UNIVERSITÀ DI PAVIA

DEPARTMENT OF ECONOMICS AND MANAGEMENT

MASTER PROGRAMME IN INTERNATIONAL BUSINESS AND ENTREPRENEURSHIP

FROM ONE CLICK TO HABIT: EXPLORING THE EFFECTIVENESS OF DIGITAL NUDGES ON REPEATED BEHAVIOR CHANGE

DAL CLIC ALL'ABITUDINE: UN'ANALISI SULL'EFFICACIA DEI DIGITAL NUDGE NEL PROMUOVERE IL CAMBIAMENTO COMPORTAMENTALE RIPETUTO

Supervisor:
Prof. Marcin Łukasz Bartosiak

Student:
Roksana Zerafatangizelangroudi

Academic Year 2025-2026

Abstract

Digital nudging has become a popular mechanism in shaping user behavior across digital environments, yet its long-term effectiveness in sustaining repeated actions remains underexplored. This study investigates how consistent exposure to digital nudges influences continued engagement over time, using Duolingo as the research context. Drawing on behavioral economics and habit-formation theory, the study examines whether users with longer engagement histories and stronger behavioral consistency (measured through streak length) are more likely to sustain app use in the absence of reminders. A survey was conducted with Duolingo users who had maintained a learning streak of at least 365 days. After filtering invalid responses, 146 valid cases were analyzed using logistic regression to test the relationship between streak length and three outcomes: continued use without reminders, app opening behavior without reminders, and self-reported continuation for streak maintenance. The results showed a small but statistically significant relationship between longer streaks and continued use without reminders, suggesting that consistent engagement may strengthen autonomous habits. The other models showed similar but non-significant trends. These findings highlight that digital nudges can facilitate behavior internalization when reinforced consistently, but their influence depends on consistency and user motivation. The study contributes to the literature by emphasizing the distinction between short-term compliance and long-term habit formation.

Sommario

Il *digital nudging* è diventato uno strumento sempre più diffuso per orientare il comportamento degli utenti nei contesti digitali; tuttavia, la sua efficacia nel mantenere comportamenti ripetitivi nel lungo periodo rimane poco approfondita. Questo studio analizza in che modo un'esposizione costante ai *digital nudge* influenzi la continuità del coinvolgimento nel tempo, utilizzando Duolingo come contesto di ricerca. Basandosi sui principi dell'economia comportamentale e della teoria della formazione delle abitudini, la ricerca esplora se gli utenti con una storia di utilizzo più lunga e una maggiore costanza comportamentale (misurata attraverso la lunghezza della streak) siano più propensi a mantenere l'uso dell'app anche in assenza di promemoria. È stato somministrato un questionario a utenti Duolingo che avevano mantenuto una streak di apprendimento di almeno 365 giorni. Dopo la rimozione delle risposte non valide, sono stati analizzati 146 casi attraverso regressione logistica, per testare la relazione tra la lunghezza della streak e tre esiti comportamentali: l'uso continuato dell'app senza promemoria, l'apertura dell'app senza promemoria e la dichiarata intenzione di proseguire per mantenere la streak. I risultati hanno mostrato una relazione modesta ma statisticamente significativa tra streak più lunghe e uso continuato senza promemoria, suggerendo che un coinvolgimento costante può rafforzare la formazione di abitudini autonome. Gli altri modelli hanno evidenziato tendenze simili ma non significative. Nel complesso, i risultati indicano che i *digital nudge* possono facilitare l'interiorizzazione dei comportamenti quando vengono applicati con coerenza, ma la loro efficacia dipende dalla costanza e dalla motivazione dell'utente. Lo studio contribuisce alla letteratura distinguendo chiaramente tra *compliance* a breve termine e formazione di abitudini a lungo termine.

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Professor Bartosiak, for his exceptional guidance, patience, and expertise throughout this research journey. The completion of this thesis would not have been possible without his dedicated supervision and mentorship.

I also wish to thank my parents, Dr. Madjid Zerafat and Dr. Afsaneh Meishani, whose love and support have shaped every step of my academic and personal path. This work is dedicated in loving memory of my father, whose research will continue to contribute to advances in mathematics, and with profound appreciation for the unconditional support of my mother, who inspires me every single day.

List of Contents

List of Figures

List of Tables

Introduction

1.1 Background and Context

Over the past decade, digital environments have become powerful spaces for shaping human behavior. Platforms that once aimed only to inform or entertain now actively guide users toward specific actions, from exercising and learning languages to saving money or conserving energy. Within this context, digital nudging has emerged as an influential concept. It refers to the deliberate design of digital interfaces that steer decisions and actions without restricting individual choice (Weinmann et al., 2016). Rooted in behavioral economics and choice architecture, digital nudges use subtle cues such as reminders, progress bars, and feedback loops to influence behavior. Their strength lies in their simplicity, and it can be a small notification or a visible streak that prompts the user to take action in a way that still feels voluntary.

As technology becomes part of daily life, understanding how these nudges influence repeated behavior is increasingly relevant. Most existing research on digital nudging has concentrated on short-term behavioral outcomes or immediate decision effects (Weinmann et al., 2016; Mertens et al., 2022), and little attention has been given to whether these effects last over time or whether they lead to habit formation. The distinction between temporary compliance and long-term habit is important. A person might open an app after receiving a notification, but continuing to do so without any reminder indicates a deeper behavioral change. This shift from external prompting to self-driven action represents the point at which digital nudges may transform from simple behavioral triggers into tools for lasting behavioral change.

Duolingo, one of the most widely used language-learning platforms in the world, with over 116.7 million monthly active users (MAUs) (Duolingo, 2025), provides a meaningful context for exploring this process. Its design relies heavily on gamification, with features such as streaks, reminders, badges, and visual progress feedback that reflect core principles of digital nudging. These elements are intentionally designed to encourage daily engagement and create motivation to continue. The streak feature, in particular, rewards repetition and reinforces the value of consistency, while push notifications act as external cues that prompt users to return to the app. Together, these mechanisms illustrate how digital environments can convert motivation into routine.

Despite the prevalence of these design strategies, there is still limited understanding of how digital nudges influence long-term behavior once a routine is established. It is unclear whether users who have maintained consistent engagement continue to rely on these nudges or whether they can eventually act independently of them. Understanding this distinction matters for both theory and practice. From a theoretical perspective, it offers insight into how digital environments shape habit formation, and from a practical standpoint, it helps designers understand when nudges remain effective and when they become redundant.

By examining digital nudges in the context of sustained app use, this study aims to understand how subtle design cues contribute to repeated behavior over time. It focuses on when nudges continue to matter, when their influence declines, and how consistent engagement may reduce dependence on external reminders. In doing so, the research looks beyond single instances of behavior to explore the broader process of how digital interactions evolve into habits.

1.2 Research Problem

Multiple studies show that such nudges can succeed in prompting immediate actions (Grüning et al., 2023; Forberger et al., 2022; Karlan et al., 2016), but their long-term effectiveness remains underexplored. Most existing research is limited to short-term impacts, leaving open whether nudges can sustain behavior once users become familiar with them or when prompts are removed.

The gap between short-term behavioral compliance and long-term habit formation is central to this research problem and becomes especially important in contexts where continued engagement is critical to user success. In the case of language learning, platforms like Duolingo rely on repeated, consistent practice. Their architecture includes nudging features such as streaks and reminders to encourage users to return daily, but it remains unclear whether users continue to revisit simply because of ongoing nudges or whether their behavior gradually internalizes into routine. In effect, we need to know if nudges retain their influence or if their effect decays as habits emerge.

Answering this distinction has theoretical and practical implications. Theoretically, it touches on how behavior change mechanisms interact with habit formation. Practically, it may guide digital platform designers on temporal dynamics, in order to avoid user fatigue while supporting lasting engagement. Without clarity, designers risk overusing nudges, which could undermine their long-term effectiveness.

Due to the scarcity of digital nudge research on long-term effects and the dominance of short-term interventions, our understanding of how user motivation functions in the absence of a given nudge remains limited. To fill this gap, this study investigates digital nudges and their impact on repeated behavior within a real-world learning environment. Specifically, it explores whether consistent exposure to nudging mechanisms gradually shift users from externally prompted action to autonomous engagement.

1.3 Research Question and Hypotheses

This study investigates how digital nudges influence repeated user behavior over time, focusing on their effectiveness in fostering sustained engagement rather than short-term compliance.

The study is guided by the following research question:

RQ: How do consistent digital nudges change or impact repeat behaviors over a long period of time?

This question stem from the need to differentiate between temporary behavioral triggers and durable behavioral change. While many studies demonstrate that nudges can prompt immediate responses, far fewer explore whether such effects persist once external prompts are removed or reduced. Understanding this dynamic is essential to assessing the true behavioral impact of digital interventions.

Accordingly, the study sets out to examine both the direct effects of digital nudges on repeated engagement and the temporal dimension of those effects, assessing whether their influence strengthens, stabilizes, or declines as user experience increases. Using Duolingo as an illustrative context, the research investigates how consistent mechanisms such as reminders, streak tracking, and progress feedback affect users' ongoing app engagement.

From these objectives, the following hypotheses are proposed:

H_1: Digital nudges impact repeat behaviors after a long period of consistency

H_2: Over time, users require fewer nudges for the same behavior

Together, these hypotheses operationalize the central idea that sustained behavioral change depends not only on the presence of nudges but also on their consistency and the evolution of user autonomy. The analysis seeks to determine whether digital nudges act as temporary motivators or as catalysts for long-term behavioral adaptation.

1.4 Scope and Limitations

While this study offers valuable insights into how digital nudges influence repeated behaviors over time, several limitations should be acknowledged. The data were collected through self-reported surveys, which depend on participants' memory and honesty. This means some responses may not perfectly reflect actual app behavior. In addition, because the study design was cross-sectional, it captures a single point in time and cannot establish causal relationships between nudges and long-term habit formation.

Another limitation concerns the sample itself. Participants were mainly long-term Duolingo users with streaks exceeding 365 days, meaning they represent highly engaged "super-users." As a result, the findings may not fully generalize to newer or less active users.

These boundaries do not undermine the study's contribution; rather, they provide clarity regarding what can and cannot be inferred from the findings. The focus on a single, widely used learning platform enables a detailed understanding of digital nudges in a real-world behavioral setting, offering a foundation for future comparative and longitudinal research. A more detailed and elaborative discussion of methodological and contextual limitations is provided in section 6.4.

Literature Review

2.1 Overview of Digital Nudges

2.1.1 What Is Digital Nudging?

Digital nudging is the deliberate design of user interface elements to influence decision-making in digital environments without limiting choices (Meske & Potthoff, 2017). Nudging is the altering of choice architecture to guide decisions predictably without coercion or incentives (Thaler & Sunstein, 2008), and digital nudging extends this idea into online contexts. Weinmann et al. (2016) define it concisely: digital nudging is “the use of user-interface design elements to guide people’s behavior in digital choice environments”. Unlike offline nudges, such as changing food placement in a cafeteria, digital nudges occur within screens and workflows, often subconsciously shaping decisions (Weinmann et al., 2016). In essence, the layout, wording, defaults, or timing within interfaces can subtly steer users. Mirsch et al. (2017) further emphasize that digital nudges incorporate design, wording, visual elements, or small user interface tweaks to influence decisions in real time within digital platforms.

Digital nudging is becoming ubiquitous. Whether in e-commerce checkout flows, privacy setting defaults, or progress reminders in learning apps, interfaces shape how user decisions are made. Because digital environments are highly scalable, the cumulative effect of small nudges can be significant even when individual effects are minor (Weinmann et al., 2016). However, ethical concerns arise around transparency and manipulation. Nudges designed without user consent or visibility can erode trust or impact autonomy. Scholars argue that design ethics must accompany behavioral intentions, ensuring nudges remain aligned with user welfare (Meske & Amojo, 2019).

2.1.2 Why Are Digital Nudges Important?

Digital nudging offers a unique advantage in influencing behavior across large populations at minimal cost. Digital platforms mediate countless daily decisions, which highlights how integrating minor interface adjustments, such as defaults or prompts, can lead to measurable shifts in user actions (Weinmann et al., 2016). For instance, Square, a Silicon Valley-based mobile payment company, introduced a payment interface that nudges consumers toward tipping. After swiping their credit or debit card, users are prompted to select from preset tip amounts, enter a custom tip, or choose not to tip. This design change was reported to increase the proportion of transactions including a tip by 38 percent (Weissmann, 2014), demonstrating how thoughtfully designed digital interfaces can significantly influence user behavior.

Additionally, in comparison with traditional behavioral interventions, digital nudging is exceptionally cost-effective. Interface tweaks and message prompts can be implemented with low marginal costs and minimal maintenance, making them suitable for large-scale adoption across domains like healthcare, education, finance, and sustainability (Mirsch et al., 2017). In many cases, the infrastructure already exists, and nudges can be implemented with minimal issues.

Digital nudges are also valuable for encouraging behaviors that require repetition or sustained engagement. They reduce cognitive friction, issue reminders at the point of action, and simplify decision-making, thereby aiding in habit formation. Empirical studies in digital education platforms show that regular reminders significantly improve students' participation, attendance, and assignment completion (Kizilcec et al., 2020; Liu et al., 2023), and well-designed digital nudges can align individuals' decisions with better outcomes by gently leveraging cognitive biases without coercion (Weinmann et al., 2016; Meske & Amojo, 2020).

However, even when highly effective, digital nudging must be executed with ethical considerations at the forefront. Nudges should preserve autonomy of users, maintain their transparency, and avoid manipulative practices. It is important for users to understand when a nudge is occurring, why it is occurring, and what goal it serves (Meske & Amojo, 2020). Transparency and consent are key to building trust and support long-term engagement with digital platforms.

2.1.3 What Are the Techniques?

Various techniques are used in digital nudging to influence decisions without manipulating or coercing the user. One of the most straight-forward types of nudging is simplification, which involves reducing the complexity of information or steps, in turn making it more likely for users to follow through and complete the steps. Creating more streamlined flows and requiring fewer decisions decreases user resistance by making the process faster or easier. Similarly, defaults, which are pre-selected options that can be changed by the user, increase the likelihood of acceptance due to user inertia. While prompting changes subtly, defaults can be highly effective as a nudging method, especially in relation to instances with ticked boxes or cookie consent forms. Furthermore, timely reminders are prompts sent at moments when people are most likely to act and are especially effective for deadline-driven behaviors. However, while timing and simplicity play important roles in influencing user decision-making, people are bound to be influenced by the actions of others. That is where social norms come in as a nudging technique, as providing information about how a majority has acted can increase compliance. Likewise, how a message is presented affects user decisions, bringing us to the technique of framing. Framing refers to the way statements are worded to exert maximum impact on the reader and to subtly encourage a certain behavior or outcome. Loss-framing is deemed to be particularly motivating, with statements such as “don’t lose your streak” or “don’t miss out.” Messages can also be infused with emotional framing (Troussard et al., 2016).

Alternatively, digital nudging techniques can be classified based on the underlying behavioral mechanism they employ. One such type is social nudges, which leverage social influence to encourage specific actions by highlighting what others do. Closely related are reinforcement nudges, designed to promote repeated behavior through positive rewards or incentives. Disclosure nudges provide users with additional information to help inform their choices, improving decision quality without restricting options. On the other hand, friction nudges work by increasing the effort needed to perform undesirable behavior, thereby discouraging it. In addition, there are feedback nudges, which give individuals real-time insights into their behavior in order to promote self-awareness and adjustment. Warning nudges alert users of potential negative consequences of certain behaviors, promoting caution. Meanwhile, Scarcity nudges create a sense of urgency or limited availability, motivating quicker action. Lastly, deception nudges manipulate information to steer decisions, although their ethical implications remain controversial (Gyulai & Revesz, 2023).

2.2 Digital Nudging at Scale

One of the defining strengths of digital nudging is in its scalability, meaning that it can be delivered to thousands or even millions of users at once without requiring extra resources. Unlike traditional behavior change methods that often require face-to-face contact, printed materials, or in-person interventions, digital nudges can be embedded in apps, websites, or digital services and made instantly available to a vast number of users. This scalability makes digital nudging a powerful tool where large-scale behavior change is desired.

Empirical studies have examined the real-world impact of nudging at scale. In a large-scale randomized field experiment conducted by Hyytinen et al. (2022), over 40,000 Finnish citizens who had not adopted a new e-government digital ID service were targeted with different nudge messages. The researchers tested various methods, including reminders, social norm cues, and simplification. The most effective intervention was a loss-framed message emphasizing the risk of missing out on critical services, which led to a statistically significant increase in adoption compared to the control group. The intervention was fully automated and delivered through existing government communication channels, highlighting how small behavioral prompts can effectively be implemented at a very large scale within existing infrastructure.

Similarly, nudges have been embedded in consumer health technology, demonstrating effectiveness across large user bases. In a longitudinal observational study of Apple Watch users, Nazaret and Sapiro (2023) examined the effects of hourly "stand reminders". Automated haptic nudges were used to prompt users to stand up after prolonged sitting. The study tracked over 160,000 users and found that the nudges increased standing behavior by 43.9%, with this behavioral change persisting over 21 months. These findings illustrate that even simple nudges can generate results across massive user populations.

Algorithmic personalization further enhances the scalability of digital nudges. Chiam et al. (2024) introduced *NudgeRank,* a graph-based algorithm designed to personalize health nudges based on users' digital profiles and social network positioning. NudgeRank was deployed on a physical activity platform with tens of thousands of users and generated a 6-7% increase in exercise engagement compared to non-personalized messages. What made this approach scalable was its automated adaptability. As the system continuously refined itself based on user responses, it was able to deliver the right nudge to the right user at the right time without additional human input.

However, scalability also comes with risks. It is misleading to assume that a successful nudge in smaller contexts will maintain its success rate in larger populations. For example, a campaign in the United States sent text message nudges to over 800,000 high school seniors to encourage completion of the Free Application for Federal Student Aid (FAFSA) but found no statistically significant improvement in application rates or college enrollment (Bird et al., 2019). The campaign's failure was due to a mismatch between message content and the recipients' individual circumstances, reducing its effectiveness despite its extensive reach. This case highlights that scale requires sensitivity to heterogeneity rather than solely focusing on volume.

Meta-analytical evidence confirms that scaling nudges involves trade-offs. DellaVigna and Linos (2020) analyzed over 126 Randomized Controlled Trials (RCT) conducted by behavioral science teams, which were embedded within U.S. government agencies and covered over 23 million individuals. They found that while the size of impact tended to be modest at scale, with an average of 1.4 percentage points, it was still statistically meaningful and cost-effective. In particular, interventions that were relevant in terms of context, that tested well in pilot stages, and that were informed by behavioral theory were more likely to retain their effectiveness when scaled.

In summary, digital nudging at scale provides an unparalleled opportunity to influence behavior across populations. However, its success depends on factors such as context-sensitive design, personalization, and rigorous testing. Without these, there is a risk of ineffective large-scale implementations despite expansive reach.

2.3 Effect of Digital Nudges on Long Term and Repeat Behaviors

2.3.1 Empirical Evidence on Long-Term and Repeated Effects

One of the debates in nudge literature concerns whether effects persist beyond the immediate decision. Brandon et al. (2017, revised 2022) addressed this directly through 38 natural field experiments on consumption reduction in homes with the Home Energy Report (HER) treatment. The study illustrated that while nudges can initially change behavior, their effects often decay once the intervention is removed. However, they concluded that the influence of nudges show higher persistence via technology adoption and in the case of higher technology abundance. The present study complements

Brandon et al. (2017, revised 2022) by exploring the effects of digital nudges such as daily reminders on habit-formation in the long term, raising the question of if consistent digital nudges reduce decay.

Meta-analytic evidence paints a bigger picture. Mertens et al. (2022) drew on over 200 studies and 440 behavioral interventions, finding that choice architecture is a widely applicable and effective tool for behavior change, with considerable heterogeneity, but did not study long-term persistence or outcome of nudges. However, on the other hand, in a systematic review of digital behavior change interventions, Zhu et al. (2024) emphasized that the most commonly applied behavioral change techniques for habit formation include self-monitoring of behavior, self-decided goals, reward system, and consistent prompts. This aligns directly with this study's investigation into whether Duolingo's streak system and reminders function not only as short-term nudges but as habit-forming mechanisms in the long run.

A field-experiment collecting behavioral user data from 280 participants over the span of 6 weeks on the effectiveness of the self-nudging app *one sec* reported a 57% decrease in users' opening of target apps after six consecutive weeks (Grüning et al., 2023). The app is installed by the user to shrink the mindless use of designated smartphone apps. The interventions used by Grüning et al. (2023) include an explicit dismiss option to choose against opening the target app, friction by a short waiting time, and a deliberation message to help counter impulsive intentions, listed in order of effectiveness from most to least effective. Improvements were seen over the 6 weeks, as there were 37% less attempts to open the target apps than in the first week, highlighting the effectiveness of the nudges over time. The study, however, only examined the impact of nudges while they were being administered and over a short period of time, leaving a gap for research on long-term effects of digital nudges on habit-building, as opposed to habit-breaking.

Weijers et al. (2021) provide a conceptual framework for applying nudging within educational contexts. They highlight that understanding the long-term effects of a nudge is crucial for successful implementation, although their research is highly theory-based and not practice-based. They introduce a decision matrix grounded in the framework presented by Hansen and Jespersen (2013), distinguishing four nudge categories comprising Type 1 nudges, which target behaviors not conscious and deliberate, and Type 2 nudges, which target deliberate actions and choices, in addition to transparent and non-transparent nudges. Type 2 nudges were found to be more effective for attaining long-term and sustainable behavioral changes, providing a basis for further research on the impact of repeated, long-term digital interventions.

Moving from theory to practice within the context of education and digital learning, Vermetten et al. (2017) provide a long-term experimental study on nudging in higher education, testing how subtle interventions could support study success over an extended period. Conducted at the Dutch University of Applied Sciences, their experiment examined how digital nudges in the form of injunctive feedback incorporated into the university's app influenced students' persistence and learning outcomes in four months. There was no significant interrelation between the treatments and changes in GPA or ECT points. The paper underscores the need for longer-term assessment of digital nudge effectiveness in similar contexts, but does not address nudge removal and subsequent sustained effects.

Weijers et al. (2023) extended the discussion on educational nudges by conducting three field experiments aimed at encouraging autonomous learning behavior via nudging.

Their results showed that simple interventions can positively influence autonomous learning behaviors, though the strength of effects varied depending on the context and the type of nudge. In contrast, Baker et al. (2016) tested a scheduling nudge in a large Massive Open Online Course (MOOC) and found modest or even negative long-term effects on engagement and persistence. These surprising results highlighted the risk of undermining motivation if the nudges are perceived as intrusive or are mismatched with the learners' goals. Taken together, the studies demonstrate both the opportunities and pitfalls of digital nudges in education.

Moreover, Kay and Bostock (2023) provide empirical evidence of how digital nudges can drive sustained engagement at scale through their institution-wide evaluation of the University of Canterbury's *Early Alert System.* This initiative automatically delivered text and email nudges to students identified as at risk of disengagement. Results showed that students who received the nudge demonstrated a higher re-engagement rate and also spent additional time engaging with online material. These effects also spilled over into other enrolled courses where no nudge was provided, pointing to the transferability of the benefits. The study highlights that timely nudges can influence persistence and digital learning behaviors, providing critical support for this thesis' focus on the effects of digital nudges on repeat behaviors in long-term contexts.

In the context of health and well-being, a few key studies have explored the effects of nudges. Johannsdottir et al. (2025) examined how digital nudges can be used within digital health technologies in healthcare settings to support patient treatment, safety, and management. They found that digital nudges are effective in improving lifestyle-related diseases by promoting more exercise, weight loss, reducing unhealthy behaviors including smoking and excessive drinking, and having appreciable influence on adherence to treatment as well as improvements in mental health. However, Johannsdottir et al. did not look into the long-term impact of digital nudges on repeat behaviors, instead focusing on their effectiveness during the period of occurrence.

Similarly, Forberger et al. (2022) provide evidence on the application of nudges in increasing physical activity levels and reducing sedentary behavior through workplace health promotion. While positive effects were observed during the intervention period, the authors highlighted the need for further research on long-term effects and habituation, as well as the impact of digital nudges in isolation. Meanwhile, Venema et al. (2018) assessed the effect of a default nudge to decrease sedentary behavior at work over time, finding that stand-up working rates rose from 1.82% at baseline to 13.13% during the intervention. After the removal of the nudge, the percentage lingered at 10.01% after two weeks and 7.78% after two months. A multilevel analysis also showed a significant rise in both intention and social norms following the intervention. The present study directly extends these findings by expanding both the types of nudges tested and the timeline of interventions, specifically in the context of digital learning.

Beyond workplace contexts in health and wellness, Sasaki et al. (2021) conducted research on repeated nudges and their persistence during Covid-19. They found that a gain-framed altruistic message on protecting loved ones was highly effective for short-term behavior change, while its impact on perception change and stronger intentions showed some degree of durability. These findings have important implications for the current study, which extends this line of inquiry into digital learning. While Sasaki et al.'s objective leaned more toward habit-breaking than habituation, given that their research was set against the backdrop of a crisis that forced changes to daily routines, this thesis investigates whether consistency in digital nudges can foster sustained engagement in a non-crisis, everyday setting, and in an app environment.

In the context of environment and sustainability, Ling et al. (2023) investigated the impact of social norm nudges for household recycling, their direct and spillover effects, and the persistence of said effects in a longitudinal study. The nudge led to a rise in recycling participation and policy support during the five-month treatment period, demonstrating positive direct and spillover effects. However, the results decayed within the following five months after the removal of the messaging, indicating that the effects were not sustainable. Correspondingly, Linder et al. (2018) studied the effectiveness of an information intervention in promoting food waste recycling. The results showed a statistically significant increase in recycled food waste in comparison to a control group. Both studies conclude that nudges can improve target behavior, but neither study involved digital nudges or examined the role of high-frequency nudges.

In a different domain, Karlan et al. (2016) investigated the role of reminders in helping individuals meet their savings commitments through field experiments with three different banks. They found that reminder messages significantly increased the likelihood of clients making deposits into commitment savings accounts, particularly among those who had recently opened an account. This finding parallels Duolingo's daily notifications, which nudge users to continue their language learning streak. Both contexts involve voluntary, repeated behaviors that require ongoing effort over time, one for savings deposits and the other for daily practice. However, while Karlan et al. demonstrate that digital nudges are effective in driving commitment attainment, their study does not examine whether these effects persist once the nudges are removed and how the frequency or consistency of nudges alters sustained behavior.

Extending this line of evidence to another domain, Van Gestel et al. (2017) conducted a longitudinal study to examine the effect of a food repositioning nudge at checkout counters on healthy food choice. They found that making healthier items more visible and accessible increased their purchase during the four-week intervention period in comparison to the four-week baseline phase, promoting a growth of over 50%. Not unlike the other studies previously mentioned, exposure to nudges were found to be effective during the intervention period, and while the study focused on offline consumer behavior rather than digital nudging, it reinforces the idea that simple, repeated interventions may induce sustainable engagement, providing a foundation for exploring similar mechanisms in a digital environment.

2.3.2 Theoretical Foundations and Mechanisms Explaining Repeated Behavior

While empirical studies demonstrate that nudges can influence behavior in the short term, a theoretical lens is required for understanding why certain behaviors persist while others fade. Exploring the mechanisms that explain repeated behaviors and how they are formed, maintained, or abandoned is essential for the interpretation of the long-term effectiveness of digital nudges. This section looks beyond surface-level outcomes to uncover what causes sustained engagement over time.

Lally et al. (2010) provide seminal empirical investigations into how habits form over time in everyday contexts. Their longitudinal study tracked participants as they chose an eating, drinking, or activity behavior to perform daily in a consistent manner for 12 weeks. Habit strength was assessed through the Self-Report Habit Index (SRHI), focusing particularly on automaticity, which is the extent to which behaviors became effortless and automatic. The authors found that habit formation follows a nonlinear trajectory. Behavioral automaticity typically increases gradually at first, accelerates, and then levels off, eventually reaching asymptotic habit strength, where additional repetition adds little further strength. The time required to reach this asymptote varied among participants, ranging from 18 to 254 days, and consistency of behavior played a key role in the duration: more frequent repetition led to stronger habit formation. Importantly, occasional lapses, such as missing a single day, did not negatively affect the development of the habit, indicating a degree of resilience in the process. The implications of this model are profound for digital nudging. It emphasizes that sustained and repeated engagement in a behavior, supported by consistent cues, is critical to transitioning from deliberate actions to automatic, habitual ones. Therefore, digital nudges designed to promote repeated behavior need to be delivered over long periods of time and with consistency to facilitate habit formation. Furthermore, the study emphasizes the variability among individuals in the rate of habit creation, implying that one-size-fits-all interventions may not be equally effective for all users. The present research builds on Lally et al.'s findings by exploring how digital learning platforms can leverage repeated nudges to stimulate automatic engagement and sustained user behavior over time.

Lally et al. (2010) focused on how habits form over time, but this raises the question of which conditions need to be in place for habit-building behavior to occur. Fogg (2009) proposes a behavioral framework that highlights three fundamental factors required for any behavior to occur: motivation, ability, and triggers. According to the Fogg Behavior Model (FBM), these three components must occur together, and if one is missing or insufficient, the target behavior will not take place. Motivation refers to the individual's desire to perform the behavior, while ability represents how easy or difficult the behavior is to execute, factoring in elements like time, effort, and complexity. Finally, triggers act as cues that induce the behavior when motivation and ability align. What makes Fogg's model especially relevant for digital nudging is its emphasis on understanding underlying psychology for successful persuasive design and designing interventions that balance these three factors. Moreover, Fogg identifies three types of triggers: the *spark* for motivation is low, the *facilitator* for when motivation is high but ability is constrained, and the *signal,* which simply serves as a reminder and works when both ability and motivation are available and aligned to perform the target behavior, paralleling the reminder nudges on the Duolingo app. Following this framework, digital nudges can spark motivation through persuasive messaging, facilitate ability by simplifying tasks or reducing friction, and provide triggers via notifications or reminders. The FBM provides a useful blueprint for analyzing how Duolingo's digital nudges are designed to guide users toward consistent engagement.

While Fogg's model emphasizes the micro-mechanics of behavior, broader frameworks such as the Behavior Change Wheel extend the concept further. Michie et al. (2011) introduced the Behaviour Change Wheel (BCW) as a framework to both understand behavior and design interventions to change it. At its core lies the COM-B model, which states that behavior occurs when capability, opportunity, and motivation come together. Capability includes the skills and knowledge needed to act, opportunity relates to the external environment that makes the behavior possible, and motivation covers both conscious processes, such as planning and decision-making, and automatic processes, such as habits and emotions. The BCW was constructed to overcome the limitations of existing frameworks, and for digital nudging research, the BCW underscores that long-term engagement is not only a matter of triggering motivation but also of shaping capability and opportunity within the user's environment, offering a more holistic counterpart to Fogg's model.

While the models discussed shed light on behavior, existing research highlights unresolved questions in nudging. Marchiori, Adriaanse, and De Ridder (2017) emphasize that much of the literature has overlooked important psychological factors, calling for a closer examination of what truly drives behavior. This observation reveals that we still lack a clear understanding of not just whether nudges are effective, but why they work and under what conditions they lead to lasting engagement rather than mere short-term compliance. From a behavioral science perspective, Van Bavel et al. (2020) expanded on behavioral change theory by showing how social and collective dynamics shape responses, particularly in times of crisis such as the COVID-19 pandemic. They demonstrate that compliance is not an individual decision but rather is a part of wider social and cultural contexts, which can lead to stronger or weaker adherence over time. They argue that effective interventions must look at collective responsibility and shared action. This perspective is relevant to digital nudging where interventions that tap into social norms may prove more effective in sustaining engagement and supporting long-term habit formation.

2.3.3 Moderators, Challenges, and Dynamics in Sustaining Nudges

Sustaining the effects of nudges over time is not straightforward. While numerous studies show that nudges can successfully trigger immediate behavioral responses, their long-term impact remains unclear. The effectiveness of nudges is shaped by a range of moderators, such as individual differences, contextual variances, and frequency of exposure, as well as by challenges including nudge fatigue and unstable motivation. It’s crucial to look into how the influence of nudges changes over time.

A promising start in assessing the extended influence of nudges is the study conducted by Van Rookhuijzen et al. (2021), who examined the temporal spillover effects of a default nudge. Two of the three experiments demonstrated spillover effects, while one did not, suggesting that default nudges may cause short-term spillover effects for certain behaviors. While Van Rookhuijzen et al. provide a useful foundation for the current study, the observation period was only two days and they focused solely on default nudges, highlighting the need for further research on the longer-term dynamics of nudging and on a wider range of nudge types.

Kovacs et al. (2021) highlight another challenge in sustaining behavioral change: the tendency for users to gradually weaken their adherence to a regimen while expecting to recommit in the near future. Their findings suggest a temporal dynamic in which individuals anticipate re-engagement but often postpone doing so. This leads to a cycle of delay rather than sustained progress. This dynamic highlights a challenge for nudging strategies. Even when initial compliance is achieved, users' actual behaviors may differ over time. This points to the importance of designing interventions that not only prompt immediate action but also counteract resistances in habit formation.

Saig and Rosenfeld (2023) add to this discussion by emphasizing the importance of timing in sustaining user engagement. Their study shows that interventions do not always need to push for more activity; instead, strategically suggesting breaks can prevent fatigue and ultimately extend participation. This challenges the assumption that continuous nudging is always beneficial and highlights the risk of overexposure. For a platform like Duolingo, where daily practice is encouraged, this indicates that striking a balance between consistency and recovery could supports more durable language learning habits.

2.4 Gaps in Literature

Despite growing research on digital nudging across domains such as health, sustainability, finance, and education, several important gaps remain. First, much of the existing literature concentrates on short-term responses to nudges or immediate changes in behavior during the intervention period. Even when studies include longitudinal elements, the outcomes are typically measured only while nudges are still active, leaving uncertainty about whether the behavior continues once prompts are removed. Additionally, while some studies show persistence under specific circumstances, there is no clear consensus on the conditions under which digital nudges translate into stable, repeated behaviors. Findings often differ depending on the context, the type of nudge used, and the nature of the target behavior, highlighting important inconsistencies across results.

Furthermore, current evidence rarely examines high-frequency nudges that operate daily over long periods of time. Most interventions involve limited exposure windows and do not capture the full duration typically required for habit formation. Similarly, there is limited understanding of what happens once a nudge-supported routine becomes more automatic, and whether continued nudging helps or becomes unnecessary. Finally, many studies rely on institutional or platform-level interventions but pay little attention to real-world, voluntary engagement behaviors where users must repeatedly decide to return, such as daily learning on Duolingo. This represents a practical and theoretical gap, as repeated engagement requires not only triggering action but supporting motivation, self-regulation, and habit stabilization.

Together, these gaps show that we still lack a clear understanding of how and when digital nudges transition from driving short-term compliance to supporting long-term behavioral consistency. This gap directly motivates the present study, which examines whether consistent exposure to digital nudges is associated with sustained engagement even in the absence of reminders. By focusing on long-term Duolingo users, this research extends prior work on the role of digital nudges in supporting ongoing habit formation in the form of self-driven repeat behaviors shaped by consistent digital nudges over a long period of time.

Contextual Background: Duolingo and Digital Nudges

3.1 Overview of Duolingo and Its Nudging Strategies

Duolingo is an available-to-all language learning platform providing quick, bite-sized lessons with a creative twist. Founded in 2010, Duolingo has been using research-backed teaching methods and engaging content to create courses that target all four areas of language learning: speaking, reading, listening, and writing. The app focuses on habit-building through gamification, challenges, and consistent reminders. The company places profound importance on the concept of universal accessibility, stating that Duolingo is used by the richest man in the world and numerous Hollywood stars, while also being used by public schools students in developing countries, portraying their belief that equality is when spending more can't buy you a better education (Duolingo, n.d.).

Duolingo stands as one of the world's most popular language learning platforms, boasting over 500 million registered users and more than 40 million active monthly learners as of 2024 (Duolingo, 2024). Its widespread adoption is a result of the platform's ability to engage diverse audiences across more than 40 languages (Duolingo, 2024). Users complete millions of lessons daily, generating interaction data that informs allows the app to employ ongoing feature optimization (Duolingo, 2024). Importantly, Duolingo's success in retaining users and encouraging regular practice is often attributed to its effective use of digital nudges within the app. These nudges, including streak reminders, leaderboards, and personalized goals, are designed to sustain motivation and promote habitual learning behaviors (Duolingo, n.d.). The platform's scale and high user engagement provide a valuable context for studying the effectiveness of digital nudging in shaping repeated behaviors.

3.2 Types of Digital Nudges Used by Duolingo

Duolingo integrates a variety of digital nudges directly into its user experience to encourage consistent learning and long-term engagement. Understanding these nudging techniques is crucial for analyzing how digital interventions can influence repeated behaviors effectively.

One of the most prominent features is the use of streak reminders, which notify users to maintain consecutive days of learning. These reminders tap into the behavioral principle of loss aversion, motivating users not to break the chain and lose their progress streak. This type of nudge falls under timely reminders and framing nudges because it prompts action at critical moments and uses loss framing to increase motivation. Alongside this, the platform incorporates leaderboard rankings, which rank users weekly based on experience points earned. This element introduces social comparison as a motivating force, encouraging users to compete with friends and other learners. Leaderboard rankings are classic examples of social nudges, leveraging social norms to increase engagement.

In addition, Duolingo offers XP boosts and rewards, providing immediate positive reinforcement to users who reach specific milestones or complete challenges. These features represent reinforcement nudges that encourage repeated behavior through rewards. Users can also set daily goals, which serve as default nudges by offering personalized targets that guide users toward regular practice. Daily goals also function as feedback nudges by allowing users to track their progress. Finally, the platform's social features, such as friend lists and challenges, foster a sense of community and accountability. These social nudges strengthen motivation by providing peer influence and encouragement. By combining these techniques, Duolingo creates a comprehensive behavioral ecosystem designed to promote sustained learning and long-term user retention.

3.3 User Engagement Metrics: How Duolingo Measures Success

Duolingo frames success as a combination of broad platform indicators and smaller learning and engagement signals. The company reports usage numbers such as Daily Active Users (DAUs) and subscriber counts as the primary measures of platform reach (Duolingo, 2025). Internally, Duolingo prioritizes a learning-centered metric called Time Spent Learning Well (TSLW) that looks at sessions via their expected learning value rather than the raw time spent on the app (Duolingo, 2024). Meanwhile, at product level, Duolingo tracks behavioral engagement metrics including streak length, XP earned, lessons completed, and frequency of sessions to monitor habit formation inside the app (Duolingo, n.d.). Gamified social measures such as leaderboard position and league movement are captured separately to assess competitive and social engagement effects (Duolingo, n.d.). The company maintains an experiment-first culture and routinely uses A/B testing to optimize retention and learning outcomes (Duolingo, 2023). The experiments are combined with a Markov growth model that decomposes top-line metrics, such as the DAUs into user segments based on activity states (Duolingo, 2023). Academic work on online education experiments reinforces that trials embedded within platforms and small-grained engagement metrics are the right way to infer whether nudges causally improve learning and persistence (Kizilcec et al., 2020). Overall, Duolingo's measurement stack is designed to align product optimizations to both business goals and genuine learning gains (Duolingo, 2024).

Methodology

4.1 Research Design

This study adopts a survey-based research design to examine the long-term and repeated effects of digital nudges in the context of language learning on Duolingo. Unlike purely observational or platform-restricted datasets, the survey allowed the collection of self-reported data on motivational factors, perceptions of nudges, and engagement habits, thereby combining both behavioral and attitudinal measures.

The survey was structured to reduce ambiguity and enhance reliability, with 17 core questions (excluding those filtering for consent and participant eligibility). The questions were varied in type and format, combining Likert-scale items, multiple-choice responses, and open-ended questions to generate both quantitative and qualitative insights. This mix enabled the collection of standardized data suitable for statistical analysis while also allowing respondents to elaborate on their experiences with digital nudges.

The survey yielded 146 valid responses, which provided a sufficient sample size for conducting regression analysis. In line with the study's hypotheses, the quantitative data was primarily analyzed using logistic regression techniques to test the influence of nudges on repeated engagement behaviors. Regression models were chosen because they allow for the assessment of relationships between independent variables and dependent variables.

By combining a structured survey instrument with statistical modeling, this design ensures that the analysis moves beyond descriptive insights to provide evidence of causal or correlative relationships. The inclusion of open-ended survey responses also contributes qualitative depth, supporting the exploration of the hypotheses, particularly Hypothesis 3 regarding the potential reduction in required nudges as habits become established. Overall, the design provides the framework necessary to investigate the role of digital nudges in shaping sustained user behavior on Duolingo.

4.2 Sample Selection

The participants of this study were exclusively Duolingo users. To ensure that the analysis captured the long-term effects of digital nudging mechanisms within the platform, a specific inclusion criterion was applied: respondents were required to have used Duolingo consistently for more than one year and to have achieved a streak of at least 365 consecutive days. This threshold was chosen because it reflects both sustained engagement and repeated exposure to Duolingo's nudging strategies. By setting this criterion, the study ensured that participants had sufficient experience to provide meaningful insights into the persistence effects of digital nudges.

Loss of a streak was not treated as grounds for exclusion, provided the participant had previously achieved a streak of 365 days or more and the streak-loss was recent. This adjustment acknowledged the reality of temporary disengagement while still recognizing the value of long-term behavior patterns. It also allowed for a more accurate reflection of the dynamic nature of engagement with digital learning platforms, where breaks in participation may occur but do not necessarily diminish the overall influence of nudging interventions.

In total, the survey collected 163 responses. However, data screening was conducted to ensure reliability and validity. Subsequently, complete-case analysis was used to filter out invalid responses. A subset of participants did not complete the full survey, while others reported implausible streak lengths that rendered their responses unusable. These cases were excluded to preserve the accuracy of the dataset. After this filtering process, the final sample consisted of 146 valid participants, whose responses were used in the subsequent statistical analysis.

The resulting sample represents a population of highly engaged language learners. Such a group is particularly well-suited for investigating how repeated exposure to digital nudges influences persistence over time. Nonetheless, it is acknowledged that this sampling strategy may limit the generalizability of findings to less-engaged or casual Duolingo users, a consideration that is addressed in the limitations of this study.

4.3 Data Analysis Techniques: Quantitative and Qualitative Approaches

To address the research questions on the long-term effectiveness of digital nudges, this study employs a combination of quantitative and qualitative data analysis techniques. This mixed-methods approach allows for a comprehensive understanding of both the measurable outcomes of nudge interventions and the contextual insights that shape user behaviors.

On the quantitative side, the analysis primarily examines Duolingo user data with a focus on repeat behaviors such as streak continuation, daily goal adherence, and lesson completion over time. Descriptive statistics are used to establish baseline patterns of engagement, while logistic regression models are applied to test the probability of streak continuation under varying conditions, such as with or without reminder nudges. These methods provide a means of assessing the relationship between nudging frequency and sustained engagement, providing insight on the long-term effects of consistency and frequency of digital nudges.

The qualitative dimension complements this statistical analysis by interpreting the broader implications of engagement behaviors. Specifically, thematic analysis is conducted on secondary data to capture perceptions of nudges, including whether they are seen as motivating, intrusive, or neutral. These insights provide nuance to the quantitative findings by identifying mechanisms through which nudges influence behavior, and by exploring whether perceived intrusiveness could explain declining effectiveness over time (as suggested by Baker et al., 2016). Furthermore, qualitative interpretation is essential for evaluating Hypothesis 3, which posits that users may require fewer nudges as behaviors become habitual.

Together, the quantitative and qualitative techniques enable triangulation, strengthening the validity of the results. Quantitative models establish measurable patterns in user engagement, while qualitative analysis explains how and why nudges work, or fail to work, over extended periods.

4.4 Ethical Considerations

Ethical integrity was maintained throughout the research process to ensure the protection of participants' rights and the credibility of the findings. The study followed standard ethical guidelines for research involving human participants, with particular emphasis on informed consent, confidentiality, and data protection.

Participation in the survey was entirely voluntary, and respondents were informed about the purpose of the study, its academic nature, and their right to withdraw at any point without penalty. To ensure consent, the survey began with a mandatory agreement statement clarifying these terms. No participant was allowed to continue without providing explicit consent.

To protect privacy, no personally identifiable information such as names, email addresses, or IP data was collected. Responses were recorded anonymously and stored securely in password-protected databases. Only the researcher had access to the raw dataset, and the results were analyzed and reported in aggregate form to prevent the identification of individual participants. These safeguards align with the data protection principles outlined by the European Union's General Data Protection Regulation (GDPR, 2016).

The survey questions focused solely on digital learning behaviors, perceptions, and experiences, and care was taken to frame all questions in a neutral and respectful manner to avoid biasing responses or causing discomfort.

Finally, approval was obtained under relevant university supervision before the commencement of data collection. This ensured that the study complied with institutional requirements and broader ethical standards for research.

Results

5.1 Overview

The results of the statistical analyses conducted to examine the relationship between users' streak length and their self-reported engagement behaviors on Duolingo are outlined in the following section. The analyses aimed to determine whether longer streaks were associated with higher likelihoods of continued use without reminders, app openings without notifications, and continued usage motivated by streak maintenance.

All models were tested using binary logistic regression with complete-case data. Participants with missing responses on any of the variables required for a specific model were excluded listwise. Because the proportion of missing data was minimal and appeared to occur at random, no data imputation was performed. Consequently, the effective sample size varied slightly across models depending on the availability of complete data for each variable.

The following results are presented separately for each behavioral outcome, followed by a discussion of model fit, effect sizes, and interpretive comments concerning the magnitude and direction of associations.

5.2 Demographic Analysis of Participants

Participants in this study reported exceptionally long streaks compared to the average Duolingo user, specifically targeted to fulfil the aim of the research. The representation of highly engaged, long-term users limits the generalizability of the findings to the broader Duolingo population, particularly those in earlier habit-formation stages or with shorter usage histories, and this limitation is discussed further in the following chapter. In the third quarter of 2024, there were nearly 8 million learners on Duolingo with a streak of 365 days or over, while the number of daily active users (DAUs) was 37.2 million in the same period (Duolingo, 2024). Therefore, approximately 21.5% of users had a streak over 365 days.

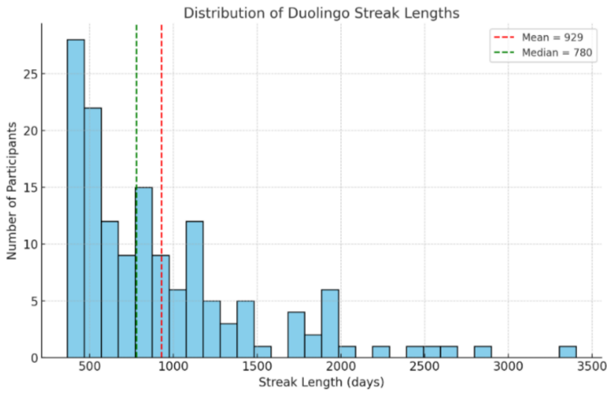

Figure 5.1: Histogram illustrating the distribution of Duolingo streak lengths among participants (N = 146). The mean streak length was 929 days (red dashed line), while the median was 780 days (green dashed line).

In this study, the average streak length was approximately 929 days (SD ≈ 566), with a median of 780 days. Streaks ranged from the minimum eligibility of 365 days to a maximum of 3,405 days, meaning that some participants had maintained continuous use for nearly a decade. The distribution was positively skewed: most streaks fell between 400 and 1,200 days, while a smaller group of "super-users" reported streaks above 2,000 days. This shows that while the sample was anchored by users with one to three years of sustained engagement, a minority represented extreme long-term app use.

When asked why they began using Duolingo, the overwhelming majority cited learning a new language as their primary motivation. Many participants also reported secondary reasons, including learning for fun, building a habit, and school or work-related purposes. A smaller number mentioned more specific personal contexts, such as preparing to move abroad, improving communication with family members, or refreshing previously studied languages. This mix of intrinsic (fun, self-improvement) and extrinsic (school, relocation, career) drivers aligns with prior findings that language learning apps attract users for both personal enrichment and practical needs.

The sample was geographically diverse, though certain regions were more strongly represented. A large proportion of respondents were located in Italy, while sizeable groups came from the United States, United Kingdom, and Germany, alongside smaller numbers from countries such as Canada, Australia, Norway, Spain, and Iran. Several participants reported multiple countries of residence or recent mobility, suggesting an internationally oriented user base. This global spread reflects the accessibility of Duolingo as a free, mobile-first platform.

In terms of language learning goals, Italian was the most commonly reported language, consistent with the sample's demographic concentration in Italy and among those with Italian heritage or family ties. Other popular study choices included German, French, and Spanish, often in combination with Italian. A subset of participants also reported experimenting with less widely studied languages (e.g., Japanese, Arabic, Latin, Welsh), underscoring Duolingo's appeal as a tool for multilingual exploration. The dominance of Italian within the dataset, however, reinforces the need to control for language choice in subsequent analyses, as motivation and engagement patterns may vary by target language.

Overall, the participants represent a highly engaged subset of Duolingo's user base, characterized by long streaks, strong motivation to learn languages, and diverse international backgrounds. Their demographic and usage profiles provide a solid foundation for examining how digital nudges operate over time in an app-based environment and among a highly embedded population.

5.3 Effectiveness of Digital Nudges on Repeated Behaviors

5.3.1 Streak Length and Continued Usage Without Reminders

Model Summary - would_continue_without_reminder

Model	Deviance	AIC	BIC	df	ΔX^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
M_0	94.527	96.527	99.323	120			0.000		0.000	
M_1	87.702	91.702	97.294	119	6.825	0.009	0.072	0.101	0.048	0.055

Note. M_1 includes streak_length

Table 5.1

Coefficients

Model		Estimate	Standard Error	z	Wald Test			95% Confidence interval	
					Wald Statistic	df	p	Lower bound	Upper bound
M_0	(Intercept)	1.881	0.268	7.010	49.144	1	< 0.001	1.355	2.407
M_1	(Intercept)	0.463	0.635	0.730	0.532	1	0.466	-0.781	1.707
	streak_length	0.002	0.001	2.126	4.521	1	0.033	0.000	0.003

Note. would_continue_without_reminder level '1' coded as class 1.

Table 5.2

A binary logistic regression was conducted to examine whether streak length predicted the likelihood that participants would continue using Duolingo without receiving reminder notifications. The outcome variable would_continue_without_reminder was coded as 1 = Yes and 0 = No. Cases with missing data on either the predictor or outcome were excluded using complete-case analysis (no imputation applied).

The model including streak length (M_1) fit the data significantly better than the null model containing only the intercept, $\Delta\chi^2(1) = 6.825$, $p = 0.009$. Model fit and explained variance were modest (AIC = 91.7; BIC = 97.3; df = 119; McFadden $R^2 = 0.072$; Nagelkerke $R^2 = 0.101$; Tjur $R^2 = 0.048$; Cox & Snell $R^2 = 0.055$), which is typical for behavioral data.

The regression coefficient for streak length was positive and statistically significant ($B = 0.002$, $SE = 0.001$, Wald $\chi^2(1) = 4.521$, $p = 0.033$, 95% CI [0.000, 0.003]), indicating a small but reliable association between streak length and the likelihood of continuing to use the app without reminders. The corresponding odds ratio was OR = 1.002 (95% CI [1.000, 1.003]), suggesting that each additional streak day increased the odds of continuing use by about 0.2%. Although the per-day effect is small, it compounds meaningfully over time. A 100-day increase in streak length corresponds to an estimated odds ratio of $\exp(0.002 \times 100) = 1.22$, or approximately 22% higher odds of continuing to engage with the app without reminders, while an additional 365-day streak corresponds to $OR = \exp(0.002 \times 365) \approx 2.08$, which approximately doubles the odds of continued autonomous use of the app. This suggests that as users maintain their streaks, their behavior becomes increasingly self-sustaining, consistent with habit formation theory and reinforcement through repetition (Lally et al., 2010; Wood & Rünger, 2016).

To assess whether this relationship was influenced by language group, a language-control model (M_2) was estimated with categorical predictors for Italian, German and Other languages. The inclusion of language did not significantly improve model fit, $\Delta\chi^2(3) = 7.820$, $p = 0.099$, and none of the language coefficients reached significance. Thus, linguistic context did not meaningfully affect the streak–continuation association.

An exploratory interaction term testing whether the strength of this effect depended on total Duolingo usage duration (streak_length × usage_length_years) was also non-significant ($B = -0.001$, $SE = 0.001$, $p = 0.366$). This indicates that the relationship between streak length and continued use without reminders remained stable regardless of how long respondents had been using the app.

In summary, Model M_1 provides evidence of a modest yet statistically significant positive relationship between streak length and autonomous continued use. Although the incremental impact per day is small, its cumulative effect over time, or extended streaks, supports the notion that digital nudges can foster self-reinforced engagement over time.

5.3.2 Streak Length and App Open Behavior Without Reminders

Model Summary - opened_without_reminder

Model	Deviance	AIC	BIC	df	ΔX^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
M_0	36.668	38.668	41.652	145			0.000		0.000	
M_1	32.586	36.586	42.554	144	4.081	0.043	0.111	0.124	0.027	0.028

Note. M_1 includes streak_length

Table 5.3

Coefficients

					Wald Test			95% Confidence interval	
Model		Estimate	Standard Error	z	Wald Statistic	df	p	Lower bound	Upper bound
M_0	(Intercept)	3.570	0.507	7.041	49.570	1	< 0.001	2.576	4.563
M_1	(Intercept)	0.948	1.595	0.595	0.354	1	0.552	-2.178	4.074
	streak_length	0.004	0.003	1.399	1.959	1	0.162	-0.002	0.010

Note. opened_without_reminder level 'Yes' coded as class 1.

Table 5.4

A binary logistic regression was conducted to assess whether streak length predicted the likelihood of participants opening the Duolingo app without being reminded. The dependent variable opened_without_reminder was coded as 1 = Yes and 0 = No. Cases with missing data on either the predictor or outcome were excluded using complete-case analysis (no imputation applied).

The inclusion of streak length (M_1) modestly improved model fit compared to the intercept-only model (M_0), $\Delta\chi^2(1) = 4.081$, $p = 0.043$. Although the improvement was small, it indicates that streak length offered some incremental explanatory value for app-opening behavior without reminders. Model fit and pseudo-R^2 values were low to moderate (AIC = 36.6; BIC = 42.6; df = 144; McFadden $R^2 = 0.111$; Nagelkerke $R^2 = 0.124$; Tjur $R^2 = 0.027$; Cox & Snell $R^2 = 0.028$), consistent with typical behavioral data patterns.

The regression coefficient for streak length was positive but not statistically significant ($B = 0.004$, $SE = 0.003$, Wald $\chi^2(1) = 1.959$, $p = 0.162$, 95% CI [−0.002, 0.010]). The corresponding odds ratio (OR = 1.004, 95% CI [0.998, 1.010]) suggests that each additional streak day increased the odds of opening the app without reminders by approximately 0.4%. If the estimated trend were reliable, a 100-day increase in streak length would correspond to an odds ratio of approximately 1.49, indicating about 49% higher odds of opening the app without reminders. However, given the wide confidence intervals and non-significant coefficient, this figure should be interpreted only as an illustrative directional estimate.

While the direction aligns with the hypothesis that longer streaks foster greater autonomous engagement, the confidence interval spans zero, indicating that the effect is uncertain and should be interpreted with caution. The intercept was also non-significant ($B = 0.948$, $SE = 1.595$, $p = 0.552$), reflecting substantial variability in baseline autonomous engagement across users. Although model diagnostics (AIC and BIC) suggest a small improvement in fit relative to the null model, the effect does not reach conventional thresholds for reliable inference.

Language group was added as a categorical control variable (Italian, German, and Other, with Italian as the reference category). None of the language coefficients reached statistical significance. These results confirm that language group membership did not meaningfully predict autonomous engagement behavior.

In summary, Model M_1 exhibits a trend that is directionally consistent but imprecise. Users with longer streaks tended to open the app without reminders slightly more often, but the association lacks statistical robustness. This suggests that while streak length may relate to intrinsic engagement, the present evidence is insufficient to confirm a meaningful behavioral effect.

5.3.3 Streak Length and Self-Reported Continuation for Streak Maintenance

Model Summary - continued_usage_for_streak

Model	Deviance	AIC	BIC	df	ΔX^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
M_0	136.794	138.794	141.778	145			0.000		0.000	
M_1	136.794	140.794	146.761	144	0.000	0.993	5.055×10^{-7}	7.787×10^{-7}	4.282×10^{-7}	4.736×10^{-7}

Note. M_1 includes streak_length

Table 5.5

Coefficients

Model		Estimate	Standard Error	z	Wald Test			95% Confidence interval	
					Wald Statistic	df	p	Lower bound	Upper bound
M_0	(Intercept)	1.529	0.216	7.070	49.985	1	< 0.001	1.105	1.953
M_1	(Intercept)	1.532	0.417	3.677	13.523	1	< 0.001	0.716	2.349
	streak_length	-0.000	0.000	-0.008	6.919×10^{-5}	1	0.993	-0.001	0.001

Note. continued_usage_for_streak level 'Yes' coded as class 1.

Table 5.6

A binary logistic regression was conducted to assess whether streak length predicted participants' likelihood of continuing to use Duolingo primarily to maintain their streak. The dependent variable continued_usage_for_streak was coded as 1 = Yes and 0 = No. Cases with missing data for either the predictor or outcome were excluded using complete-case analysis (no imputation was performed).

Adding streak length (M_1) did not improve model fit relative to the intercept-only model (M_0), $\Delta\chi^2(1) = 0.000$, p = 0.993. The AIC (140.8) and BIC (146.8) values were nearly identical to those of the null model, confirming the absence of explanatory improvement. Pseudo-R^2 values were effectively zero (McFadden $R^2 < 0.001$;

Nagelkerke $R^2 < 0.001$; Tjur $R^2 < 0.001$; Cox & Snell $R^2 < 0.001$), indicating that streak length accounted for virtually none of the variance in participants' self-reported motivation.

The regression coefficient for streak length was non-significant and essentially zero ($B = -0.000$, $SE = 0.000$, Wald $\chi^2(1) < 0.001$, $p = 0.993$, 95% CI [−0.001, 0.001]), corresponding to an odds ratio (OR = 1.000, 95% CI [0.999, 1.001]). This indicates no measurable association between streak length and the probability of reporting that continued app use was motivated by maintaining a streak. The intercept was statistically significant ($B = 1.532$, $SE = 0.417$, $p < 0.001$, 95% CI [0.716, 2.349]), implying that some participants generally identified streak preservation as a motivational factor, but the duration of their streak did not meaningfully influence this likelihood.

To explore potential confounding by language group, an extended model was estimated with language (Italian, German, Other) included as a categorical control variable. This model did not significantly improve fit, $\Delta\chi^2(3) = 2.390$, $p = 0.664$, and streak length remained non-significant ($B = 0.000$, $p = 0.932$). None of the language categories emerged as significant predictors, although the "Other" group displayed a small, non-significant positive trend ($B = 0.633$, $p = 0.155$).

In summary, there was no evidence of association between streak length and participants' self-reported continuation of app use for the purpose of maintaining their streak. This finding suggests that even among users with extended streaks, the accumulation of streak days does not systematically shape whether they explicitly identify streak preservation as their primary motivation. This lack of association highlights a possible gap between self-reported motives and underlying behavior, meaning that streak effects may operate implicitly, reinforcing engagement through habit formation and consistency rather than through conscious goal pursuit.

Discussion

6.1 Interpretation of Findings: Implications for Theory and Practice

The findings of this study provide a nuanced perspective on how digital nudges shape repeated behaviors in a language learning context, offering both theoretical insight and practical guidance. Across the three models, the results indicate that nudges, particularly those tied to streak length, are associated with sustained engagement, though their influence varies depending on the specific behavior examined. The small but statistically significant effect of streak length on continued use without reminders suggests that persistent exposure to nudges may indeed contribute to longer-term engagement patterns. In more specific terms, an additional 100 days of streak corresponded to approximately 22% higher odds of continuing to use the app without reminders, and an additional 365 days corresponded to roughly double the overall odds. At the same time, the non-significant yet directionally consistent trend in the app-open-without-reminders model points to a weaker and less certain relationship, while the absence of any association between streak length and continued usage for streak maintenance underscores the complexity of linking perceived motivation to actual behavior. Taken together, these results paint a picture of digital nudges as potentially influential in shaping sustained behaviors, but not uniformly so and within the methodological limits of this study.

From a theoretical standpoint, these results align closely with existing models of habit formation, which suggest that repeated behaviors become increasingly automatic over time as they are reinforced by contextual cues (Lally et al., 2010; Wood & Rünger, 2016). The significant association between streak length and continued usage without reminders supports the idea that consistent engagement over long periods can strengthen behavioral automaticity. In other words, once a user has repeated the behavior enough times, such as completing a daily lesson in this context, the need for external prompts decreases. This is consistent with habit theory's emphasis on the transition from action requiring effort to habitual behavior that occurs almost automatically in response to environmental cues.

However, it is crucial to interpret this result cautiously. The current study cannot establish a causal relationship between nudges and habit formation. The cross-sectional, self-reported nature of the data means that stronger habits among long-streak users are plausible but cannot be definitively attributed to the nudges themselves. It is equally possible that individuals with a natural predisposition toward consistent behavior are more likely to maintain long streaks, or that other factors, such as intrinsic motivation or learning goals, play a mediating role. Nevertheless, the observed pattern implies that nudges may help support the shaping of behavior in its early stages, but as repetition strengthens automaticity, the external intervention becomes progressively less necessary.

The findings also resonate with literature on commitment devices, which are mechanisms individuals use to bind their future behavior to current intentions (Bryan et al., 2010). Duolingo's streak system is a quintessential example of this. By visibly displaying records of past effort, it leverages psychological principles such as loss aversion. The significant relationship between streak length and continued engagement even without reminders suggests that once users have built a sufficiently long streak, they feel compelled to maintain it, not merely because of external prompts, but because breaking it would mean losing something they have already invested in. This aligns with research showing that commitment devices are most effective when the potential loss of progress is salient (Thaler & Benartzi, 2004), creating a powerful internal motivation to persist even in the absence of external cues. The results in the app-open-without-reminders model further support this interpretation. While the relationship is not statistically significant, the observed trend implies that the psychological weight of a long streak might influence user behavior beyond the presence of nudges alone. This suggests that the streak may evolve from a nudge-based feature into a self-imposed accountability tool that continues to shape behavior even when the original nudge (e.g. a notification) is removed.

Although the regression analysis found no significant association between streak length and self-reported motivation to maintain the streak, the descriptive survey data suggest a more complex relationship. Over 80% of respondents reported that they had completed lessons primarily to preserve their streak, yet this motivation did not vary systematically with streak length. This indicates that while streak-based nudges serve as a widely acknowledged source of motivation, their psychological function may change as engagement becomes habitual. For many users, the streak operates as an explicit commitment device, whereas for long-term users, it may function more as an implicit behavioral cue, sustaining action through habit rather than conscious goal pursuit. This reconciles the statistical and self-reported findings, indicating that the lack of association in model 3, app-open-without-reminders model, does not imply that streak maintenance is unimportant, but rather that its influence changes from conscious motivation to automated behavioral reinforcement as users become more experienced.

Beyond theory, the findings offer several implications for practice, particularly in the design of digital products that aim to support long-term engagement. First, the evidence suggests that nudges are most impactful in the early phases of behavior formation, when repetition and reinforcement are still needed to establish a routine. This implies that product designers should focus on delivering frequent, salient nudges during onboarding and early use stages. Once behaviors become more habitual, however, over-reliance on external prompts could backfire, as users may perceive them as intrusive or unnecessary. At this stage, strategies that emphasize self-monitoring, visible progress markers, and goal-setting may be more effective than simple reminders.

Second, the results highlight the value of integrating commitment-based design features, such as streaks or cumulative achievement indicators, into digital platforms. These features can transform an externally nudged behavior into a self-sustained one by leveraging users' natural aversion to loss and their desire to maintain consistency with past actions. Designing nudges that evolve into commitment devices rather than remaining static notifications may be key to sustaining engagement over the long term.

Finally, the disconnect between self-reported motivation and actual behavior suggests that users' awareness of why they act may not always align with the factors that drive their engagement. This has practical implications for how digital products gather feedback and evaluate success as self-reported satisfaction surveys may fail to capture the underlying mechanisms of engagement, highlighting the importance of behavioral metrics and longitudinal tracking in product evaluation.

6.2 Comparison with Existing Literature

The findings of this study extend the existing body of literature on digital nudging and habit formation by examining their interaction over a sustained period of user engagement. Prior studies have largely investigated the short-term influence of nudges on behavior, emphasizing their ability to prompt immediate decisions or temporary engagement (Thaler & Sunstein, 2008; Weinmann et al., 2016). In contrast, the present study highlights that digital nudges can evolve in function over time. While they initially serve as external motivators, their impact appears to diminish as habits form, therefore making them useful at the earlier stages of behavior formation, but less essential once intrinsic mechanisms of motivation and automaticity take over.

This pattern resonates with research on habit formation, which proposes that consistent repetition in stable contexts strengthens automatic behavioral responses (Wood & Rünger, 2016). In the case of Duolingo users with long streaks, the platform's reminder and streak features appear to have initiated behaviors that later became self-sustaining. The observed continuity of use without reminders suggests that these users have internalized the cues originally introduced by the platform. This is consistent with Lally et al. (2010), who found that the median time required for a behavior to become automatic was approximately 66 days, with substantial individual variability within a range of 18 to 254 days. The extended time frames observed in this study suggest that for digital behaviors reinforced by gamified systems and consistent reminders, automaticity can be maintained over far longer periods once the motivational basis is established.

At the same time, the modest size of the observed effects supports prior evidence that the behavioral influence of nudges is typically small to moderate in magnitude. Mertens et al. (2022) found that across multiple domains, choice architecture interventions produced an average effect size of Cohen's d = 0.43, underscoring the more subtle nature of many nudging mechanisms. The results here align with that outcome. Duolingo's nudges appear to maintain engagement over time but do not necessarily predict stronger motivation or consistent self-reported behavior change. This indicates that while nudges may play an important supportive role, they are unlikely to serve as independent drivers of long-term behavior without reinforcement from other psychological mechanisms such as habit formation or self-regulation.

The diminishing effectiveness of frequent nudges observed in this study also reflects a recurring theme in the literature. Saig and Rosenfeld (2023) caution that repetitive digital prompts can lead to nudge fatigue. In the present findings, users who had maintained long streaks appeared to require fewer reminders to continue engaging with the app, suggesting a transition from external to internalized behavioral control. This phenomenon also echoes arguments from the broader digital behavior literature that overexposure to reminders can reduce their effect, reinforcing the importance of managing frequency and timing in nudge design.

The relationship observed between external prompts and sustained engagement mirrors the mechanisms described in research on nudging systems. On Duolingo, features such as streaks and progress notifications initially act as extrinsic motivators by providing immediate feedback and a sense of reward. Over time, however, these same features appear to contribute to the stabilization of use patterns, helping transform externally prompted actions into self-sustaining routines. The fact that many participants reported a willingness to continue learning without reminders supports this interpretation, suggesting that these gamification elements can evolve from short-term incentive structures into habits, a process consistent with behavioral reinforcement models (Weinmann et al., 2016; Wood & Rünger, 2016).

Overall, these findings contribute to a more nuanced understanding of digital nudging as a time-sensitive mechanism. In alignment with previous work, the results suggest that the strength of a nudge is not static but evolves alongside the user's behavioral consistency. Nudges appear most effective when behaviors are still under construction, and become more redundant as habits form. This interpretation reinforces the importance of considering temporal dynamics in the design of behavioral interventions.

6.3 Evaluation of Research Questions and Hypotheses

The central research question guiding this study asked how digital nudges influence repeated behaviors over a long period of time. The results provide evidence that digital nudges, such as Duolingo's streak reminders and motivational notifications, can play a meaningful role in establishing and maintaining user engagement.

Research Question: How do consistent digital nudges change or impact repeat behaviors over a long period of time?

The results demonstrate that digital nudges initially act as effective behavioral triggers that help reinforce consistency in engagement. In the context of Duolingo, nudges such as streak reminders and progress notifications encourage users to return to the app and complete daily lessons. Over time, however, the relationship between nudging and continued engagement becomes less direct. Users with longer streaks, referring to those who have maintained consistent app use for over 365 days, showed a great tendency to continue practicing without reminders. This finding suggests that the function of nudges evolves: they facilitate early engagement and consistency but eventually lead behaviors to transition from deliberate to automatic. This interpretation directly answers the research question by showing that digital nudges influence behavior strongly during the initial stages of engagement, leading to habituation in the long term. This dynamic aligns with broader habit formation theory (Lally et al., 2010; Wood & Rünger, 2016), where repeated behaviors eventually become self-sustaining through internalized cues. Thus, while nudges remain important for initiating and stabilizing behavior, their long-term influence is mediated by the development of automaticity.

Hypothesis 1: Digital Nudges Impact Repeat Behaviors After a Long Period of Consistency

Hypothesis 1 proposed that digital nudges would influence repeated behaviors after extended exposure, gradually promoting sustained engagement. The results largely support this assumption, though the effects observed were modest. The regression model examining continued use without reminders indicated a positive association between streak length and engagement, suggesting that users repeatedly exposed to nudges over time were more likely to maintain app use even when external prompts were absent. To further put this into context, extending a streak by 100 days translates into about a 22% increase in the likelihood of continuing without reminders, and after an additional full year or 365 days of consistent daily use, the odds of autonomous engagement are roughly 100% higher. This pattern aligns with the notion that consistent nudging can facilitate habit formation, where behavior becomes progressively self-sustaining. Additionally, the app-open-without-reminders model, which used binary data on whether participants ever opened the app and completed a lesson without a reminder, suggested that users with longer streaks tended to engage with the app without relying on reminders slightly more often, and while the evidence was not sufficient for a solid conclusion to be drawn, the trend was directionally consistent and should be further investigated in future studies.

However, the relatively small effect sizes indicate that nudges alone do not account for long-term persistence. Their influence appears to diminish as behaviors become internalized, with habit strength, intrinsic motivation, and personal commitment playing a role in sustaining engagement. In this sense, the long-term impact of nudges can be explained in the following way: they serve as early-stage catalysts that help establish behavioral regularity, after which the behavior is maintained through automaticity rather than continued external prompting.

Hypothesis 2: Over Time, Users Require Fewer Nudges for the Same Behavior

Hypothesis 2 proposed that users require fewer nudges as their behavior become more habitual. The results provide strong indicative support for this hypothesis. Long-term Duolingo users demonstrated the ability to continue using the app and complete lessons without the need for reminder notifications. This finding is consistent with the idea of behavioral automaticity described in habit theory (Wood & Rünger, 2016), where environmental cues or established routines trigger behavior without conscious deliberation.

The statistical evidence indicated that although nudges were effective in encouraging early engagement, their absence did not prevent continued use among experienced users. This suggests a developmental trajectory in which external interventions gradually reduce in necessity as users internalize the behavior. Such a pattern also aligns with the concept of nudge fatigue (Saig and Rosenfeld, 2023), wherein repeated prompts may lose salience once users are sufficiently self-regulated. Thus, the data support the interpretation that digital nudges serve as an initial catalyst for behavior formation but are eventually superseded by intrinsic or habitual processes.

In summary, the findings provide an understanding of the relationship between digital nudges and long-term behavioral engagement. The results support Hypotheses 1 and 2, confirming that nudges influence behavior over time but that their impact diminishes as consistency and habit strength increase. The study demonstrates that digital nudges can serve as both initiators and stabilizers of behavioral change, though their necessity decreases as automaticity develops.

6.4 Limitations of the Study and Future Directions

While this study provides meaningful insights into the long-term effects of digital nudges, a number of limitations should be acknowledged. The most notable limitation concerns the sample composition. The participants were predominantly long-term Duolingo users with established streaks of over 365 days. This group represents a population of highly engaged and motivated individuals, who can be described as “super-users,” who are likely to exhibit stronger commitment and habit strength than average users. As a result, the findings may not generalize to new or less consistent users, whose engagement with nudges may differ substantially. It is also plausible that the effects observed in this study are partly driven by selection effects: only users with relatively high intrinsic motivation remain active long enough to accumulate these long streaks. In other words, the reduced need for reminders among long-streak users may reflect the characteristics of those users, rather than the causal impact of nudges themselves. Future studies should therefore include a more heterogeneous sample that captures early-stage users and varying levels of commitment, allowing for a clearer understanding of how nudges operate across different stages of habit formation.

Another limitation involves the cross-sectional design of the study. Because the data were collected at a single point in time and relied on self-reported behaviors, it was not possible to establish causal relationships between nudges and behavioral outcomes. This means that, while the analysis revealed associations between streak length, reminders, and continued engagement, the direction of causality cannot be definitively determined. It remains unclear whether nudges directly lead to sustained use, or whether already motivated users are more responsive to nudging mechanisms. A longitudinal or experimental design would provide greater clarity by observing how engagement changes over time in response to specific nudge interventions.

The study also relied on self-reported data, which can introduce biases such as social desirability or inaccurate recall. Although measures were taken to minimize ambiguity and standardize question wording, self-perception of motivation and engagement may differ from actual in-app behavior. This also means that the interpretation of these results must be conservative on the grounds that if users systematically overestimate their autonomy or underreport their reliance on reminders, the observed trend that long-streak users appear more "self-sustaining" could partially reflect a distortion introduced by biased self-report rather than an actual behavioral shift. Future research could address this limitation by integrating self-reported measures with objective behavioral data from platforms such as Duolingo's analytics. This would allow researchers to verify patterns of engagement, reminder response rates, and streak continuity with higher accuracy.

Finally, the study's focus on Duolingo limits the generalizability of findings to other digital learning platforms or behavioral contexts. The underlying mechanisms in nudging and habit formation may differ depending on the type of behavior being encouraged, the platform's reward structure, and the user's intrinsic motivation. Comparative research across multiple platforms could help determine whether the patterns observed here, particularly the diminishing role of nudges over time, are unique to Duolingo or represent a broader trend in digital behavior design.

Conclusion

7.1 Summary of Key Findings

This study examined how digital nudges influence repeated behaviors over time, using Duolingo as a case study. The main aim was to understand whether consistent exposure to digital nudges encourages long-term engagement and whether users eventually maintain the same behaviors with reduced reliance on these prompts.

The findings indicate that consistent nudging plays an important, albeit modest, role in sustaining repeated engagement. The logistic regression analysis showed a small but statistically significant relationship between streak length and continued app use without reminders. This suggests that users exposed to digital nudges for longer periods are more likely to sustain the behavior even when external prompts are absent. In contrast, the model predicting self-reported continuation for streak maintenance showed no significant association, while the model predicting app-open behavior was not statistically significant but followed the same positive trend.

Overall, these results imply that the influence of digital nudges may change over time, shifting from externally prompted engagement to more autonomous, habitual behavior. In this sense, nudges appear most effective in the early stages of use, when users are still building a routine. As engagement becomes more consistent, the role of nudges may diminish, giving way to internal motivators as habituation occurs. These findings support the notion that digital nudging can help establish behavioral regularity and impact repeat behaviors over a long period of consistency.

7.2 Contributions to the Field of Behavioral Economics

This research contributes to behavioral economics by extending the understanding of how digital nudges operate over prolonged periods of use. While previous studies have primarily focused on short-term behavioral change or one-time decision outcomes (Weinmann et al., 2016; Mertens et al., 2022), this study highlights their potential role in supporting ongoing, repeated engagement. By examining long-term Duolingo users with extended streaks, the research adds empirical evidence that consistency in digital nudging can facilitate a gradual transition from externally driven compliance to internalized habit formation.

The findings also reinforce the importance of considering temporal dynamics in behavioral interventions. In line with habit formation theory (Wood & Rünger, 2016; Lally et al., 2010), the study shows that the effectiveness of nudges is not static but evolves as users progress from deliberate engagement arising from external cues to automated behavior. Furthermore, the results echo ethical perspectives within digital nudging research (Meske & Amojo, 2020), emphasizing that nudges must be applied with contextual sensitivity to avoid overuse.

Overall, this thesis contributes to the growing body of evidence that digital nudges, when designed and implemented consistently, can bridge the gap between short-term decision support and long-term behavioral reinforcement. However, it also cautions that their influence is modest and likely indirect, most effective as early catalysts within a larger behavioral ecosystem rather than as stand-alone tools for sustained change. Future research should continue to explore these mechanisms in longitudinal and cross-platform contexts.

7.3 Practical Recommendations for Digital Nudge Design

The findings of this study offer several practical insights for designers and organizations that implement digital nudges in learning or behavioral-change platforms. Firstly, consistency matters. Regular, well-timed reminders help users establish behavioral patterns, especially during early engagement. However, once habits begin to form, the same interventions could gradually be reduced or diversified. Secondly, nudges should support self-regulation, not replace it. The goal of any digital intervention should be to help users internalize desired behaviors, making them sustainable without continuous prompting. Integrating features that encourage reflection, such as progress summaries or milestone achievements, can help reinforce intrinsic motivation while reducing dependency on external cues.

Given that a 100- and 365-day increase in streak length were associated with roughly 22% and 100% higher odds of continued use without reminders, nudge designers should prioritise intervention intensity early in the user lifecycle, when behaviors are still fragile. In practice, this means increasing prompt frequency, reward signalling, and loss-framing during the early phase of routine establishment, and then progressively dialing back the intensity once consistent daily action is established. After a period of regular engagement, nudges should transition toward maintenance rather than motivational triggers. For example, reminders could shift from direct prompts like “Do your lesson now!” to progress anchors like “You are 120 days into your streak” or identity cues such as “You are becoming consistent”. This approach uses nudges to build the initial routine, while allowing habit momentum to take over later, which is consistent with the finding that long streaks predict autonomous usage even in the absence of external reminders.

Personalization is key. Users respond differently to nudging frequency, tone, and timing. Systems that adapt nudge intensity based on user history, such as decreasing reminders for long-streak users, can be more effective in maintaining engagement without diminishing satisfaction. Designers should also remain mindful of ethical considerations, ensuring that nudges remain transparent, non-intrusive, and aligned with users' goals (Meske & Amojo, 2020). Ethical and context-aware design not only protects user autonomy but can also enhance trust and long-term engagement.

References

- Meske, C., & Amojo, I. (2019). Status quo, critical reflection and road ahead of digital nudging in information systems research – A discussion with Markus Weinmann and Alexey Voinov. *Communications of the Association for Information Systems*, 46, Paper 17. https://doi.org/10.17705/1CAIS.04617
- Mirsch, T., Lehrer, C., & Jung, R. (2017). Digital Nudging: Altering User Behavior in Digital Environments. In *Proceedings of the 13th International Conference on Wirtschaftsinformatik (WI 2017)* (pp. 634–648). Association for Information Systems. http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1076&context=wi2017z
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.
- Weinmann, M., Schneider, C., & vom Brocke, J. (2016). Digital nudging: Guiding judgment and decision-making in digital choice environments. *Business & Information Systems Engineering*, 58(6), 433–436. https://doi.org/10.1007/s12599-016-0453-1
- Kizilcec, R. F., Reich, J., Yeomans, M., Dann, C., Brunskill, E., Lopez, G., & Tingley, D. (2020). Scaling up behavioral science interventions in online education. Proceedings of the National Academy of Sciences, 117(26), 14900–14905. https://doi.org/10.1073/pnas.1921417117
- Liu, X., Xu, Y., Yang, Q., & Zhou, H. (2023). Examining the impact of digital nudging on user interaction with e-learning platforms: A study of working professionals in distance education. Education and Information Technologies, 28, 5829–5849. https://doi.org/10.1007/s10639-025-13600-x
- Meske, C., & Amojo, I. (2020). Ethical guidelines for the construction of digital nudges. arXiv preprint. https://arxiv.org/abs/2003.05249

- Mirsch, T., Lehrer, C., & Jung, R. (2017). Digital nudging: Altering user behavior in digital environments. Proceedings of the 13th International Conference on Wirtschaftsinformatik, 634–648. https://aisel.aisnet.org/wi2017/track03/paper11
- Bird, K. A., Castleman, B. L., Denning, J. T., Goodman, J., Lamberton, C. L., & Rosinger, K. O. (2019). Nudging at scale: Experimental evidence from FAFSA completion campaigns. *National Bureau of Economic Research Working Paper Series*, No. 26158. https://doi.org/10.3386/w26158
- Chiam, J., Lim, A., Teredesai, A., et al. (2024). NudgeRank: Digital algorithmic nudging for personalized health. *arXiv preprint*. https://arxiv.org/abs/2407.20241
- DellaVigna, S., & Linos, E. (2020). RCTs to scale: Comprehensive evidence from two nudge units. *National Bureau of Economic Research Working Paper Series*, No. 27594. https://doi.org/10.3386/w27594
- Hyytinen, A., Tuomala, J., & Hammar, M. (2022). Enhancing the adoption of digital public services: Evidence from a large-scale field experiment. *Government Information Quarterly*, 39(3), 101687. https://doi.org/10.1016/j.giq.2022.101687
- Nazaret, A., & Sapiro, G. (2023). A large-scale observational study of the causal effects of a behavioral health nudge. *Science Advances*, 9(38), eadi1752. https://doi.org/10.1126/sciadv.adi1752
- Duolingo. (n.d.). *What is a streak?* Duolingo Help Center. Retrieved August 1, 2025, from https://www.duolingo.com/help/what-is-a-streak
- Duolingo. (n.d.). *What are leaderboards and leagues?* Duolingo Help Center. Retrieved August 1, 2025, from https://www.duolingo.com/help/leaderboards-and-league

- Gyulai Zsófia, Revesz Balazs (2023), Nudging in the Digital World: An Up-to-Date Systematic Literature Review. Proceedings of the European Marketing Academy, (117268)
- Troussard, X., Sousa Lourenço, J., Rafael Almeida, S., & Ciriolo, E. (2016). *Behavioural insights applied to policy : European report 2016*, Publications Office. https://data.europa.eu/doi/10.2760/903938
- Duolingo. (2024). About us. Retrieved August 8, 2025, from https://www.duolingo.com/about
- Duolingo. (n.d.). How Duolingo keeps you motivated. Retrieved August 8, 2025, from https://www.duolingo.com/help
- Weissmann, J. (2014, March 26). *The empty jar vs. the glowing screen: How do e-payments affect tipping?* Slate. http://www.slate.com/articles/business/moneybox/2014/03/starbucks_square_and_e_payments_do_new_point_of_payment_systems_increase.html
- Meske, C., & Potthoff, T. (2017). The DINU-model – A process model for the design of nudges. *Proceedings of the 25th European Conference on Information Systems (ECIS)*, 2587–2597. https://aisel.aisnet.org/ecis2017_rip/11
- Duolingo. (2023, February 17). *Meaningful metrics: How data sharpened the focus of product teams*. Duolingo Blog. https://blog.duolingo.com/growth-model-duolingo/
- Duolingo. (2024, June 13). *Understanding Duolingo's Time Spent Learning Well metric*. Duolingo Blog. https://blog.duolingo.com/time-spent-learning-well/
- Duolingo. (2025). *Duolingo reports second quarter 2025 financial results* [Press release]. Duolingo Investor Relations. https://investors.duolingo.com
- Duolingo. (n.d.). *What is a streak?* Duolingo. https://www.duolingo.com/help/what-is-a-streak

- Duolingo Help Center. (n.d.-b). *What are leaderboards and leagues?* Duolingo. https://www.duolingo.com/help/leaderboards-and-league
- Kizilcec, R. F., Reich, J., Yeomans, M., Dann, C., Brunskill, E., Lopez, G., & Tingley, D. (2020). Scaling up behavioral science interventions in online education. *Proceedings of the National Academy of Sciences*, *117*(26), 14900–14905. https://doi.org/10.1073/pnas.1921417117
- Karlan, D., McConnell, M., Mullainathan, S., & Zinman, J. (2016). Getting to the Top of Mind: How Reminders Increase Saving. *Management Science*, *62*(12), 3393–3411. http://www.jstor.org/stable/44166531
- Brandon, A., Ferraro, P. J., List, J. A., Metcalfe, R. D., Price, M. K., & Rundhammer, F. (2017, revised 2022). Do the effects of nudges persist? Theory and evidence from 38 natural field experiments (NBER Working Paper No. 23277). National Bureau of Economic Research. https://doi.org/10.3386/w23277
- Mertens, S., Herberz, M., Hahnel, U. J. J., & Brosch, T. (2022). The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. Proceedings of the National Academy of Sciences, 119(1), e2107346118. https://doi.org/10.1073/pnas.2107346118
- Zhu, Y., Long, Y., Wang, H., Lee, K. P., Zhang, L., & Wang, S. J. (2024). Digital behavior change intervention designs for habit formation: Systematic review. Journal of Medical Internet Research, 26, e54375. https://doi.org/10.2196/54375
- Grüning, D. J., Riedel, F., & Lorenz-Spreen, P. (2023). Directing smartphone use through the self-nudge app one sec. Proceedings of the National Academy of Sciences of the United States of America, 120(8), e2213114120. https://doi.org/10.1073/pnas.2213114120

- Weijers, R. J., de Koning, B., & Paas, F. (2021). Nudging in education: From theory towards guidelines for successful implementation. European Journal of Psychology of Education, 36(4), 883–902. https://doi.org/10.1007/s10212-020-00495-0
- Weijers, R., de Koning, B., Vermetten, Y., & Paas, F. (2023). Nudging autonomous learning behavior: Three field experiments. Education Sciences, 13(1), 49. https://doi.org/10.3390/educsci13010049
- Vermetten, Y. J. M., de Wild, A. F., Stam, K., Bastiaansen, M., & Klatter, E. B. (2017, June). Nudging the student for study success: A long-term experiment in a Dutch University of Applied Sciences [Poster]. WINK Conference, Utrecht, the Netherlands. Rotterdam University of Applied Sciences. Retrieved from https://www.rotterdamuas.com/research/projects-and-publications/pub/nudging-the-student-for-study-success-a-long-/f7da828d-f492-4ac5-9e86-be100d176ae4/
- Kay, E., & Bostock, P. (2023). The power of the nudge: Technology driving persistence. Student Success, 14(2), 8–18. https://doi.org/10.5204/ssj.2848
- Sasaki, S., Kurokawa, H., & Ohtake, F. (2021). Effective but fragile? Responses to repeated nudge-based messages for preventing the spread of COVID-19 infection. Japanese Economic Review, 72(3), 371–408. https://doi.org/10.1007/s42973-021-00076-w
- Forberger, S., Wichmann, F., & Comito, C. N. (2022). Nudges used to promote physical activity and to reduce sedentary behaviour in the workplace: Results of a scoping review. Preventive medicine, 155, 106922. https://doi.org/10.1016/j.ypmed.2021.106922
- Venema, T. A. G., Kroese, F. M., & de Ridder, D. T. D. (2018). I'm still standing: A longitudinal study on the effect of a default nudge. Psychology & Health, 33(5), 669–681. https://doi.org/10.1080/08870446.2017.1385786

- Johannsdottir, L., Sigurðardóttir, S., Óskarsdóttir, M., & Islind, A. (2025). Digital nudging in digital health technologies: A systematic review. Health and Technology. https://doi.org/10.1007/s12553-025-01000-7
- Ling, M., Xu, L., & Yang, H. (2023). Direct and spillover effects of social norm nudges for household recycling: A longitudinal field experiment. Sustainable Production and Consumption, 37, 978–988. https://doi.org/10.1016/j.spc.2023.06.001
- Linder, N., Borgström, S., Devine-Wright, H., & Västfjäll, D. (2018). Using behavioural insights to promote food waste recycling in urban households—Evidence from a longitudinal field experiment. Frontiers in Psychology, 9, Article 352. https://doi.org/10.3389/fpsyg.2018.00352
- van Gestel, L. C., Kroese, F. M., & de Ridder, D. T. D. (2017). Nudging at the checkout counter: A longitudinal study of the effect of a food repositioning nudge on healthy food choice. Psychology & Health, 33(6), 800–809. https://doi.org/10.1080/08870446.2017.1416116
- Lally, P., van Jaarsveld, C. H. M., Potts, H. W. W., & Wardle, J. (2010). How are habits formed: Modelling habit formation in the real world. European Journal of Social Psychology, 40(6), 998–1009. https://doi.org/10.1002/ejsp.674
- Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. Implementation Science, 6, 42. https://doi.org/10.1186/1748-5908-6-42
- Fogg, B. J. (2009). A behavior model for persuasive design. In Proceedings of the 4th International Conference on Persuasive Technology (pp. 1–7). ACM. https://doi.org/10.1145/1541948.1541999

- Marchiori, D. R., Adriaanse, M. A., & De Ridder, D. T. D. (2017). Unresolved questions in nudging research: Putting the psychology back in nudging. Social and Personality Psychology Compass, 11(1), e12297. https://doi.org/10.1111/spc3.12297
- Van Bavel, J. J., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., Crockett, M. J., Crum, A. J., Douglas, K. M., Druckman, J. N., Drury, J., Dube, O., Ellemers, N., Finkel, E. J., Fowler, J. H., Gelfand, M., Han, S., Haslam, S. A., Jetten, J., Kitayama, S., Mobbs, D., Napper, L. E., Packer, D. J., Pennycook, G., Peters, E., Petty, R. E., Rand, D. G., Reicher, S., Schnall, S., Shariff, A., Skitka, L. J., Smith, S. S., Sunstein, C. R., Tabri, N., Tucker, J. A., van der Linden, S., van Lange, P. A. M., Weeden, K. A., Wohl, M. J. A., Zaki, J., Zion, S. R., & Willer, R. (2020). Using social and behavioural science to support COVID-19 pandemic response. Nature Human Behaviour, 4(5), 460–471. https://doi.org/10.1038/s41562-020-0884-z
- Van Rookhuijzen, M., De Vet, E., & Adriaanse, M. A. (2021). The effects of nudges: One-shot only? Exploring the temporal spillover effects of a default nudge. Frontiers in Psychology, 12, Article 683262. https://doi.org/10.3389/fpsyg.2021.683262
- Saig, E., & Rosenfeld, N. (2023). Learning to suggest breaks: Sustainable optimization of long-term user engagement (arXiv:2211.13585v2). https://doi.org/10.48550/arXiv.2211.13585
- Kovacs, G., Wu, Z., & Bernstein, M. S. (2021). Not now, ask later: Users weaken their behavior change regimen over time, but expect to re-strengthen it imminently. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (Article 229, pp. 1–14). Association for Computing Machinery. https://doi.org/10.1145/3411764.3445695
- European Union. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the*

processing of personal data and on the free movement of such data (General Data Protection Regulation). *Official Journal of the European Union*, L119, 1–88.

- Duolingo Team. (2024, August 5). *Friend Streak: a new way to stay motivated together*. Duolingo Blog. https://blog.duolingo.com/friend-streak/?utm_source=chatgpt.com
- Duolingo, Inc. (2024, November 6). *Duolingo Q3 2024 press release / shareholder letter* [Press release]. Duolingo Investors. https://investors.duolingo.com/static-files/a503c122-ad7f-499e-91e3-d1f0d1211b17
- Thaler, R. H., & Benartzi, S. (2004). *Save More Tomorrow™: Using behavioral economics to increase employee saving.* Journal of Political Economy, 112(S1), S164–S187. https://doi.org/10.1086/380085