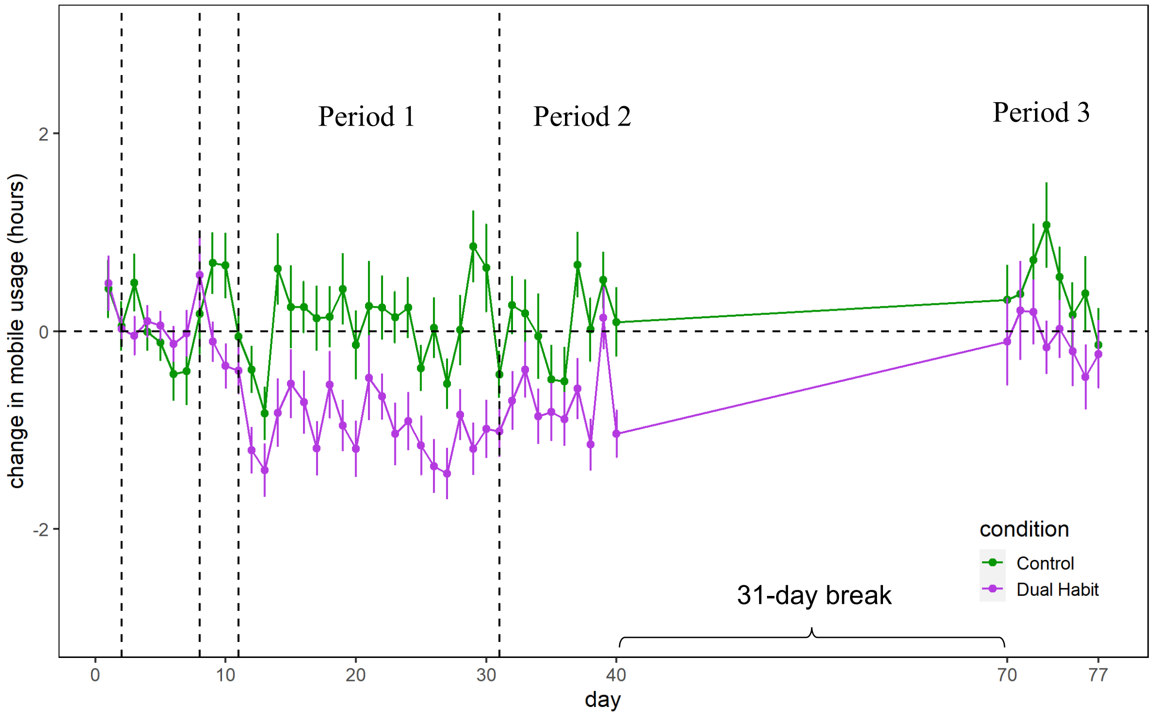
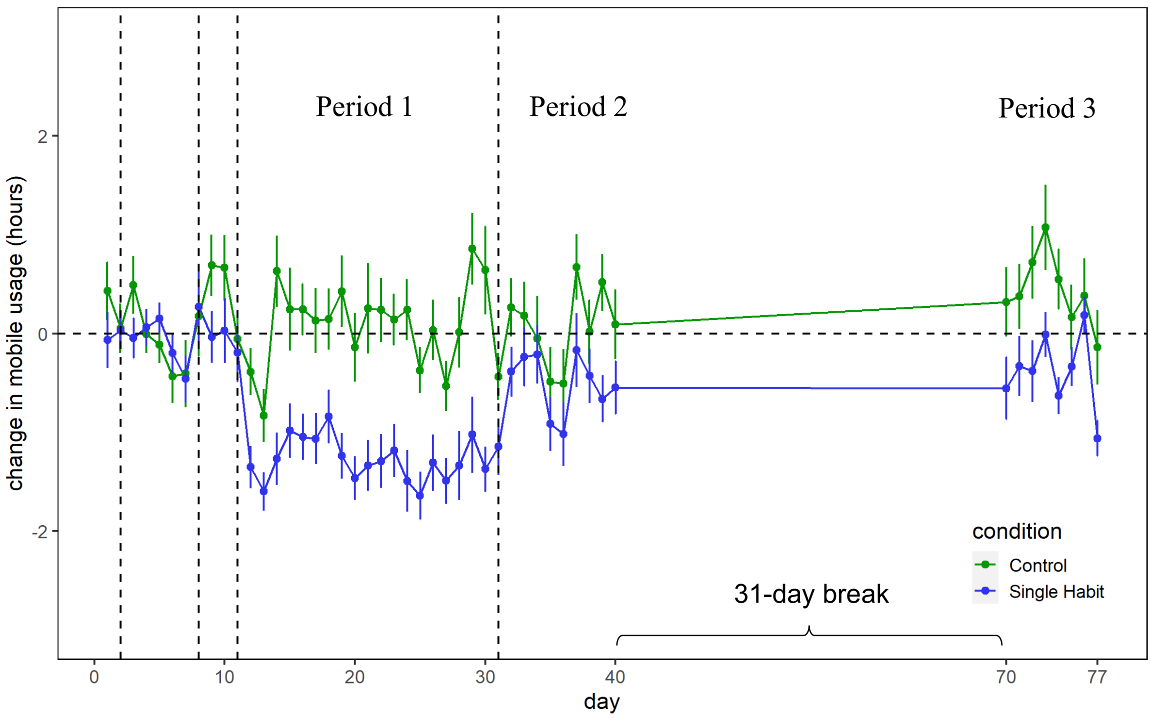


Figure C3.3

Average daily change in mobile usage from baseline (Control vs. Single-target). Standard error bars are shown.



C4. Regression analysis of screen time

I compared screen time among the three conditions using difference-in-difference ordinary least squares (OLS) regression. I regress subjects' mobile usage on (1) the treatment variables (C is the base category, DT and ST are dummy variables); (2) the period dummy variable which captures the change in usage across each period (period 0 or baseline is the base category); and (3) the interactions between treatment and period dummies to capture the changes in usage from the baseline period in the particular treatment condition. Standard errors are clustered at the subject-level to account for potential within-subject correlation. Regression estimates are presented in Table C4.1. The treatment dummy variables are not significant, indicating no baseline difference in smartphone usage. I now focus on the treatment × period interactions.

In Period 1, subjects in both ST and DT conditions decreased their usage by 1.434 hours ($p < .001$) and 1.055 hours ($p < .001$), respectively, compared to the C condition. Hence, incentives were effective in reducing usage in both treatment conditions during period 1. There was no difference between the ST and DT condition in period 1 ($p = 0.196$).

In period 2, subjects in both ST and DT conditions decreased their usage by 0.597 hours ($p = 0.07$) and 0.657 hours ($p = 0.054$), respectively, compared to the C condition.

In period 3, subjects in the ST conditions decreased their usage by 0.808 hours ($p = 0.048$) compared to the C condition. On the other hand, subjects in the DT condition did not decrease their usage compared to those in the C condition (p = 0.574). These results partially support Hypothesis 2a. I note that these results might be unreliable due to high attrition in late post-treatment responses, especially for the DT condition (Table C7). The attrition rates of the three conditions are 24% for Control, 36% for ST, and 45% for DT.

Table C4.1

Difference-in-difference OLS estimation (screen time)

	Dependent variable – Mobile Usage (hour)
ST condition	-0.251 (0.409)
	p = 0.539
DT condition	-0.253 (0.444)
	p = 0.569
Treatment (Period 1)	0.126 (0.149)
	p = 0.400
Early Post-treatment (Period 2)	0.010 (0.229)
	p = 0.966
Late Post-treatment (Period 3)	0.424 (0.284)
	p = 0.136
ST condition × Period 1	-1.434 (0.239)***
	p < .001
DT condition × Period 1	-1.055 (0.271)***
	p < .001
ST condition × Period 2	-0.597 (0.329)*
	p = 0.070
DT condition × Period 2	-0.657 (0.341)*
	p = 0.054
ST condition × Period 3	-0.781 (0.404)*
	p = 0.054
DT condition × Period 3	-0.305 (0.454)
	p = 0.502
Constant	5.130 (0.298)***
Observations	4,098
R^2	0.080
Adjusted R^2	0.077
Residual Std. Error	2.198 (df = 4,086)
F Statistic	32.236*** (df = 11; 4,086)

Note: * $p < .10$; ** $p < .05$; *** $p < .01$. The base category is control condition in baseline period. Standard errors (in parentheses) are robust and clustered at the subject level.

Table C5

Instrumental Variable Regressions (1st stage - OLS).

	Dependent variable	
	No. days achieved in treatment period – single target	No. days achieved in treatment period – dual target
	Model 1	Model 2
Control condition	-2.46***	-2.86***
	(0.18)	(0.19)
Single Target condition	3.02***	-0.89***
	(0.27)	(0.23)
Gender	-0.41	0.07
(1 = Male; 0 = Female)	(0.21)	(0.18)
Age	0.05	0.05*
	(0.03)	(0.02)
OS	1.85***	1.05***
(1 = Android; 0 = iOS)	(0.29)	(0.21)
Observations	1,760	1,760
R^2	0.24	0.12
Adjusted R^2	0.24	0.11

Note: * $p < .05$; ** $p < .01$; *** $p < .001$. Base category is Dual Target condition. Standard errors (in parentheses) are robust and clustered at the subject level. Condition is a good predictor of number of target achievement days (for both dual and single target type), and can distinguish between dual-target condition and single-target condition in terms of dual-target achievement (β = -0.89, p < .001).

Table C6

Instrumental Variable Regressions (1st stage) predicting average step count in the treatment period with Condition as the instrument. Subjects in Control condition did not walk less than those in ST condition, but walked significantly less than DT condition.

	Dependent variable = Average step count in Treatment period	
Sub-sample	ST condition vs Control	DT condition vs Control
Condition	-87.24	-886.97**
(1 = Control; 0 = Treatment condition)	(313.23)	(304.13)
Gender	755.69*	-855.27*
(1 = Male; 0 = Female)	(328.90)	(318.15)
Age	22.00	-9.99
	(35.68)	(35.12)
OS	-312.61	-476.13
(1 = Android; 0 = iOS)	(305.06)	(274.98)
Observations	639	630
R^2	0.01	0.03
Adjusted R^2	0.01	0.03

Note: * $p < .05$; ** $p < .01$; *** $p < .001$. Base category of the Condition predictor is the treatment condition in each model. Standard errors (in parentheses) are robust and clustered at the subject level.

C7

Survey measures

Well-being and Productivity measures

The following questions were asked before and after the intervention:

- How many days last week did you go to bed after midnight? (0 – 7 days)
- How many days last week did you feel tired? (0 – 7 days)
- How satisfied are you currently with your health? (1 – Not at all; 7 – Very much)
- How satisfied are you currently with your life in general? (1 – Not at all; 7 – Very much)
- How productive do you feel generally these days? (1 – Not at all; 7 – Very much)
- How anxious would you feel if you forgot to bring your mobile phone for a day? (1 – Not at all; 7 – Very much)
- Delayed gratification tendency: Using the scale provided (1 – Not at all; 7 – Very much), please indicate how much each of the following statements reflects how you typically are.
  - I am good at resisting temptation.
  - I have a hard time breaking bad habits.
  - I am able to work effectively toward long-term goals.

PANAS

This scale consists of a number of words that describe different feelings and emotions. Read each item and indicate the extent to which you generally feel that way. (1 = Very slightly or not at all; 5 = Extremely)

Interested	Irritable
Distressed	Alert
Excited	Ashamed
Upset	Inspired
Strong	Nervous
Guilty	Determined
Scared	Attentive
Hostile	Jittery
Enthusiastic	Active
Proud	Afraid

Table C8

Proportion of subjects reporting their mobile usage in different rounds (Chapter 2)

Screen Time Reports	N	Total Proportion (N/110)	Control	SH	DH
Completed baseline (report 1)	110	1.00	1.00	1.00	1.00
Completed report 2 usage (treatment)	108	0.98	1.00	1.00	0.95
Completed report 3 usage (treatment)	104	0.95	0.97	0.92	0.95
Completed report 4 usage (treatment)	102	0.93	0.94	0.92	0.92
Completed report 5 usage (treatment & early post-treatment)	97	0.88	0.91	0.87	0.87
Completed report 6 usage (early post-treatment)	92	0.84	0.88	0.85	0.79
Completed report 7 usage (early post-treatment)	96	0.87	0.91	0.90	0.82
Completed report 8 usage (late post-treatment)	71	0.65	0.76	0.64	0.55

Appendix D. Chapter 3

Table D1

Search terms

SPICE Framework keyword selection (Booth, 2006)	
SPICE	Keywords
Setting	social media, smartphone, facebook
Population	N/A (all populations are considered)
Intervention	nudg*, choice architect*, intervention, incentiv*, habit*, treatment, limit*, abstinence, restrict*
Comparator	randomized, objective data, longitudinal, experiment
Evaluation	addict*, reduc* usage, well-being, wellbeing, productivity, screen time

Example search query for PubMed (excluding reviews and meta-analyses)

("social media" OR “smartphone*” OR "facebook") AND ("nudg*" OR "choice architect*" OR "intervention" OR "incentiv*" OR "habit*" OR "treatment" OR "limit*" OR "abstinence" OR "restrict*") AND ("randomized" OR "objective data" OR “longitudinal” OR "experiment*") AND ("addict*" OR "reduc* usage" OR "well-being" OR “wellbeing” OR "productivity" OR “screen time”)

Figure D2

PRISMA diagram

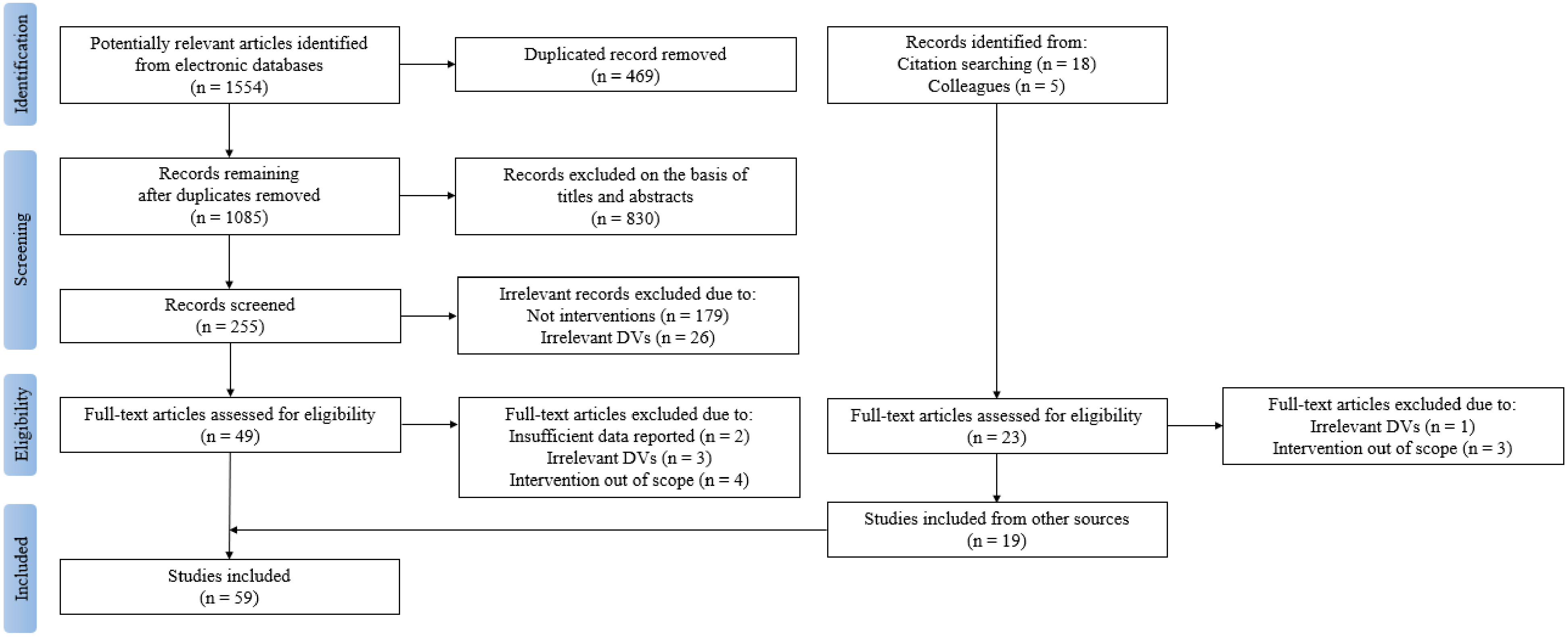


Table D3

Summary of articles

https://osf.io/vscqj/?view_only=1395397066bf451ab3202f7623687502

Table D4

Enumeration of DV subtypes (number of Effect sizes) across Intervention types

DV subtype ↓ / Intervention type →	Limit	Limit Plus	Non-Limit	Sub-total (by DV)
Screen time and social media usage amount (‘consumption’)				
Objective screen time	0	16	2	18
Objective social media time	4	12	0	16
Subjective screen time	2	2	9	13
Subjective social media time	4	6	4	14
Sub-total (by Intervention type)	10	36	15	
Downstream ‘consequence’ of intervention				
Addiction	8	18	23	49
Anxiety	7	12	11	30
Depression	9	15	9	33
Distraction	0	3	3	6
FOMO	2	0	0	2
GPA	1	3	0	4
Happiness	2	2	0	4
Life satisfaction	9	15	7	31
Loneliness	7	4	9	20
PANAS (negative)	4	6	4	14
PANAS (positive)	4	8	4	16
Self-esteem	3	0	5	8
Sleep	2	11	1	14
Stress	2	4	12	18
Sub-total (by intervention type)	60	101	88	

Table D5

Study-level characteristics across Intervention Types

	Limit	Limit Plus	Non-Limit
Number of Studies	16	35	21
Instruction Language (% in English)	62%	74%	19%
Design (% of Double-difference)	69%	74%	95%
Treatment duration (in days; mean and SD)	13.00 (19.80)	16.43 (17.99)	13.47 (9.78)
Post-treatment duration (in days; mean and SD)	15.60 (37.90)	11.57 (38.64)	23.33 (40.17)
Age (mean and SD)	23.7 (2.80)	24.4 (5.20)	19.0 (5.17)
Gender (% of female; mean and SD)	71% (10%)	64% (15%)	52% (26%)

D6

Robustness check with Impact Factor

I queried the Impact Factor (IF) of each journal for the respective year that an article in my sample was published. I used Clarivate's Journal Citation Reports as the primary source of IF. Whenever this data was not available, I used the comparable, but broader-scope open database Exaly[41].

When included as a covariate in the multivariate meta-analytic model, IF does not change the coefficient nor significance of the Instruction Language covariate. This implies that the influence of Instruction Language on the effect size is unlikely to be due to journal reputation.

I note that, when IF is included, the significance of a couple of other covariates changes. First, the effect of Design, being non-significant previously, becomes marginally significant ($\beta = -0.23$, $z = -1.83$, $p = 0.067$), which implies that single Pre-Post design might yield lower effect sizes than double-difference design. Second, the effect of gender, being significant (at the .05 level) previously, becomes marginally significant ($\beta = -0.005$, $z = -1.77$, $p = 0.077$).

[41] Journal Citation Reports: https://jcr.clarivate.com. Exaly: https://exaly.com/journals/if/.

Table D6

Multivariate Model, with and without Impact Factor as control

	With Impact Factor	*Without Impact Factor*
	β (SE)	*β (SE)*
Intercept	0.24***	0.30***
	(0.07)	(0.06)
Intervention type		
Limit	(ref)	(ref)
Limit Plus	-0.02	0.04
	(0.10)	(0.10)
Non-Limit	0.13	0.16
	(0.14)	(0.14)
DV category		
To-reduce	(ref)	(ref)
To-increase	-0.02	-0.02
	(0.04)	(0.04)
Instruction Language		
Non-English	(ref)	(ref)
English	-0.27**	-0.28**
	(0.09)	(0.09)
Design		
Double-difference	(ref)	(ref)
Single Pre-Post	-0.23^{+}	-0.10
	(0.13)	(0.12)
Treatment duration (day)	0.0004	0.0005
	(0.003)	(0.003)
Post-treatment duration (day)	0.0003	0.0003
	(0.0005)	(0.0005)
Age	-0.02*	-0.02*
	(0.01)	(0.009)
Gender (% of female)	-0.005^{+}	-0.006*
	(0.003)	(0.003)
Journal Impact Factor	0.02	-
	(0.02)	
k (no. of effect sizes)	241	247
S (no. of studies)	60	64
A (no. of articles)	48	51
R^2	0.44	0.47

Note: $^{+}p < .1$; $*p < .05$; $**p < .01$; $***p < .001$

Appendix E. Additional Analyses

Table E1

Panel regression analyses of Social Media usage time (Chapter 1). Dual-target achievement predicts post-treatment social media reduction.

	Dependent variable = Social Media Usage (hour)			
	Model 1	Model 2	Model 3	Model 4
Treatment Period	-1.029***	-0.627***	-0.627***	-0.627
(Period 1)	(0.141)	(0.159)	(0.159)	(0.159)
Post-Treatment Period	-0.235*	0.065	0.064	0.072
(Period 2)	(0.138)	(0.174)	(0.173)	(0.173)
Dual-target achievement		0.012	0.016	0.012
(days)		(0.026)	(0.028)	(0.026)
Dual-target achievement ×		-0.101***	-0.101***	-0.101***
Period 1		(0.017)	(0.017)	(0.017)
Dual-target achievement ×		-0.072***	-0.072***	-0.072***
Period 2		(0.024)	(0.024)	(0.024)
Willingness to use Duolingo			0.043	
(at baseline)			(0.077)	
Age				-0.052
				(0.061)
Gender				-0.343
(0 = Male, 1 = Female)				(0.383)
OS				0.438
(0 = Android, 1 = iOS)				(0.408)
Constant	3.081***	3.032***	2.816***	3.669**
	(0.165)	(0.208)	(0.448)	(1.494)
Observations	1,588	1,588	1,588	1,588
R^2	0.085	0.154	0.157	0.174
Adjusted R^2	0.083	0.151	0.154	0.170
Residual Std. Error	1.435 (df = 1585)	1.381 (df = 1582)	1.379 (df = 1581)	1.366 (df = 1579)
F Statistic	73.272***	57.659***	49.074***	41.515***
	(df = 2; 1585)	(df = 5; 1582)	(df = 6; 1581)	(df = 8; 1579)

Note: * p < .10; ** p < .05; *** p < .01. The base category is the baseline period. Standard errors (in parentheses) are robust and clustered at the subject level.

E2

Testing for mediation via perceived addiction

To test if post-treatment usage (relative to baseline usage, post-pre social media usage for short) is affected by dual-target achievement more than single-target achievement via perceived addiction to social media, I conducted a mediation analysis using the bootstrap method implemented by the R package *mediation* (Tingley et al., 2014). The variables entered into the model (Figure E2.1) were: dual-target achievement (independent variable, in days; higher dual-target achievement equals lower single-target achievement), perceived addiction to social media (mediator), post-pre social media usage (dependent variable).

At the first stage of the mediation model, I observed a significant dual-target achievement effect on perceived addiction to social media (β = -0.038, $t(41)$ = -2.130, p = 0.039). At the second stage of the mediation model, perceived addiction did not have a significant effect on post-pre social media usage (β = 0.215, $t(40)$ = 1.121, p = 0.269), while controlling for dual-target achievement. The direct effect on post-pre social media usage was significant (β = -0.070, $t(40)$ = -3.029, p = 0.004). Finally, a 95% confidence interval (CI) (based on 10,000 bootstrapped samples) revealed that the indirect effect of dual-target achievement on post-pre social media usage through perceived addiction to social media was not significant (point estimate = -0.008, 95% CI = [-0.029, 0.01], p = 0.286).

It is notable that, when not controlling for the effect of dual-target achievement, the effect of perceived addiction to social media on post-pre social media usage was marginally significant (β = 0.399, $t(41)$ = 2.00, p = 0.052).

Figure E2.1

Mediation model

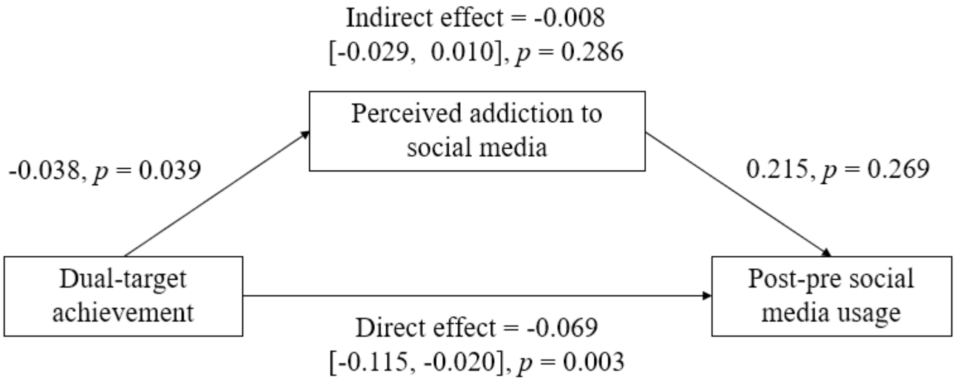


Table E3 (Chapter 2)

Instrumental Variable Regressions (2nd stage). Dual-target achievement predicts post-treatment screen time reduction above single-target achievement, even when controlling for willingness to increase step count.

	Dependent variable = Post-treatment usage - Baseline	
	Model 1	Model 2
Predicted number of days achieved in P1 - single target (β_{single})	-0.12*** (0.02)	
Predicted number of days achieved in P1 - dual target (β_{dual})		-0.22*** (0.04)
Willingness to increase step count (measured at baseline)	-0.10** (0.04)	0.02 (0.05)
Gender (1 = Male, 0 = Female)	-0.14 (0.12)	0.07 (0.12)
Age	-0.02 (0.02)	-0.04** (0.02)
OS (1 = Android; 0 = iOS)	0.23* (0.13)	0.09 (0.13)
Observations	1326	1326
R^2	0.09	0.02
Adjusted R^2	0.08	0.02

Note: * $p < .1$; ** $p < .05$; *** $p < .01$. Baseline corresponds to average mobile usage during the baseline period. Standard errors (in parentheses) are robust and clustered at the subject level.

References

42matters (n.d.). *Google Play Statistics and Trends 2024*. https://42matters.com/google-play-statistics-and-trends

Acland, D., & Levy, M. R. (2015). Naiveté, projection bias, and habit formation in gym attendance. *Management Science*, 61(1), 146-160.

Adriaanse, M. A., & Verhoeven, A. (2018). Breaking habits using implementation intentions. *The psychology of habit: Theory, mechanisms, change, and contexts*, 169-188.

Adriaanse, M. A., Van Oosten, J. M., De Ridder, D. T., De Wit, J. B., & Evers, C. (2011). Planning what not to eat: Ironic effects of implementation intentions negating unhealthy habits. *Personality and Social Psychology Bulletin*, *37*(1), 69-81.

Ahn, J., & Jung, Y. (2016). The common sense of dependence on smartphone: A comparison between digital natives and digital immigrants. *New media & society*, *18*(7), 1236-1256.

Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2020). The welfare effects of social media. *American Economic Review*, 110(3), 629-676.

Allcott, H., Gentzkow, M., & Song, L. (2022). Digital addiction. *American Economic Review*, *112*(7), 2424-63.

Alter, A. (2017). Irresistible: The rise of addictive technology and the business of keeping us hooked: Penguin.

Anderson, I. A., & Wood, W. (2021). Habits and the electronic herd: The psychology behind social media's successes and failures. *Consumer Psychology Review, 4(1)*, 83-99.

Andrews, I., & Kasy, M. (2019). Identification of and correction for publication bias. *American Economic Review*, *109*(8), 2766-2794.

Angrist, J. D., & Imbens, G. W. (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American statistical Association, 90*(430), 431-442.

Angus, C., Latimer, N., Preston, L., Li, J., & Purshouse, R. (2014). What are the implications for policy makers? A systematic review of the cost-effectiveness of screening and brief interventions for alcohol misuse in primary care. *Frontiers in psychiatry*, *5*, 114.

Anshari, M., Alas, Y., Hardaker, G., Jaidin, J. H., Smith, M., & Ahad, A. D. (2016). Smartphone habit and behavior in Brunei: Personalization, gender, and generation gap. *Computers in human behavior*, *64*, 719-727.

App Annie (2021). *State of mobile 2021*. Retrieved from https://www.data.ai/en/go/state-of-mobile-2021.

Ashraf, N., Karlan, D., & Yin, W. (2006). Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines. *The Quarterly Journal of Economics, 121(2), 635-672.*

Banfield-Nwachi, M. (2024, February 17). 'It went nuts': Thousands join UK parents calling for smartphone-free childhood. *The Guardian*. https://www.theguardian.com/technology/2024/feb/17/thousands-join-uk-parents-calling-for-smartphone-free-childhood

Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate behavioral research*, *40*(3), 373-400.

Baumeister, R. F., & Vohs, K. D. (2007). Self-Regulation, ego depletion, and motivation. *Social and personality psychology compass*, *1*(1), 115-128.

Baumeister, R. F., Sparks, E. A., Stillman, T. F., & Vohs, K. D. (2008). Free will in consumer behavior: Self-control, ego depletion, and choice. *Journal of Consumer Psychology*, *18*(1), 4-13.

Becker, G. S., & Murphy, K. M. (1988). A theory of rational addiction. *Journal of political Economy*, *96*(4), 675-700.

Booth, A. (2006). Clear and present questions: Formulating questions for evidence based practice. *Library Hi Tech 24*(3):355–368.

Brailovskaia, J., Swarlik, V. J., Grethe, G. A., Schillack, H., & Margraf, J. (2022). Experimental longitudinal evidence for causal role of social media use and physical activity in COVID-19 burden and mental health. *Journal of Public Health*, 1-14.

Brandon, A., Ferraro, P. J., List, J. A., Metcalfe, R. D., Price, M. K., & Rundhammer, F. (2017). Do the effects of nudges persist? theory and evidence from 38 natural field experiments (No. w23277). *National Bureau of Economic Research*.

Bravata, D. M., Smith-Spangler, C., Sundaram, V., Gienger, A. L., Lin, N., Lewis, R., ... & Sirard, J. R. (2007). Using pedometers to increase physical activity and improve health: a systematic review. *JAMA*, 298(19), 2296-2304.

Busch, P. A., & McCarthy, S. (2021). Antecedents and consequences of problematic smartphone use: A systematic literature review of an emerging research area. *Computers in human behavior, 114*, 106414.

Buyalskaya, A., Ho, H., Milkman, K. L., Li, X., Duckworth, A. L., & Camerer, C. (2023). What can machine learning teach us about habit formation? Evidence from exercise and hygiene. *Proceedings of the National Academy of Sciences*, *120*(17), e2216115120.

Cadario, R., & Chandon, P. (2020). Which healthy eating nudges work best? A meta-analysis of field experiments. *Marketing Science, 39*(3), 465-486.

Cappelen, A. W., Charness, G., Ekström, M., Gneezy, U., & Tungodden, B. (2017). Exercise improves academic performance. *NHH Dept. of Economics Discussion Paper*, (08).

Chadi, A., Mechtel, M., & Mertins, V. (2022). Smartphone bans and workplace performance. *Experimental Economics*, *25*(1), 287-317.

Charness, G., & Gneezy, U. (2009). Incentives to exercise. *Econometrica*, 77(3), 909-931.

Chekroud, S. R., Gueorguieva, R., Zheutlin, A. B., Paulus, M., Krumholz, H. M., Krystal, J. H., & Chekroud, A. M. (2018). Association between physical exercise and mental health in 1· 2 million individuals in the USA between 2011 and 2015: a cross-sectional study. *The Lancet Psychiatry, 5(9)*, 739-746.

Chen, B., Liu, F., Ding, S., Ying, X., Wang, L., & Wen, Y. (2017). Gender differences in factors associated with smartphone addiction: a cross-sectional study among medical college students. *BMC psychiatry*, *17*, 1-9.

Chen, I. H., Pakpour, A. H., Leung, H., Potenza, M. N., Su, J. A., Lin, C. Y., & Griffiths, M. D. (2020). Comparing generalized and specific problematic smartphone/internet use: Longitudinal relationships between smartphone application-based addiction and social media addiction and psychological distress. *Journal of behavioral addictions*, *9*(2), 410-419.

Chen, Y., Liu, X., Chiu, D. T., Li, Y., Mi, B., Zhang, Y., Ma, L., & Yan, H. (2022). Problematic Social Media Use and Depressive Outcomes among College Students in China: Observational and Experimental Findings. *International Journal of Environmental Research and Public Health, 19*(9), 4937. https://doi.org/10.3390/ijerph19094937

Cheng, C., Lau, Y. C., Chan, L., & Luk, J. W. (2021). Prevalence of social media addiction across 32 nations: Meta-analysis with subgroup analysis of classification schemes and cultural values. *Addictive behaviors*, *117*, 106845.

Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. *Psychological methods*, *19*(2), 211.

Chiang, J. T., Chang, F. C., Lee, K. W., & Hsu, S. Y. (2019). Transitions in smartphone addiction proneness among children: The effect of gender and use patterns. *PloS one, 14*(5), e0217235.

Choi, E.-H., Chun, M. Y., Lee, I., Yoo, Y.-G., & Kim, M.-J. (2020). The Effect of Mind Subtraction Meditation Intervention on Smartphone Addiction and the Psychological Wellbeing among Adolescents. *International Journal of Environmental Research and Public Health, 17*(9), 3263. https://doi.org/10.3390/ijerph17093263

Chrisafis, A. (2024, April 30). Stop children using smartphones until they are 13, says French report. *The Guardian*. https://www.theguardian.com/world/2024/apr/30/stop-children-using-smartphones-until-they-are-13-say-french-experts-in-report

Collis, A., & Eggers, F. (2022). Effects of restricting social media usage on wellbeing and performance: A randomized control trial among students. *Plos one, 17*(8), e0272416.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology, 13*(1), 3-21.

Cornil, Y., Plassmann, H., Aron-Wisnewsky, J., Poitou-Bernert, C., Clément, K., Chabert, M., & Chandon, P. (2022). Obesity and responsiveness to food marketing before and after bariatric surgery. *Journal of Consumer Psychology, 32*(1), 57-68.

Coyne, S. M., Rogers, A. A., Zurcher, J. D., Stockdale, L., & Booth, M. (2020). Does time spent using social media impact mental health?: An eight year longitudinal study. *Computers in human behavior*, *104*, 106160.

Dai, X., & Fishbach, A. (2014). How nonconsumption shapes desire. *Journal of Consumer Research, 41*(4), 936-952.

David, M. E., & Roberts, J. A. (2017). Phubbed and alone: Phone snubbing, social exclusion, and attachment to social media. *Journal of the Association for Consumer Research, 2(2)*, 155-163.

Day, G. (2021). Antitrust, Attention, and the Mental Health Crisis. *Minn. L. Rev.*, *106*, 1901.

De la Rosa, A., Olaso-Gonzalez, G., Arc-Chagnaud, C., Millan, F., Salvador-Pascual, A., García-Lucerga, C., ... & Gomez-Cabrera, M. C. (2020). Physical exercise in the prevention and treatment of Alzheimer's disease. *Journal of sport and health science, 9(5)*, 394-404.

Delfino, L. D., dos Santos Silva, D. A., Tebar, W. R., Zanuto, E. F., Codogno, J. S., Fernandes, R. A., & Christofaro, D. G. (2017). Screen time by different devices in adolescents: association with physical inactivity domains and eating habits. *The Journal of sports medicine and physical fitness, 58*(3), 318-325.

Deloitte. (2017). Device addiction shifts to device etiquette. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/us/Documents/technology-media-telecommunications/us-tmt-global-mobile-consumer-survey-2017-infographic.pdf

Deloitte. (2019). Deloitte's 2019 global mobile consumer survey. Retrieved from www2. deloitte. com/us/en/insights/industry/telecommunications/global-mobile-consumersurvey. html.

Dent, S. (2021). *Facebook whistleblower reveals identity, says company 'chooses profits over safety'*. TechCrunch. https://techcrunch.com/2021/10/04/facebook-whistleblower-reveals-identity-says-company-chooses-profits-over-safety/

Dey, M., Studer, J., Schaub, M. P., Gmel, G., Ebert, D. D., Lee, J. Y. C., & Haug, S. (2019). Problematic smartphone use in young Swiss men: Its association with problematic substance use and risk factors derived from the pathway model. *Journal of behavioral addictions, 8*(2), 326-334.

Dhaoui, C., & Webster, C. M. (2021). Brand and consumer engagement behaviors on Facebook brand pages: Let's have a (positive) conversation. *International Journal of Research in Marketing*, *38*(1), 155-175.

Dishman, R. K., & O'Connor, P. J. (2009). Lessons in exercise neurobiology: the case of endorphins. *Mental Health and Physical Activity*, 2(1), 4-9.

Dolan, P., & Galizzi, M. M. (2015). Like ripples on a pond: Behavioral spillovers and their implications for research and policy. *Journal of Economic Psychology*, *47*, 1-16.

Duval, S., & Tweedie, R. (2000). Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455-463.

Egger, F., Benzing, V., Conzelmann, A., & Schmidt, M. (2019). Boost your brain, while having a break! The effects of long-term cognitively engaging physical activity breaks on children's executive functions and academic achievement. *PloS one, 14(3)*, e0212482.

Eisingerich, A. B., Marchand, A., Fritze, M. P., & Dong, L. (2019). Hook vs. hope: How to enhance customer engagement through gamification. *International Journal of Research in Marketing*, *36*(2), 200-215.

Esmaeili Rad, M., & Ahmadi, F. (2018). A new method to measure and decrease the online social networking addiction. *Asia-Pacific Psychiatry*, *10*(4), e12330.

Fernández-Castilla, B., Jamshidi, L., Declercq, L., Beretvas, S. N., Onghena, P., & Van den Noortgate, W. (2020). The application of meta-analytic (multi-level) models with multiple random effects: A systematic review. *Behavior research methods*, *52*, 2031-2052.

Fitz, N., Kushlev, K., Jagannathan, R., Lewis, T., Paliwal, D., & Ariely, D. (2019). Batching smartphone notifications can improve well-being. *Computers in Human Behavior, 101*, 84–94. https://doi.org/10.1016/j.chb.2019.07.016

Fu, S., Li, H., Liu, Y., Pirkkalainen, H., & Salo, M. (2020). Social media overload, exhaustion, and use discontinuance: Examining the effects of information overload, system feature overload, and social overload. Information Processing & Management, 57(6), 102307.

Galizzi, M. M., & Whitmarsh, L. (2019). How to measure behavioral spillovers: a methodological review and checklist. *Frontiers in psychology*, *10*, 342.

Gardner, B., Lally, P., & Wardle, J. (2012). Making health habitual: the psychology of 'habit-formation'and general practice. *British Journal of General Practice*, *62*(605), 664-666.

Giunchiglia, F., Zeni, M., Gobbi, E., Bignotti, E., & Bison, I. (2018). Mobile social media usage and academic performance. *Computers in Human Behavior*, *82*, 177-185.

Hall, J. A., Xing, C., Ross, E. M., & Johnson, R. M. (2021). Experimentally manipulating social media abstinence: results of a four-week diary study. *Media Psychology*, 1-17.

Hancock, J., Liu, S. X., Luo, M., & Mieczkowski, H. (2022). Psychological well-being and social media use: A meta-analysis of associations between social media use and depression, anxiety, loneliness, eudaimonic, hedonic and social well-being. *Available at SSRN:* https://dx.doi.org/10.2139/ssrn.4053961.

Hanley, S. M., Watt, S. E., & Coventry, W. (2019). Taking a break: The effect of taking a vacation from Facebook and Instagram on subjective well-being. *PLOS ONE, 14*(6), e0217743. https://doi.org/10.1371/journal.pone.0217743

Harrer, M., Cuijpers, P., Furukawa, T., & Ebert, D. (2021). *Doing meta-analysis with R: A hands-on guide*. Chapman and Hall/CRC.

Harris, M. C., & Kessler, L. M. (2019). Habit formation and activity persistence: Evidence from gym equipment. *Journal of Economic Behavior & Organization, 166*, 688-708.

Heath, C., & Soll, J. B. (1996). Mental budgeting and consumer decisions. *Journal of consumer research*, *23*(1), 40-52.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in medicine*, *21*(11), 1539-1558.

Hisler, G., Twenge, J. M., & Krizan, Z. (2020). Associations between screen time and short sleep duration among adolescents varies by media type: evidence from a cohort study. *Sleep medicine*, *66*, 92-102.

Hoare, E., Stavreski, B., Jennings, G. L., & Kingwell, B. A. (2017). Exploring motivation and barriers to physical activity among active and inactive Australian adults. *Sports, 5*(3), 47.

Huang, C. (2022). A meta-analysis of the problematic social media use and mental health. *International Journal of Social Psychiatry*, *68*(1), 12-33.

Hudson, S., Huang, L., Roth, M. S., & Madden, T. J. (2016). The influence of social media interactions on consumer–brand relationships: A three-country study of brand perceptions and marketing behaviors. *International Journal of Research in Marketing*, *33*(1), 27-41.

Hughes, N., & Burke, J. (2018). Sleeping with the frenemy: How restricting 'bedroom use' of smartphones impacts happiness and wellbeing. *Computers in Human Behavior, 85*, 236–244. https://doi.org/10.1016/j.chb.2018.03.047

Hunt, M. G., Marx, R., Lipson, C., & Young, J. (2018). No more FOMO: Limiting social media decreases loneliness and depression. *Journal of Social and Clinical Psychology*, 37(10), 751-768.

Jones-Jang, S. M., Heo, Y. J., McKeever, R., Kim, J. H., Moscowitz, L., & Moscowitz, D. (2020). Good news! Communication findings may be underestimated: Comparing effect sizes with self-reported and logged smartphone use data. *Journal of Computer-Mediated Communication*, *25*(5), 346-363.

Kaye, L. K., Orben, A., A. Ellis, D., C. Hunter, S., & Houghton, S. (2020). The conceptual and methodological mayhem of "screen time*"*. *International Journal of Environmental Research and Public Health, 17*(10), 3661.

Kemp, S. (2021). Digital 2021: Global Overview Report. Retrieved from: https://datareportal.com/reports/digital-2021-global-overview-report.

Kim, J., Park, J., Lee, H., Ko, M., & Lee, U. (2019). LocknType: Lockout Task Intervention for Discouraging Smartphone App Use. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–12. https://doi.org/10.1145/3290605.3300927

King, A. L. S., Valenca, A. M., Silva, A. C. O., Baczynski, T., Carvalho, M. R., & Nardi, A. E. (2013). Nomophobia: Dependency on virtual environments or social phobia?. *Computers in Human Behavior, 29*(1), 140-144.

Kotler, P., & Zaltman, G. (1971). Social marketing: an approach to planned social change. *Journal of marketing*, *35*(3), 3-12.

Kozinets, R., Patterson, A., & Ashman, R. (2017). Networks of desire: How technology increases our passion to consume. *Journal of Consumer Research*, *43*(5), 659-682.

Kushlev, K., & Leitao, M. R. (2020). The effects of smartphones on well-being: Theoretical integration and research agenda. *Current opinion in psychology*, *36*, 77-82.

Kvam, S., Kleppe, C. L., Nordhus, I. H., & Hovland, A. (2016). Exercise as a treatment for depression: a meta-analysis. *Journal of affective disorders, 202*, 67-86.

Laibson, D. (2001). A cue-theory of consumption. *The Quarterly Journal of Economics*, *116*(1), 81-119.

Lanaj, K., Johnson, R. E., & Barnes, C. M. (2014). Beginning the workday yet already depleted? Consequences of late-night smartphone use and sleep. *Organizational Behavior and Human Decision Processes*, *124*(1), 11-23.

Lavie, C. J., Ozemek, C., Carbone, S., Katzmarzyk, P. T., & Blair, S. N. (2019). Sedentary behavior, exercise, and cardiovascular health. *Circulation research*, *124(5)*, 799-815.

Lee, I. M., & Buchner, D. M. (2008). The importance of walking to public health. *Medicine & Science in Sports & Exercise*, 40(7), S512-S518.

Lee, N. R., & Kotler, P. (2015). *Social marketing: Changing behaviors for good*. Sage Publications.

Lee, Y. K., Chang, C. T., Lin, Y., & Cheng, Z. H. (2014). The dark side of smartphone usage: Psychological traits, compulsive behavior and technostress. *Computers in human behavior*, *31*, 373-383.

Lefferts, E. C., Saavedra, J. M., Song, B. K., Brellenthin, A. G., Pescatello, L. S., & Lee, D. C. (2023). Increasing lifestyle walking by 3000 steps per day reduces blood pressure in sedentary older adults with hypertension: results from an e-health pilot study. *Journal of Cardiovascular Development and Disease*, *10*(8), 317.

Lepp, A., Barkley, J. E., & Karpinski, A. C. (2015). The relationship between cell phone use and academic performance in a sample of US college students. *Sage Open*, *5*(1), 2158244015573169.

Levenson, J. C., Shensa, A., Sidani, J. E., Colditz, J. B., & Primack, B. A. (2017). Social media use before bed and sleep disturbance among young adults in the United States: a nationally representative study. *Sleep*, *40*(9).

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE publications, Inc.

Liu, H., Ji, Y., & Dust, S. B. (2020). “Fully recharged” evenings? The effect of evening cyber leisure on next-day vitality and performance through sleep quantity and quality, bedtime procrastination, and psychological detachment, and the moderating role of mindfulness. *Journal of Applied Psychology*.

Liu, S., Xiao, T., Yang, L., & Loprinzi, P. D. (2019). Exercise as an alternative approach for treating smartphone addiction: a systematic review and meta-analysis of random controlled trials. *International Journal of Environmental Research and Public Health, 16*(20), 3912.

Loewenstein, G., Price, J., & Volpp, K. (2016). Habit formation in children: Evidence from incentives for healthy eating. *Journal of health economics*, 45, 47-54.

Long, J. A. (2019). interactions: Comprehensive, user-friendly toolkit for probing interactions. *R package version, 1(0)*, 33.

Lowe-Calverley, E., & Pontes, H. M. (2020). Challenging the concept of smartphone addiction: An empirical pilot study of smartphone usage patterns and psychological well-being. Cyberpsychology, Behavior, and Social Networking, 23(8), 550-556.

MacKay, J. (2019). Screen time stats 2019: Here's how much you use your phone during the workday. *RescueTime: Blog*. Retrieved from: https://blog.rescuetime.com/screen-time-stats-2018/

Malinauskas, R., & Malinauskiene, V. (2019). A meta-analysis of psychological interventions for Internet/smartphone addiction among adolescents. *Journal of behavioral addictions, 8*(4), 613-624.

Marlatt, G. A., & George, W. H. (1984). Relapse prevention: Introduction and overview of the model. *British Journal of Addiction, 79*(4), 261-273.

McNamee, P., Mendolia, S., & Yerokhin, O. (2021). Social media use and emotional and behavioural outcomes in adolescence: Evidence from British longitudinal data. *Economics & Human Biology*, *41*, 100992.

McShane, B. B., Böckenholt, U., & Hansen, K. T. (2016). Adjusting for publication bias in meta-analysis: An evaluation of selection methods and some cautionary notes. *Perspectives on Psychological Science*, *11*(5), 730-749.

Mead, N. L., & Patrick, V. M. (2016). The taming of desire: Unspecific postponement reduces desire for and consumption of postponed temptations. *Journal of personality and social psychology*, *110*(1), 20.

Melumad, S., & Pham, M. T. (2020). The smartphone as a pacifying technology. *Journal of Consumer Research, 47*(2), 237-255.

Méndez-Aguado, C., Aguilar-Parra, J. M., Álvarez, J. F., Trigueros, R., & Fernández-Archilla, J. A. (2020). The influence of emotions, motivation and habits in the academic performance of primary education students in French as a foreign language. *Sustainability, 12*(6), 2531.

Meta (2022). *Meta Investor Relations FAQs*. https://investor.fb.com/resources/default.aspx

Milkman, K. L., Minson, J. A., & Volpp, K. G. (2014). Holding the hunger games hostage at the gym: An evaluation of temptation bundling. *Management science*, 60(2), 283-299.

Mosquera, R., Odunowo, M., McNamara, T., Guo, X., & Petrie, R. (2020). The economic effects of Facebook. *Experimental Economics, 23*(2), 575–602. https://doi.org/10.1007/s10683-019-09625-y

Nieman, D. C., & Wentz, L. M. (2019). The compelling link between physical activity and the body's defense system. *Journal of sport and health science, 8(3)*, 201-217.

Ochs, C., & Sauer, J. (2022). Curtailing smartphone use: A field experiment evaluating two interventions. *Behaviour & Information Technology, 41*(16), 3598–3616. https://doi.org/10.1080/0144929X.2021.2007284

Olson, J. A., Sandra, D. A., Colucci, É. S., Al Bikaii, A., Chmoulevitch, D., Nahas, J., ... & Veissière, S. P. (2022). Smartphone addiction is increasing across the world: A meta-analysis of 24 countries. *Computers in Human Behavior*, *129*, 107138.

Orben, A., & Przybylski, A. K. (2019). The association between adolescent well-being and digital technology use. *Nature human behaviour, 3(2), 173-182.*

Oulasvirta, A., Rattenbury, T., Ma, L., & Raita, E. (2012). Habits make smartphone use more pervasive. *Personal and Ubiquitous computing, 16*(1), 105-114.

Panova, T., & Carbonell, X. (2018). Is smartphone addiction really an addiction?. *Journal of behavioral addictions*, *7*(2), 252-259.

Parry, D. A., Davidson, B. I., Sewall, C. J., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour, 5*(11), 1535-1547.

Pascoe, M., Bailey, A. P., Craike, M., Carter, T., Patten, R., Stepto, N., & Parker, A. (2020). Physical activity and exercise in youth mental health promotion: a scoping review. *BMJ open sport & exercise medicine, 6*(1), e000677.

Pearson, N., Biddle, S. J. H., Griffiths, P., Sherar, L. B., McGeorge, S., & Haycraft, E. (2020). Reducing screen-time and unhealthy snacking in 9–11 year old children: The Kids FIRST pilot randomised controlled trial. *BMC Public Health, 20*(1), 122. https://doi.org/10.1186/s12889-020-8232-9

Penrod, J. D., Goldstein, N. E., & Deb, P. (2009). When and how to use instrumental variables in palliative care research. *Journal of palliative medicine, 12*(5), 471-474.

Pollak, R. A. (1970). Habit formation and dynamic demand functions. *Journal of political Economy*, *78*(4, Part 1), 745-763.

Prapavessis, H., Cameron, L., Baldi, J. C., Robinson, S., Borrie, K., Harper, T., & Grove, J. R. (2007). The effects of exercise and nicotine replacement therapy on smoking rates in women. *Addictive behaviors*, 32(7), 1416-1432.

Prapavessis, H., De Jesus, S., Fitzgeorge, L., Faulkner, G., Maddison, R., & Batten, S. (2016). Exercise to enhance smoking cessation: the getting physical on cigarette randomized control trial. *Annals of Behavioral Medicine*, 50(3), 358-369.

Qudah, M. F. A., Albursan, I. S., Bakhiet, S. F. A., Hassan, E. M. A. H., Alfnan, A. A., Aljomaa, S. S., & AL-khadher, M. M. A. (2019). Smartphone addiction and its relationship with cyberbullying among university students. *International Journal of Mental Health and Addiction*, *17*, 628-643.

Radtke, T., Apel, T., Schenkel, K., Keller, J., & von Lindern, E. (2022). Digital detox: An effective solution in the smartphone era? A systematic literature review. *Mobile Media & Communication, 10*(2), 190-215.

Rajagopal, P., & Rha, J. Y. (2009). The mental accounting of time. *Journal of Economic Psychology*, *30*(5), 772-781.

Ravindran, T., Yeow Kuan, A. C., & Hoe Lian, D. G. (2014). Antecedents and effects of social network fatigue. *Journal of the Association for Information Science and Technology, 65*(11), 2306-2320.

Rebar, A. L., Gardner, B., & Verplanken, B. (2020). Habit in exercise behavior. In G. Tenenbaum, & R. C. Eklund (Eds.), *Handbook of sport psychology* (Vol. II, 4th ed., pp. 986-998). Hoboken, NJ: John Wiley & Sons.

Reed, P., Fowkes, T., & Khela, M. (2023). Reduction in Social Media Usage Produces Improvements in Physical Health and Wellbeing: An RCT. *Journal of Technology in Behavioral Science*, 1-8.

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

Rodríguez-Gutiérrez, E., Torres-Costoso, A., del Pozo-Cruz, B., de Arenas-Arroyo, S. N., Pascual-Morena, C., Bizzozero-Peroni, B., & Martínez-Vizcaíno, V. (2024). Daily steps and all-cause mortality: An umbrella review and meta-analysis. *Preventive Medicine*, 108047.

Rosenquist, J. N., Morton, F. M. S., & Weinstein, S. N. (2021). Addictive technology and its implications for antitrust enforcement. *NCL Rev.*, *100*, 431.

Ruscio, J. (2008). A probability-based measure of effect size: robustness to base rates and other factors. *Psychological methods*, *13*(1), 19.

Ryan, R., & Hill, S. Cochrane Consumers and Communication Group: meta-analysis. *CCCG* http://cccrg.cochrane.org/author-resources. La Trobe University, Melbourne.

Saboo, A. R., Kumar, V., & Ramani, G. (2016). Evaluating the impact of social media activities on human brand sales. *International Journal of Research in Marketing*, *33*(3), 524-541.

Schnauber-Stockmann, A., & Naab, T. K. (2019). The process of forming a mobile media habit: Results of a longitudinal study in a real-world setting. *Media Psychology*, *22*(5), 714-742.

Sheng, N., Yang, C., Han, L., & Jou, M. (2023). Too much overload and concerns: Antecedents of social media fatigue and the mediating role of emotional exhaustion. *Computers in Human Behavior*, *139*, 107500.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). p-curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science*, *9*(6), 666-681.

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2015). Better *P*-curves: Making *P*-curve analysis more robust to errors, fraud, and ambitious *P*-hacking, a Reply to Ulrich and Miller (2015). *Journal of Experimental Psychology: General, 144*(6), 1146–1152.

Sitzia, S., Zheng, J., & Zizzo, D. J. (2012). Complexity and smart nudges with inattentive consumers. *Available at SSRN 2180956*.

SMART Act, S. 2314, 116th Cong. (2019). Retrieved from https://www.govtrack.us/congress/bills/116/s2314

Somasundaram, J., Koch, I., & Lim, N. (2023). Raising the AC temperature in the tropics, one degree at a time. *Energy Economics, 128*, 107191

Somasundaram, J., Zimmermann, L., & Pham, Q. D. (2024). Actual and Anticipated Incentives for Reducing Mobile Usage. Available at SSRN: https://ssrn.com/abstract=3719413.

Spiller, S. A., Fitzsimons, G. J., Lynch Jr, J. G., & McClelland, G. H. (2013). Spotlights, floodlights, and the magic number zero: Simple effects tests in moderated regression. *Journal of marketing research, 50(2)*, 277-288.

Stanley, T. B., Correia, C. J., & Irons, J. G. (2022). Contingency management for smartphone and social media use: a feasibility study. *Addiction Research & Theory*, *30*(5), 323-329.

Stead, L. F., Perera, R., Bullen, C., Mant, D., Hartmann-Boyce, J., Cahill, K., & Lancaster, T. (2012). Nicotine replacement therapy for smoking cessation. *Cochrane database of systematic reviews*, (11).

Sunstein, C. R. (2014). Nudging: a very short guide. *Journal of Consumer Policy, 37*, 583-588.

Sunstein, C. R. (2017). Nudges that fail. *Behavioural public policy*, *1*(1), 4-25.

Thai, H., Davis, C. G., Stewart, N., Gunnell, K. E., & Goldfield, G. S. (2021). The Effects of Reducing Social Media Use on Body Esteem Among Transitional-Aged Youth. *Journal of Social and Clinical Psychology*, *40*(6), 481–507. https://doi.org/10.1521/jscp.2021.40.6.481

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing science*, *4*(3), 199-214.

Throuvala, M. A., Griffiths, M. D., Rennoldson, M., & Kuss, D. J. (2020). Mind over Matter: Testing the Efficacy of an Online Randomized Controlled Trial to Reduce Distraction from Smartphone Use. *International Journal of Environmental Research and Public Health, 17*(13), 4842. https://doi.org/10.3390/ijerph17134842

Toh, S. H., Howie, E. K., Coenen, P., & Straker, L. M. (2019). “From the moment I wake up I will use it… every day, very hour”: a qualitative study on the patterns of adolescents’ mobile touch screen device use from adolescent and parent perspectives. *BMC pediatrics, 19*(1), 1-16.

Tromholt, M. (2016). The Facebook experiment: Quitting Facebook leads to higher levels of well-being. *Cyberpsychology, behavior, and social networking*, 19(11), 661-666.

Tsai, C. C. (2016). The Role of Duolingo in Foreign Language Learners’ Autonomous Learning. In *Proceedings of the 2016 Asian Conference on Language Learning* (pp. 195-211).

Tudor-Locke, C., Craig, C. L., Brown, W. J., Clemes, S. A., De Cocker, K., Giles-Corti, B., ... & Blair, S. N. (2011). How many steps/day are enough? For adults. *International Journal of Behavioral Nutrition and Physical Activity*, 8(1), 1-17.

Turel, O., & Bechara, A. (2021). A Triple-System Neural Model of Maladaptive Consumption. *Journal of the Association for Consumer Research*, 6(3), 000-000.

Twenge, J. M. (2019). More time on technology, less happiness? Associations between digital-media use and psychological well-being. *Current Directions in Psychological Science, 28*(4), 372-379.

Twenge, J. M., & Campbell, W. K. (2019). Media use is linked to lower psychological well-being: evidence from three datasets. *Psychiatric Quarterly*, *90*(2), 311-331.

Twenge, J. M., & Martin, G. N. (2020). Gender differences in associations between digital media use and psychological well-being: Evidence from three large datasets. *Journal of adolescence*, *79*, 91-102.

Twenge, J. M., Martin, G. N., & Campbell, W. K. (2018). Decreases in psychological well-being among American adolescents after 2012 and links to screen time during the rise of smartphone technology. *Emotion*, *18*(6), 765.

United States Senate Committee on the Judiciary. (n.d.). *Protecting children online*. https://www.judiciary.senate.gov/protecting-children-online

Ussher, M. H., Taylor, A. H., & Faulkner, G. E. (2014). Exercise interventions for smoking cessation. *Cochrane Database of Systematic Reviews*, (8).

Vaghefi, I., Lapointe, L., & Boudreau-Pinsonneault, C. (2017). A typology of user liability to IT addiction. *Information Systems Journal*, *27*(2), 125-169.

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics*, *35*(2), 215-247.

Valkenburg, P. M. (2022). Social media use and well-being: What we know and what we need to know. *Current opinion in psychology, 45*, 101294.

van Aert, R. C., Wicherts, J. M., & van Assen, M. A. (2016). Conducting meta-analyses based on p values: Reservations and recommendations for applying p-uniform and p-curve. *Perspectives on Psychological Science*, *11*(5), 713-729.

Van der Bij, A. K., Laurant, M. G., & Wensing, M. (2002). Effectiveness of physical activity interventions for older adults: a review. *American journal of preventive medicine*, *22*(2), 120-133.

Van Deursen, A. J., Bolle, C. L., Hegner, S. M., & Kommers, P. A. (2015). Modeling habitual and addictive smartphone behavior: The role of smartphone usage types, emotional intelligence, social stress, self-regulation, age, and gender. *Computers in human behavior*, *45*, 411-420.

Van Rensburg, K. J., Taylor, A., & Hodgson, T. (2009). The effects of acute exercise on attentional bias towards smoking-related stimuli during temporary abstinence from smoking. *Addiction*, 104(11), 1910-1917.

Vanman, E. J., Baker, R., & Tobin, S. J. (2018). The burden of online friends: The effects of giving up Facebook on stress and well-being. *The Journal of social psychology*, 158(4), 496-508.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of statistical software*, *36*(3), 1-48.

Volkmer, S. A., & Lermer, E. (2019). Unhappy and addicted to your phone?–Higher mobile phone use is associated with lower well-being. *Computers in Human Behavior*, *93*, 210-218.

Volpp, K. G., John, L. K., Troxel, A. B., Norton, L., Fassbender, J., & Loewenstein, G. (2008). Financial incentive–based approaches for weight loss: a randomized trial. *JAMA*, 300(22), 2631-2637.

Volpp, K. G., Troxel, A. B., Pauly, M. V., Glick, H. A., Puig, A., Asch, D. A., . . . DeGuzman, J. (2009). A randomized, controlled trial of financial incentives for smoking cessation. *N Engl J Med*, 360, 699-709.

Vuorre, M., Orben, A., & Przybylski, A. K. (2021). There is no evidence that associations between adolescents' digital technology engagement and mental health problems have increased. Clinical Psychological Science, 9(5), 823-835.

Walsh, S. P., White, K. M., & Young, R. M. (2008). Over-connected? A qualitative exploration of the relationship between Australian youth and their mobile phones. *Journal of adolescence, 31*(1), 77-92.

Wang, C., Lee, M. K., & Hua, Z. (2015). A theory of social media dependence: Evidence from microblog users. *Decision support systems, 69*, 40-49.

Ward, A. F., Duke, K., Gneezy, A., & Bos, M. W. (2017). Brain drain: The mere presence of one's own smartphone reduces available cognitive capacity. *Journal of the Association for Consumer Research, 2(2)*, 140-154.

Wei, X. Y., Liang, H. Y., Gao, T., Gao, L. F., Zhang, G. H., Chu, X. Y., ... & Lei, L. (2023). Preference for Smartphone-Based Internet Applications and Smartphone Addiction Among Young Adult Addicts: Gender Difference in Psychological Network. *Social Science Computer Review*, 08944393231222680.

White, K., Habib, R., & Hardisty, D. J. (2019). How to SHIFT consumer behaviors to be more sustainable: A literature review and guiding framework. *Journal of Marketing, 83*(3), 22-49.

Wilson, D. B. (2023). *Practical meta-analysis effect size calculator (Version 2023.11.27)*. https://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-Home.php

Witkietvitz, K., & Kirouac, M. (2016). Relapse prevention. In C. M. Nezu & A. M. Nezu (Eds.), *The Oxford handbook of cognitive and behavioral therapies* (pp. 215–228). Oxford University Press.

Wood, W., & Neal, D. T. (2009). The habitual consumer. *Journal of Consumer Psychology, 19(4), 579-592.*

Wood, W., & Neal, D. T. (2016). Healthy through habit: Interventions for initiating & maintaining health behavior change. *Behavioral Science & Policy*, *2*(1), 71-83.

Wood, W., & Rünger, D. (2016). Psychology of habit. *Annual review of psychology*, *67*, 289-314.

Xiao, T., Jiao, C., Yao, J., Yang, L., Zhang, Y., Liu, S., Grabovac, I., Yu, Q., Kong, Z., Yu, J. J., & Zhang, J. (2021). Effects of Basketball and Baduanjin Exercise Interventions on Problematic Smartphone Use and Mental Health among College Students: A Randomized Controlled Trial. *Evidence-Based Complementary and Alternative Medicine, 2021*, 1–12. https://doi.org/10.1155/2021/8880716

Young, H. (2014, November 7). Do young people care about learning foreign languages?. *The Guardian*. https://www.theguardian.com/education/2014/nov/07/-sp-do-young-people-care-about-learning-foreign-languages-data

YouTube (n.d.). *About YouTube*. https://about.youtube/

Zhu, Z., Chen, H., Ma, J., He, Y., Chen, J., & Sun, J. (2020). Exploring the relationship between walking and emotional health in China. *International journal of environmental research and public health*, *17*(23), 8804.

Zimmermann, L., & Chakravarti, A. (2022). Not just for your health alone: Regular exercisers' decision-making in unrelated domains. *Journal of Experimental Psychology: Applied*, *28*(2), 379–398.

Zimmermann, L., & Sobolev, M. (2022). Digital Strategies for Screen Time Reduction: A Randomized Field Experiment. *Cyberpsychology, Behavior, and Social Networking*. https://doi.org/10.1089/cyber.2022.0027