# Mobile Applications for Health and Wellness: A Systematic Review

ALAA ALSLAITY, Dalhousie University, Canada
BANUCHITRA SURULIRAJ, Dalhousie University, Canada
OLADAPO OYEBODE, Dalhousie University, Canada
JONATHON FOWLES, Acadia University, Canada
DARREN STEEVES, Dalhousie University, Canada
RITA ORJI, Dalhousie University, Canada

Mobile health (mHealth) apps show potential contributions as interactive systems for managing users' health conditions. They are also used to improve health habits using behaviour change strategies. However, the trends, effectiveness, and design practices of these apps in terms of behaviour change are unclear yet. With a collaboration between researchers, domain experts, interactive systems developers and professionals, this paper aims to fill this gap by systematically investigating 70 mHealth apps using two popular behaviour change frameworks, namely App Behaviour Change Scale (ABACUS) and the Persuasive System Design (PSD) model. The study investigates the most common strategies and how these strategies were designed and implemented in the apps to achieve the targeted design objectives. Furthermore, the study evaluates apps' behaviour change potential using the behaviour Change Score (BCS), a measure we introduced to evaluate how the apps employ behaviour change strategies. The results show that 1) Journaling is the most common category of apps. 2) the most employed strategies are Self-monitoring, Customize and Personalize, and Reminders. And 3) there is a positive correlation between apps' ranks (based on ratings and installation) and the BCS score of most strategies. Based on our findings, we offer recommendations for designing and developing mHealth apps and present opportunities for future work in this area.

CCS Concepts: • Ubiquitous and mobile computing - Ubiquitous and mobile computing systems and tools-Applied computing-Life and medical sciences - Consumer health.

**Additional Key Words and Phrases**: Health and wellness, mobile Apps, Behaviour change techniques, ABACUS, PSD.

**ACM Reference format:**
Alaa Alslaity, Banuchitra Suruliraj, Oladapo Oyebode, Jonathon Fowles, Darren Steeves, And Rita Orji. 2022. Mobile Applications for Health and Wellness: A Systematic Review. Proc. ACM Hum.-Comput. Interact, 6, EICS, Article 171 (June 2022), 29 pages, https://doi.org/10.1145/3534525



This work is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).
Author's addresses: A. Alslaity, Dalhousie University, Halifax, NS, Canada; B. Suruliraj, Dalhousie University, Halifax, NS, Canada; O. Oyebode, Dalhousie University, Halifax, NS, Canada; J. Fowles, Acadia University, Halifax, NS, Canada; D. Steeves, Dalhousie University, Halifax, NS, Canada; R. Orji, Dalhousie University, Halifax, NS, Canada. 2022. Magnetic Normal Modes of Bi-Component Permalloy Structures.


2573-0142/2022/06 – Article#171… $15.00
https://doi.org/10.1145/3534525

## 1. Introduction

Digital means (e.g., robots, smartphones, and chatbots) have become essential supporters of traditional health practices [1, 2]. Particularly, mobile health (mHealth) apps show potential contributions as interactive systems for supporting the practice of medicine and public health and managing individuals' health conditions. These apps have gained increased attention in the last two decades due to several reasons; mobile devices, such as smartphones, are equipped with advanced technologies and devices that allow them to collect various data such as physical activities, sleep patterns, physiological data (e.g., Heart rate, breathing rate) and psychological metrics [3], which are all helpful for health care purposes [4]. Also, the presence of

these advanced technologies supports the use of Artificial Intelligent (AI) to make mHealth more innovative and adaptive. Besides, the rapid rise of mobile phone penetration worldwide helps provide health interventions easily and cost-efficiently.

All the aforementioned advantages of mobile devices, in addition to users' passion for using these technologies, allow mobile apps to deliver various health services, such as behavioural health interventions. This, in turn, helps users achieve their health and wellness goals. Thus, mobile apps are gaining increasing familiarity in the domain of health and wellness. As a result, many mobile health apps have been developed and made available to the public. According to Levine et al. [5], mobile health apps are one of the most rapidly proliferated healthcare innovations. Over 300,000 mobile health apps were available in 2020, and more than half of smartphone users have installed a health app [5].

The increased attention to mobile health apps has resulted in a large but incohesive body of literature. For instance, few available apps have shown whether they achieve their goals of improving users' health [5]. Additionally, mobile apps are continuously being developed, and reviews should be done at a similar pace to keep track of the features deployed and theories used in these apps. Thus, the literature needs further investigation of mHealth apps to get a clearer picture of the usefulness of these apps and the effectiveness of behavioural change techniques used in the wild. Having such knowledge would, in turn, guide the development of future mobile apps for health and wellness.

This paper represents a step towards filling this need by systematically reviewing mobile apps related to behaviour change interventions for health and wellness. The objective is to investigate different types of mHealth apps and examine the engineering and operationalization of behaviour change strategies in these apps. To achieve this goal, we review 70 highly-ranked apps against 28 behaviour change strategies selected from two major behaviour change frameworks, namely, the App Behaviour Change Scale (ABACUS) [6] and the Persuasive System Design (PSD) model [7]. Only top-ranked apps (based on their ratings and number of installations) were considered. Furthermore, we evaluate apps' behaviour change potential using the Behaviour Change Score (BCS), a measure we introduced to evaluate how the apps employ behaviour change strategies. The review process was done collaboratively between researchers, domain experts, and professionals. The results show that: 1) the Android platform hosts more mHealth apps than Apple's iOS, 2) Journalling emerged as the most popular app category, 3) the number of behaviour change strategies employed in each app ranges between 5 and 26, with most apps implementing less than 14 strategies, 4) the most commonly operationalized strategies are Self-monitor Behaviour, Customize and Personalize Features, and Reminders, 5) the least operationalized strategies are Recognition, Distraction or Avoidance, and Willingness for Behaviour Change, and 6) apps' ranking is positively related to the BCS of most behaviour change strategies. Based on our findings, we offer suggestions for designing and developing interactive mobile apps for health and wellness and present opportunities for future work in this area.

This study is important, and it is different from previous works for several reasons; First, this work is the first to combine two well-known frameworks to produce a comprehensive review of apps that advance state of the art not only in academia but in the industry. Most existing work use either one framework or none and are not usually focused on apps used in the wild. Second, considering the rapid and continuous innovation in this area that leads to the frequent release of new mobile apps, it is necessary to investigate the latest state and trends in this domain frequently. Finally, in contrast to most existing theoretical- and research-oriented reviews, this review is practice-based in the sense that it explores the actual implementation of behaviour change strategies in apps available in the wild and are used by the public. Our work can bridge the gap between research and practice by providing a comprehensive review of the state of the art of mHealth applications in practice and highlighting opportunities for future work in the area.

The main contribution of this paper is providing an inclusive view of the state of mobile apps for health and wellness and identifying their potential for behaviour change. We offer in-depth insights beyond what exists in the previous work. This paper also provides suggestions and design recommendations for future mobile health apps.

The rest of this paper is organized as follows: Section 2 introduces the paper's main topics and discusses the related work. Section 3 presents the study design, while Section 4 shows the results. Then, the key findings and design recommendations are discussed in Section 5. Finally, Section 6 concludes the paper.

## 2. Background and Related Work

This section introduces the main concepts related to the paper's theme. Specifically, it discusses ABACUS and PSD frameworks. It also investigates related works.

### 2.1. App Behaviour Change Scale (ABACUS)

The App Behaviour Change Scale (ABACUS) is a reliable, theory-based scale primarily developed by McKay et al. to assess the behaviour change potential of smartphone apps [6]. It comprises 21 items categorized into four categories: Knowledge and information (five items), goals and planning (three items), feedback and monitoring (seven items), and actions (six items).

Table 1 summarizes the 21 items of the ABACUS scale. It is built based on the literature on health behaviour change interventions and has been widely adopted by other researchers. We used ABACUS in this study because the primary goal of ABACUS is to evaluate the potential behaviour change of smartphone apps (which is the main goal of our study). Also, it is built based on the literature on health behaviour change interventions [2]. Thus, we found this scale suitable for our study because it focuses on evaluating mobile apps (the main concern of ABACUS) in the health-related interventions (the domain based on which the scale was built). Besides, ABACUS is the first scale that can measure the potential for behaviour change in smartphone apps [6]. It has been shown to have good interrater reliability and is a valid tool for evaluating the potential behaviour change [5]. Furthermore, ABACUS has been widely adopted by other researchers in persuasive interventions for health and wellbeing, and it has been found very useful [8].

Table 1. App Behaviour Change Scale (ABACUS) strategies [6]

| Category | Item | Explanation |
|---|---|---|
| **Knowledge and Information** | Customize and personalize features | Elements of the app can be personalized through specific tools or functions that are specific to the individual using the app. |
| | Consistent with national guidelines or created with expertise | This would be found in the about section or generally in the app. |
| | Baseline information | This includes Body Mass Index (BMI[a]), weight, smoking rate, exercise, or drinking behaviours |
| | Instruction on how to perform the behaviour | The app is clear in telling the person how to perform a behaviour or preparatory behaviours, either verbally, through video, or in written form. |
| | Information about the consequences of continuing and/or discontinuing behaviour | The app gives the user information about the consequences of behaviour in general; this includes information about the relationship between the behaviour and its possible or likely consequences in the general case. |
| **Goals and Planning** | Willingness for behaviour change | Is there a feature during setup where you describe how ready you are for behaviour change? |
| | Goal setting | The person is encouraged to set a general goal that can be achieved by behavioural means. This includes subgoals or preparatory behaviours and/or specific contexts in which the behaviour will be performed. |
| | Review goals, update, and change when necessary | Involves a review or analysis of the extent to which previously set behavioural goals (regardless of short or long) were achieved. |
| **Feedback and Monitoring** | Understand the difference between current action and future goals | Allows users to see how they are tracking against a goal and the difference between what they want and what they are currently doing. |
| | Self-monitor behaviour | The app allows for regular monitoring of the activity. |
| | Share behaviours with others and/or allow for social comparison | The app allows the person to share their behaviours on social media or in forums. This could also include a buddy system or a leaderboard. |
| | User feedback (in person or automatically) | The app is able to provide the person with feedback, comments, or data about their own recorded behaviour. |
| | Export data | The app allows for exporting information and progress to an external user. |
| | Material or social reward or incentive | App provides rewards for attempts at achieving a behavioural goal. This might include efforts made toward achieving the behaviour or progress made in preparatory steps toward the behaviour or in achieving a goal. |
| | General encouragement | The app provides general encouragement and positive reinforcement on actions leading to the goal. |

*Table Continued.*

Table 1 (continued). App Behaviour Change Scale (ABACUS) strategies [6]

| Category | Item | Explanation |
|---|---|---|
| **Actions** | Reminders and/or prompts or cues for activity | The app prompts the user to engage in the activity. The app has the ability to give notifications or reminders to cue the behaviour. |
| | Encourage positive habit formation | The app prompts explicit rehearsal and repetition of the behaviour–not just tracking or logging. |
| | Practice or rehearsal, in addition to daily activities | The app does not have a lock on activities or a number that you cannot exceed daily. |
| | Opportunity to plan for barriers | The app encourages the person to think about potential barriers and identify ways of overcoming them. |
| | Restructuring the physical or social environment | The app prompts the person to alter the environment in ways so that it is more supportive of the target behaviour. |
| | Distraction or avoidance | The app gives suggestions and advice on how the person can avoid situations or distract themselves when trying to reach their goal. |

### 2.2. Persuasive System Design (PSD)

The Persuasive Systems Design (PSD) model [7] is a framework for analyzing, designing, and evaluating persuasive systems [9]. It suggests four categories of persuasive system principles, which are:

- *Primary Task*: this category compromises design features that support users in carrying out their primary tasks or the steps the users do in the application.
- *Dialogue*: this category focuses on the features related to the computer-human interaction to help users achieve their goals or target behaviour. This kind of interaction between systems and users is common in any interactive system, which provides system feedback to the users.
- *System Credibility*: the design features that belong to this category focus on increasing the system's credibility. That is, it provides a means to design more credible and persuasive systems.
- *Social Support*: the social support design features provide approaches to design the system that motivates users by leveraging the power of social influence.

Each of these categories compromises several persuasive features. Table 2 summarizes these features based on their categories, according to the PSD framework [7].

We also used the PSD in our study for several reasons: 1) it is one of the leading frameworks for persuasive and behaviour change design and evaluation. 2) The PSD is a comprehensive framework that combines many behaviour theories and frameworks to guide the design, evaluation, and analysis of behaviour change apps. 3) It also defines the content and software functionality of these systems. 4) PSD fits well for evaluating design specifications of behaviour change apps [10]. Therefore, integrating a framework for evaluating apps design specifications (using PSD) and their potential behaviour change (using ABACUS) allows for an in-depth and comprehensive assessment of the apps.

Table 2. Persuasive System Design Strategies

| **Category** | **Principle** | **Definition** |
|---|---|---|
| **Primary Task Support** | Reduction | A system that reduces complex behaviour into simple tasks helps users perform the target behaviour, and it may increase the benefit/cost ratio of a behaviour. |
| | Tunnelling | Using the system to guide users through a process or experience provides opportunities to persuade along the way. |
| | Tailoring | Guide users through a process or experience |
| | Personalization | A system that offers personalized content or services has a greater capability for persuasion. |
| | Self-monitoring | A system that keeps track of one's status supports the user in achieving goals. |
| | Simulation | Simulations can persuade by enabling users to observe the cause-effect link. |
| | Rehearsal | A system providing means to rehearse a behaviour can enable people to change their attitudes or behaviour in the real world. |
| **Dialogue Support** | Praise | By offering praise, a system can make users more open to persuasion. |
| | Rewards | Systems that reward target behaviours may have great persuasive powers. |
| | Reminders | If a system reminds users of their target behaviour, the users will more likely achieve their goals |
| | Suggestion | Systems offering fitting suggestions will have greater persuasive power |
| | Similarity | People are more readily persuaded through systems that remind them of themselves in some meaningful way. |
| | Liking | Visually attractive system is likely to be more persuasive. |
| | Social Role | If a system adopts a social role, users will more likely use it for persuasive purposes |
| **System credibility support** | Trustworthiness | A system that is viewed as trustworthy will have increased powers of persuasion. |
| | Expertise | A system that incorporates expertise will have increased powers of persuasion. |
| | Third-party endorsements | Third-party endorsements, especially from well-known and respected sources, boost perceptions of system credibility |
| | Surface Credibility | People make initial assessments of the system's credibility based on a firsthand inspection |
| | Real-world Feel | A system that highlights people or organizations behind its content or services will have more credibility. |
| | Authority | A system that leverages roles of authority will have enhanced powers of persuasion |
| | Verifiability | Credibility perceptions will be enhanced if a system makes it easy to verify the accuracy of site content via outside sources |

*Table Continued.*

Table 2 (continued). Persuasive System Design Strategies

| Category | Principle | Definition |
|---|---|---|
| **Social Support** | Social learning | A person will be more motivated to perform a target behaviour if (s)he can use a system to observe others performing the behaviour. |
| | Normative influence | A system can leverage normative influence or peer pressure to increase the likelihood that a person will adopt a target behaviour |
| | Social facilitation | System users are more likely to perform target behaviour if they discern via the system that others are performing the behaviour along with them |
| | Cooperation | A system can motivate users to adopt a target attitude or behaviour by leveraging human beings' natural drive to cooperate. |
| | Recognition | By offering public recognition for an individual or group, a system can increase the likelihood that a person/group will adopt a target behaviour. |
| | Social Comparison | System users will have a greater motivation to perform the target behaviour if they can compare their performance with the performance of others. |
| | Competition | A system can motivate users to adopt a target attitude or behaviour by leveraging human beings' natural drive to compete. |

### 2.3. Related Work

Mobile apps have been used in several health-related domains and have been extensively studied from different perspectives. For instance, Calisto et al. [11] emphasiz the design of a multimodality medical image user interface for breast screening. Another study [12] investigated the interaction perspective based on AI techniques in medical imaging, focusing on workflow efficiency and quality, preventing errors and variability of diagnosis in Breast Cancer. Calisto et al. [13] focus on providing easy-to-use interfaces that help radiologists perform a reliable visual inspection. Several other studies have been conducted for a particular health domain, such as weight control and physical activities [14, 15], nutrition monitoring and healthy food [16], management of chronic diseases [17, 18], and Mental health [19]. Due to the multidimensional nature of the mHealth domain, we cannot comprehend the whole literature in a single study. Therefore, this section focuses on presenting the most related work. Specifically, it investigates previous literature studies (or reviews) on the effectiveness of behaviour change mobile apps for health and wellness.

Several research was conducted to evaluate and investigate existing mHealth apps. For instance, Mckay et al. [2] reviewed mobile apps available on the Australian Apple iTune and Google Play Store. The study included apps in five domains, including smoking, alcohol use, physical activity, nutrition, and mental health. The selected apps were evaluated against their functionality and potential behaviour change. The results of evaluating 344 apps indicated that the ability of an app to encourage practice or rehearsal and the user's ability to easily self-monitor behaviour were the most and second most commonly identified feature across all apps.

Mathews et al. [20] surveyed mobile apps for health behaviour change. The paper aimed to identify the emphasis, gaps and commonly present design features in mobile apps for motivating physical activity. The PSD framework was used to evaluate the persuasive strategies used in the evaluated apps. The study concluded that system credibility support is the least considered category as persuasive systems feature in mHealth apps for promoting physical activity. The other three categories: primary task support, dialogue support, and social support, were moderately represented in the selected articles. Alhasani et al. [21] also used the PSD framework to review persuasive strategies implementation in mHealth apps for stress management. The study evaluated 60 apps against the PSD framework. The study aimed to uncover the persuasive strategies employed in these apps and how they were implemented. The results show that Personalization and self-monitoring are the most and second most commonly employed strategies, respectively.

Ko et al. [22] reviewed mHealth literature and apps for sleep-related purposes. The search included app stores, such as Apple App Store, Google Play, and Microsoft Windows Store. The results showed that using digital intervention changes the landscape of sleep health and clinical sleep medicine. Particularly, mHealth interventions can improve and impair collective and individual sleep health. This study focuses on the functionality and features of the apps, but it neglects the behaviour change potential of the considered apps.

To explore the behaviour change features of mobile apps, Payne et al. [23] provided a review of the literature on mobile apps used in health behavior interventions. The paper described the behavioral features and the focus of the considered apps and evaluated these apps regarding their potential to disseminate health behavior interventions. The review of 24 studies revealed that self-monitoring is the most common strategy, followed by cues to action and feedback. It also concluded that "mobile apps may be considered a feasible and acceptable means of administering health interventions, but more studies and more rigorous research and evaluations are needed to determine efficacy and establish evidence for best practices." Roffarello and Russis [24] reviewed the features of 42 digital wellbeing apps retrieved by searching the Google Play store. Based on the results of the review and thematic analysis, the authors concluded that "digital wellbeing apps are appreciated and useful for some specific situations. However, these apps do not promote the formation of new habits, and they are perceived as not restrictive enough". This paper is limited to the "smartphone addiction" domain.

Wright et al. [25] examined the effectiveness of mHealth apps in changing health behaviours and clinical health outcomes. The authors used databases (e.g., CHINAHL, EMBASE, and PubMed) to search articles on health behavior changes using mobile apps. The study included peer-reviewed articles published in journals between January 2000 and May 2017. A total of 20 articles were included, and most of them reported that mHealth apps positively impact health-related behaviours. A more recent study [26] also studied the effectiveness of mobile apps in changing health behaviors based on a literature review of peer-reviewed papers published between 2014 and 2019. The study examined the inclusion of behavior change techniques in mHealth apps focusing on four health domains: Physical Activity, Diet, Drug and Alcohol Use, and Mental Health. The results show that most of the studies were done in the Physical Activity or Diet domains, and participants perceived mHealth apps positively. However, There was no strong evidence of the effectiveness of mHealth apps in promoting the desired health behaviours because few studies found significant differences between the app and control groups.

Some research focuses on the measures and evaluation approaches of mobile apps in general and mHealth apps in particular. Powell et al. [27] considered 22 measures that can be used to evaluate mHealth apps. The measures were identified based on The Anxiety and Depression Association of America (ADAA) and PsyberGuide, which are two websites that rate mobile apps. The study considered 20 apps distributed equally between two domains: smoking cessation and depression. The results showed wide differences in the performance of measures between different domains, and clinical quality measures such as effectiveness, ease of use, and the performance had relatively poor reliability. Another study [28] investigated evaluation approaches of mobile health apps. The result of reviewing 38 papers revealed no single best practice method for evaluating mobile health and well-being apps. The study indicated that most evaluation approaches lack sufficient information or evaluation. Therefore, the study suggested that the evaluation of apps should include a review of the functionality and an assessment of the potential behavior change of the app. Based on this review study, the authors developed a scale specifically designed to determine the behavior change potential of smartphone apps, which is known as the App Behavior Change Scale [6].

Some papers concern the general use of smartphones as ubiquitous systems for the health domain rather than focusing on the available mHealth apps for behaviour change. For instance, Trifan et al. [4] conducted a study that aims to identify recent publications on the use of smartphones as ubiquitous systems for health monitoring. Specifically, it investigates the existing literature about the sensing technologies used to monitor the health status of smartphones users. 118 papers were considered in the study. Results show that mental health and physical activity are the most common domains. Batra et al. [29] systematically reviewed published literature concerning technologies for serious mental illness. The study aims to identify available health technologies and their intended uses for patients with mental health issues. Helga et al. [30] reviewed 35 papers published in the domain of self-assessed stress using smartphone apps. The goal of the study is to investigate the use of smartphones to measure self-assessed stress in healthy adult individuals. Kao et al. [31] propose a collaborative filtering approach to predict users' health conditions based on similar users’ conditions.

#### *2.6.1. Why is this study important?*

As presented above, there are a number of research and literature surveys concerning the use of smartphones in the health domain. Despite this number of studies, there is a continuous need for new research to cover the proliferated increase in mHealth apps. Our review of the related work revealed several limitations in the existed literature, which can be summarized as follows:

- We noticed that most of the studies investigated the domain based on the academic literature (e.g., [30], [22], [23], [25], and [20]) rather than reviewing the apps in the real world. Also, only a few published academic studies explicitly reported the behavioural change technologies incorporated into the apps [32]. Thus, investigating apps under actual use is essential to bridge the gap between academia and practitioners. Examining mobile apps based on real users’ experiences provides more accurate opinions, which leads to more informing insights.
- Most published literature on mHealth apps is descriptive [23]. These studies (e.g., [2], [20], and [23]) emphasize the gaps in the literature, employed strategies, and other descriptive statistics, but they neglect to evaluate the effectiveness of mHealth apps in promoting behaviour change.
- Most of the existed reviews are limited in terms of domain (e.g., [24], [20], [21], [22], and [27]), audience groups (e.g., [2]), or both (e.g., [33]). This limitation was confirmed by a previous study [32], which also called for the need for new studies to explore the mHealth domain more comprehensively.
- Most of the previous reviews are not theory-based. That is, they do not rely on well-defined theories or frameworks to conduct their studies. Examples of these studies include ([27], [24], [22], [23], [26], and [30]). Few studies have considered well-known frameworks. However, even these studies were limited to a single framework (e.g., [2], [20], and [21]). Our work is the first to combine two well-known frameworks for a comprehensive evaluation of apps existing in the wild.

- The available systemic reviews were conducted before 2020 (e.g., [2], [27], [20], [23], [22], and [25]). Newer research, such as [26] and [21], are based on academic research literature, and they are limited in terms of considered domain. Considering the rapid and continuous production of mobile apps, it is necessary to continuously investigate the latest state and trends in this domain.

Given these limitations of existing work, we can say that there is a need for new studies that explore the latest state of the art of mHealth apps for behaviour change. The literature witnessed several calls for more research and evaluations to determine efficacy and establish evidence for best practices in this domain [23]. According to Milne-Ives et al. [32], “Further evaluation of the effectiveness of mobile health apps is needed to determine which apps are most useful and which behavioral change theories and techniques best promote positive behavior change, which in turn can guide future development.” Also, Mckay et al., [2] stated, “the strength or potential of these apps to lead to behavior change remains uncertain”. New apps and innovative technologies are continuously being introduced and developed. Thus, reviews could keep pace to match this rapid development and investigate trends in the features and effectiveness of mHealth apps [32]. To fill this need, this paper builds on previous reviews to update and expand previous systematic reviews.

## 3. Study Design

The review process was done as a collaboration between researchers, domain experts, and professionals. Particularly, the review went through three stages. In the first stage, application keywords were collected by our domain experts and professionals team. These keywords were used to build the search queries. In the second stage, the team selected 28 behaviour change strategies from ABACUS and PSD frameworks. Finally, the search strategy was applied, and the selected applications were installed and evaluated.

### 3.1. Inclusion and Exclusion criteria

This study focuses on the current state of mobile health app technology. Only apps in the English language were considered. There was no restriction on the considered populations in terms of demographic information. That is, the study considers apps that target any population regardless of their age, gender, country, or ethnicity. However, each application has at least five installations, they should be designed for at least one of the Android or iOS platforms, and they aim to achieve any health and wellness goals. Due to the high number of available apps and to build our study on stable ground, only highly-rated apps were considered. That is, all the considered apps have a rating of more than 3 out of 5. It is worth mentioning that the ratings of an application in app stores show the average rating across all app versions. We considered the average rating because it is the most common criterion for apps selection. Also, the goal of our study is to evaluate the top apps available on Android and iOS. So, we believe that considering ratings and the number of downloads is a rationale approach to achieving our goals.

In addition, since our work aims to study the app with the most recent designing methods, tools, and techniques, we focused our search on apps that deploy AI techniques. This condition helps limit our search to the apps that use the most recent technologies. The considered apps use AI-based solutions for diverse purposes, including mood tracking, journaling, virtual chat, and other implementations. However, we will not discuss this aspect in detail because it is out of the scope of this paper; This paper concerns the behaviour change strategies and their implementation rather than the technical details of the apps.

### 3.2. Sample Selection

The apps were selected by searching major app stores and search engines, including Google Play Store and Apple's App Store. Before starting our search process, experts from the health domain identified the most relevant search terms. Based on this identification, the search query included a combination of the following keywords "Journal", "Life coach", "CBT or Cognitive Behaviour Therapy", "Goal tracking", "Habit tracking", "Personal growth", "Resilience", and "Wellness". Figure 1 depicts the logical relation between these search terms. Our domain experts selected these search terms through discussion sessions and with reference to existing literature.

```
("Journal" OR "Life coach" OR "CBT" OR "Cognitive Behaviour Therapy"
OR "Goal tracking" OR "Habit tracking" OR "Personal growth" OR
"Resilience" OR "Wellness")
```

Figure 1. Search query

The search in both app stores resulted in 988 apps. We carefully filtered these apps against the inclusion/exclusion criteria based on their descriptions in the app stores. After this step, we included 359 apps. Then, these apps were ranked based on their ratings and number of installations (i.e., by multiplying the number of installs by rating). It is worth noting that the word "rank" in this paper refers to this ranking. The top 70 ranked apps were reviewed using the behaviour change review scales (more details in Section 3.3). Figure 2 depicts the selection and filtering process.

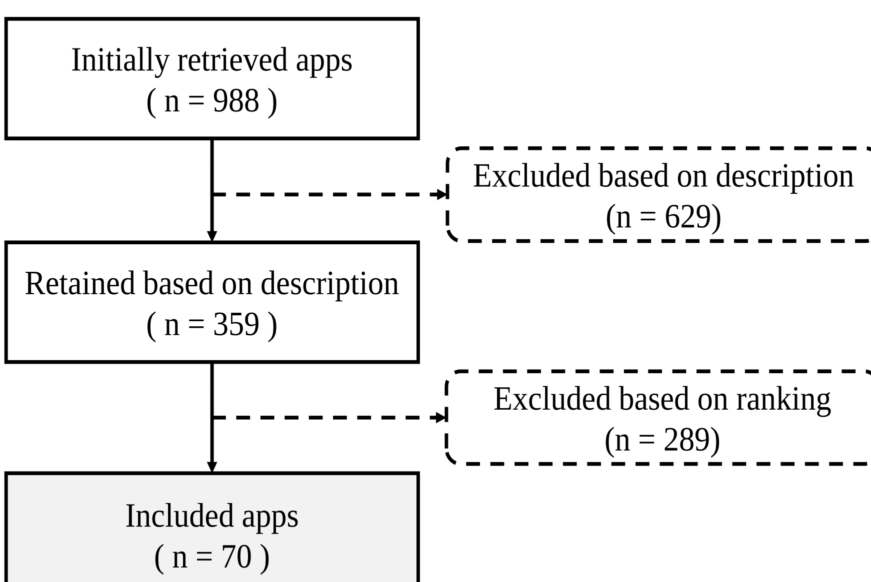


Figure 2. Selection process

All the considered apps can be installed and used for free, so we installed and used all the apps. However, some features are only available for premium accounts (i.e., paid accounts). To investigate these features, we relied on several resources, including the descriptions of the apps in the app stores, the apps' websites where these paid features are advertised with screenshots or video demos, and users' reviews and comments. It is also worth mentioning that many mHealth apps are available, and no single study can comprehensively review them all. The 70 apps selected in our study were deemed representative and a good number for an in-depth review. Most previous studies considered less than 60 apps, and several others considered less than 25. So, we believe that 70 apps are sufficient for our study. Also, app stores involve apps ranging from authentic and theory-driven apps to small apps developed for practicing without theoretical background or health expertise. To make our results more reliable, we limited our search to the highly ranked apps and apps with a high number of installations. High ratings along with a high number of installations indicate that the app is popular and generally liked by the users.

### 3.3. Behaviour Change Review Scale

As mentioned before, we used well-established frameworks ABACUS and PSD to analyze and investigate the behaviour change potential of the apps. Based on these two frameworks and with the help of behaviour change researchers and practitioners, and health practitioners, our team formulated our behavioural review scale by merging both frameworks to enable comprehensive review beyond what is obtainable in the existing literature. Specifically, the team initially selected a total of 37 strategies that are commonly used in the health and wellness domain. These 37 strategies involve 21 items from the ABACUS model and 16 strategies from the PSD model. Through several meetings and brainstorming, persuasive technology and behaviour change experts selected these strategies based on their applicability and popularity in the health and wellness domain. Upon further review of both frameworks, we found that the two frameworks share some items in common. Specifically, we could match nine (9) strategies from the ABACUS to the PSD scales. As a result, after merging duplicate items from both frameworks, we identified 28 distinct items; nine items were common between the two scales, 12 items were distinct from ABACUS, and seven items were distinct from the PSD scale. These 28 distinct items were used for the app review. Table 3 depicts these items where the first column shows the items selected from the ABACUS framework and the second column shows the items selected from the PSD framework.

## 4. Results

This section discusses the analysis results based on the themes considered in this study. Specifically, the section considers the following themes: platforms (Android and iOS), apps categories, popularity (or the number of installations), and the behaviour change potential of the considered apps.

### 4.1. Platform

All the apps included in this study are highly rated, such that the lowest rating included was 3 out of 5. The average rating for all apps in the Google play store is 4.5, and Apple's app store is 4.6. Figure 3 shows the distribution of the apps between the most common smartphone operating systems, namely iOS and Android. It shows that more applications are available for Android systems. Particularly, all the considered apps (n=70) are available for Android systems, while only 44 are available for the Apple iOS system, which means (37%) of the applications do not have an iOS version.

Table 3. Selected items of the Behavioural Review Scale

| # | **ABACUS Items** | **PSD Items** |
|---|---|---|
| 1 | Customize and personalize features | Personalization |
| 2 | Consistent with national guidelines or created with expertise | Expertise |
| 3 | Instruction on how to perform the behaviour | Tunnelling |
| 4 | Self-monitor behaviour | Self-monitoring |
| 5 | Share behaviours with others and/or allow for social comparison | Social facilitator |
| 6 | Material or social reward or incentive | Rewards |
| 7 | General encouragement | Praise |
| 8 | Reminders and/or prompts or cues for activity | Reminders |
| 9 | Encourage positive habit formation | Rehearsal |
| 10 | Baseline information | -- |
| 11 | Information about the consequences of continuing and/or discontinuing behaviour | -- |
| 12 | Willingness for behaviour change | -- |
| 13 | Goal setting | -- |
| 14 | Review goals, update, and change when necessary | -- |
| 15 | Understand the difference between current action and future goals | -- |
| 16 | User feedback (in person or automatically) | -- |
| 17 | Export data | -- |
| 18 | Practice or rehearsal, in addition to daily activities | -- |
| 19 | Opportunity to plan for barriers | -- |
| 20 | Restructuring the physical or social environment | -- |
| 21 | Distraction or avoidance | -- |
| 22 | -- | Reduction |
| 23 | -- | Suggestion |
| 24 | -- | Trustworthiness |
| 25 | -- | Cooperation |
| 26 | -- | Recognition |
| 27 | -- | Normative Influence |
| 28 | -- | Third-party Endorsements |

The popularity of the Android phones itself could explain the popularity of the apps available for it; Android systems are more common compared to iOS because of several reasons: first, Android is the operating system used by a wide range of smartphone makers. In contrast, iOS is limited to Apple's phones. Second, Android devices are available in all price ranges, making them affordable for more users. Third, Android is compatible with more smart devices compared to iOS. Finally, iOS requires developers to have a paid account in order to release an iOS app, whereas such accounts are free for Android developers. These reasons make Android-based devices more accessible to a broader range of users, which, therefore, attract companies and investors to develop apps for Android devices.

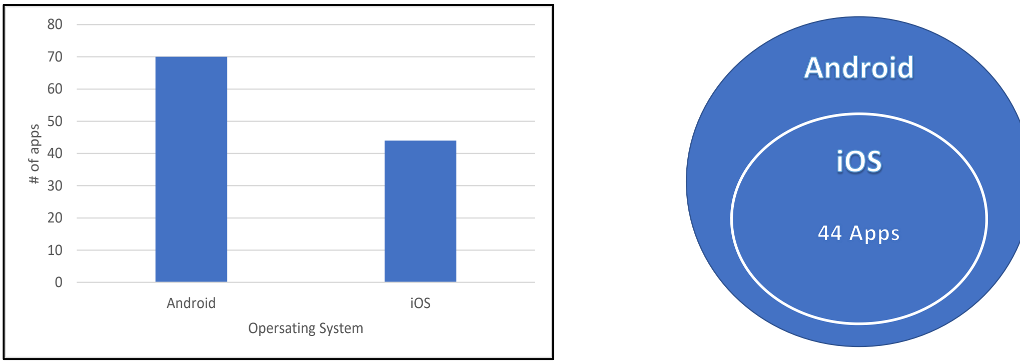


Figure 3. Apps distribution based on Operating Systems.

### 4.2. Applications Categories

Based on our analysis of the app features and target purpose, we classified them into different categories. For instance, apps that concern sleep issues are categorized under the "*Sleep*" category, while apps that focus on mental health are grouped under the "*Mental Health*" category. We classified the apps into 19 categories, as depicted in Figure 4. The classification was done based on these apps' descriptions of their purpose and use. It is worth mentioning that some apps are focused on multiple health and wellness goals. These apps were classified under the "*Health and Wellbeing*" category. The results show that *Journaling* (which involves apps based on users' journaling where they write their feelings, thoughts, emotions, and other life events) is the most common category. Twenty-five apps (36%) of the reviewed apps are *Journaling* apps. The popularity of the *Journaling* app is not surprising, considering that journaling has been shown to be successful in several domains [34], and it can be easily implemented. The *habit Tracking* category is found as the second most common (n=12). This category involves apps that help users take control of their habits by setting goals, maintaining good habits and tackling bad habits. Habit tracking is more automated and quantitative compared to Journaling. The third most common category is *Mental Health* (n = 5), followed by *Routine Task* tracking (n = 4) and *Meditation* (n = 4) apps.

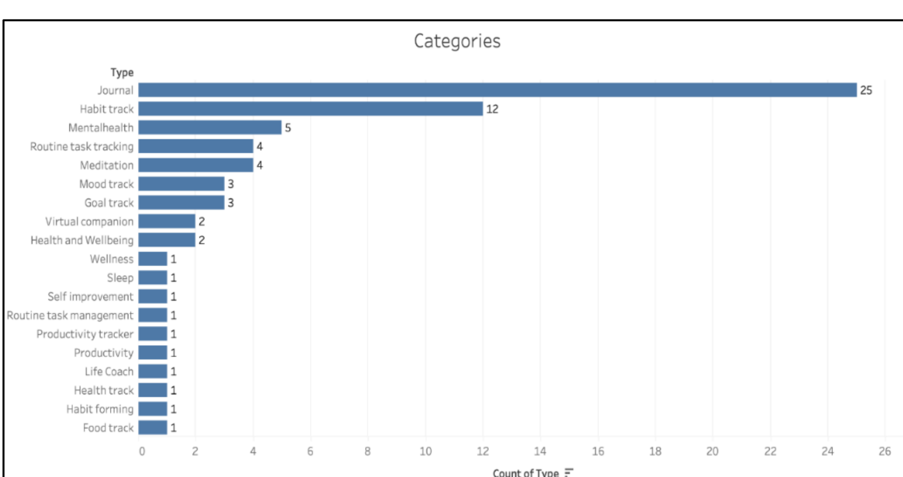


Figure 4. Number of apps based on categories

### 4.3. Installations

This section discusses the number of installations for each application category. The number of installations is considered an indicator of apps popularity of apps because installing an app implies an intention to use it. According to Mardiana et al. (2015), most studies for validating the Technology Acceptance Model (TAM) have demonstrated that the *intention* to use is the antecedent of *actual* use of the system. Therefore, we investigated the number of installations as an indicator of popularity that may also predict application acceptance and perhaps success.

The total installs are calculated as the total number of installs for each application category, whether on Android or iOS. That is, if an app has both iOS and Android versions, the total installs is the sum of the iOS and Android installs. Figure 5 shows the total number of installations for each category. As Figure 5 shows, *Journaling* apps have the highest number of installations (about 31 million), with the mean equals 1.24 million and the median being 1 million installs. *Meditation* comes next with a total install of 11.7 million, the mean is 2.9 million, and the median is 764 thousand installs. *Health and Wellbeing* category is the third most installed, with around 11 million installations. On the other side, *Food track*, *Life coach*, and *Wellness* applications are the least installed applications, with just above 100 thousand installations for each category.

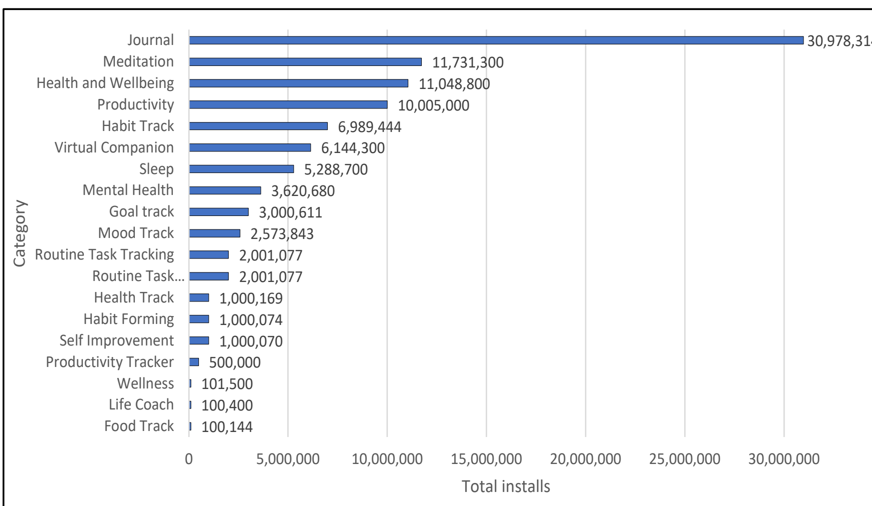


Figure 5. Total number of installations

Figure 6 depicts the average number of installations for each app category. The figure shows that *Productivity* has the highest average installations, followed by *Health and Wellbeing* and *Sleep* categories. On the other side, *Food Track*, *Life Coach*, and *Wellness* have the least average number of downloads. An important observation in this context is that the *Journal* category is the sixth in terms of average installations, although it has the highest number of installations. Similarly, *Meditation* (the second most installed app category) is the fifth in average installations. On the other hand, *Productivity* has the highest average number of installations, although it emerged as the fourth most installed category. The Health and Wellbeing category emerged as the third most installed category, and it has the second-highest average of installations. These results show that total installs, although it indicates the application's popularity and potential acceptance, should not be considered in isolation from other measures.

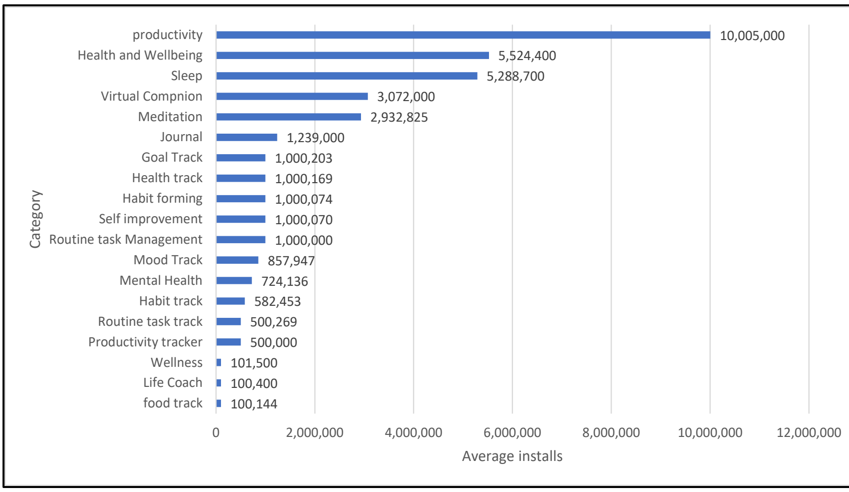


Figure 6. The average number of installations per category

Figure 7 and Figure 8 depict the average number of installations per app category for Android and iOS, respectively. The results show that average installations in Android are in the millions, whereas iOS is only seen in thousands. In particular, the average number of installations for Android and iOS are (1,390,000) and (20,100), respectively. This huge difference in the number of installations can be attributed to the popularity differences between both systems, as discussed in Section 4.1. Figure 7 shows that the Productivity category has the highest average number of downloads on Android, whereas, in iOS applications, the *Sleep* category has the highest average number of installations. Figure 7 also shows that Android apps with the highest average downloads belong to *Productivity*, *Health and Wellbeing, and Sleep* categories. In the iOS system, apps that belong to the *Sleep*, *Virtual Companion*, and *Meditation* categories have the highest average downloads.

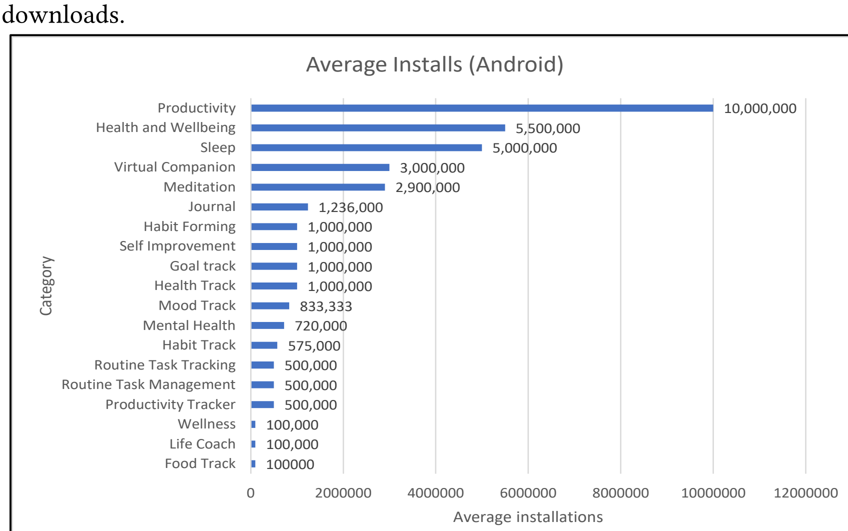


Figure 7. The average number of installations for Android

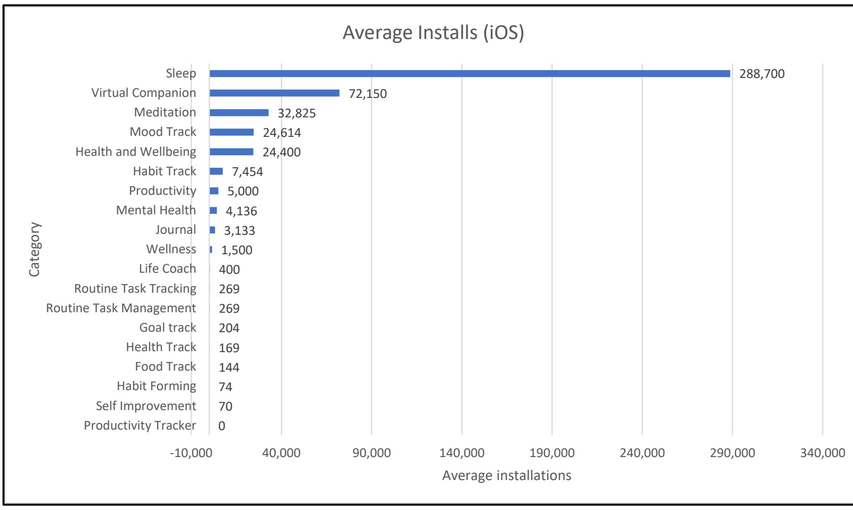


Figure 8. The average number of installations for iOS

## 4.4. Application Domain

This section studies the distribution of the behaviour change strategies over different domains or application areas. Based on the classification of the app stores, the apps included in this study belong to seven different categories: Health and Fitness, Lifestyle, Medical, Parenting, Productivity, Social, and Video_Players. Figure 9 depicts the number of apps that operationalize each strategy. It also shows the distribution of these apps over the seven domains.

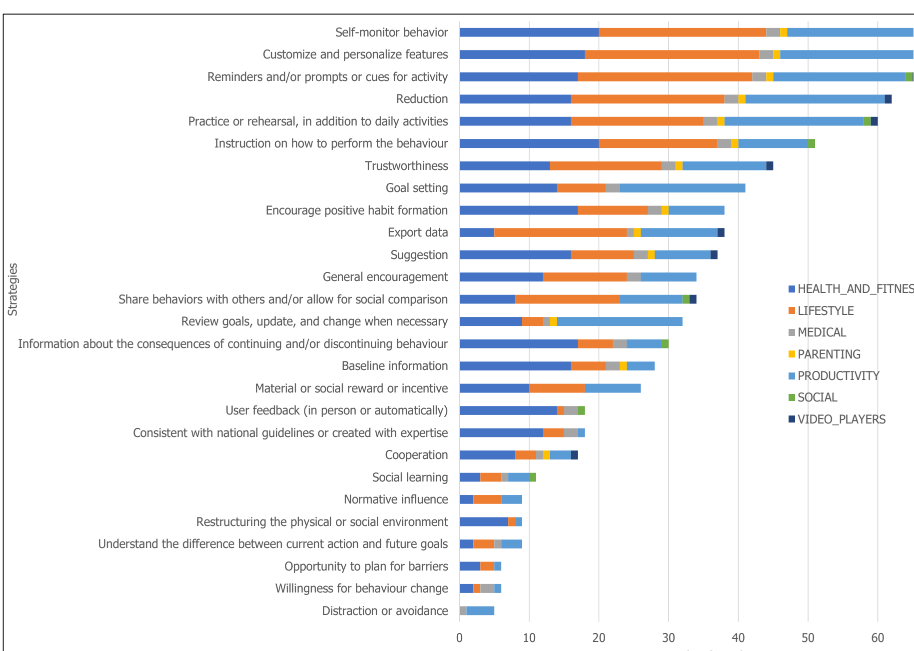


Figure 9. Distribution of Behaviour Change Strategies over domains

Our results show that the most implemented strategies are Self-monitor behaviour, Customize and Personalize, and Reminders and Prompts, implemented in 68, 67, and 66 apps, respectively. As the figure shows, most apps belong to three domains: Life Style, Health and Fitness, and Productivity. The most operationalized strategies in the Health and Fitness domain are *Self-monitor Behaviour*, and *Instruction on How to Perform the Behaviour*. Whereas in Lifestyle domains, *Self-monitor Behaviour*, and *Reminders* are the most common strategies. *Practice or rehearsal*, *Reduction*, and *Customize and Personalize* emerged as the dominant strategies in the Productivity domain. Regarding the other four categories (Medical, Parenting, Social, and Video Players), strategies are almost evenly distributed. All strategies are implemented 1 or 2 times, or they are not operationalized in that domain based on the apps included in the study.

### 4.5. Behaviour Change Potential

To evaluate the extent to which the 28 behaviour change strategies were implemented in the corresponding apps, the apps were assessed against these strategies on a four-point Likert scale. This Likert scale ranges from zero to three, where zero represents None (i.e., the app did not implement the corresponding strategy), while one, two, and three indicate Low, Medium, and High implementation, respectively. That is, if the app did not involve the strategy in its implementation, it would get a zero rating for that strategy. Otherwise, it will get a rating of more than zero depending on to what extent it implements that strategy. Specifically, the rating is given to the app based on the number of different implementations of the same behaviour change strategy. If the app uses one implementation of a strategy, it was given the rating (1), but if it uses two implementations, it was given the rating (2). Apps were given the highest rating (3) if the strategy was implemented in three or more different ways. For example, for the *Customize and Personalize Features* strategy, an app that only has one option to customize notification settings (e.g., time to notify or the notification sound) was given a rating of (1). On the other hand, if the app offers users three or more customization controls (e.g., customize the app's look and feel and app content, in addition to notifications preferences), it was given a rating of 3.

We evaluated all 28 behaviour change strategies for each app and rated each strategy according to the aforementioned Likert scale. Then, these ratings were used to calculate the overall score, which we called the *Behaviour Change Score (BCS)*. This score is calculated either for strategies ($BCS_{str}$) or apps ($BCS_{app}$). For each strategy (X), the $BCS_{str}$ is the summation of rating values given to X in all apps. On the other hand, for each app (Y), the $BCS_{app}$ is the summation of rating values given to all strategies corresponding to Y. That is, $BCS_{app}(Y)$ is the summation of scores for all 28 strategies in the Y. It is worth mentioning that we identified this score based on previous studies, such as [32], that also associated the effectiveness of behaviour change apps to the number of persuasive strategies implemented.

The process of evaluating the apps against persuasive strategies was done by the authors, and it involved three main stages: first, based on a literature review of the most common strategies in mHealth apps, our behaviour change researchers formulated the behavioural review scale (as mentioned in Section 3.3). Second, two experts in behaviour change mobile apps evaluated the apps against the 28 strategies. Then both researchers met and discussed their evaluation results. Finally, other experts who were not part of the review and scoring were consulted for any doubts that might appear in the previous stage. That is, if the experts who evaluated the apps had any concerns, they discussed it through discussion sessions with other experts and reached an agreement.

#### 4.5.1. Strategies-Based Analysis

As mentioned above, to evaluate the operationalization of each behaviour change strategy, we calculated BCS for each strategy. We called this value the ($BCS_{str}$), and we calculated it as the summation of all scores for the corresponding strategy in the 70 apps. Figure 10 depicts the ($BCS_{str}$) for all 28 strategies considered in this study. The mean value of overall $BCS_{str}$ is found to be (61.78). The results indicate that *Self-monitoring* was the highest employed strategy, with a score of 153. This was not surprising as Self-monitoring is an essential strategy in interventions for health and wellness [35]. Also, previous studies have shown that *Self-monitoring* is among the most commonly used strategies [2, 21]. *Customize and Personalize Features* are the second with a score of 141, Followed by *Reminders, Reduction,* and *Practice or Rehearsal* with scores of 138, 128, and 117, respectively. On the lower end, *Recognition* strategy was not used in any of the apps, *Distraction and Avoidance* received a score of 8, *Opportunity to Plan for Barriers* scored 9, and *Willingness for Behaviour Change* got a score of 12. These statistics provide a picture of highly employed and least employed strategies in health and wellbeing apps.

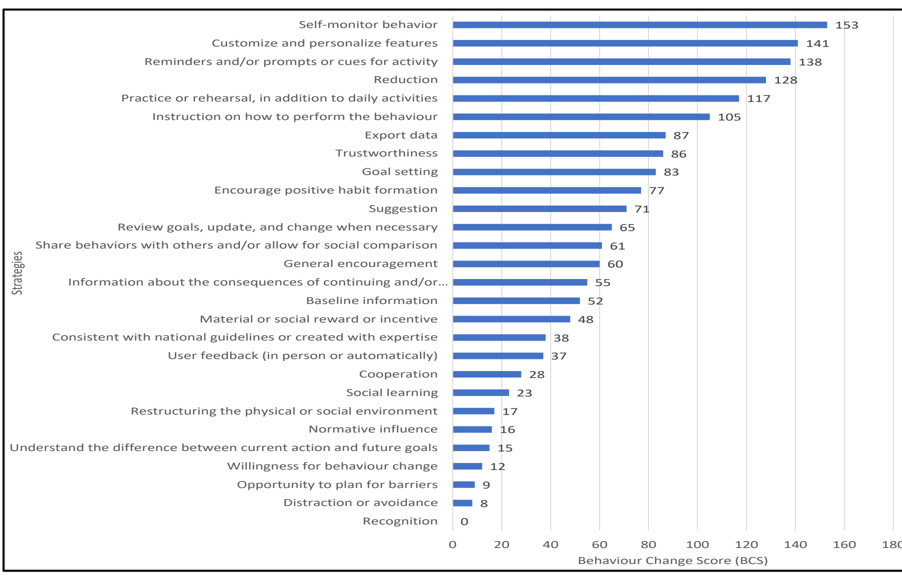


Figure 10. Behaviour changes score (28 behaviour change strategies)

#### 4.5.2. App-Based Analysis

In regards to the app-based behaviour change score (or $BCS_{app}$), we calculated it for each app as the summation of scores for all 28 strategies in the corresponding app. The histogram presented in Figure 11 shows the distribution of the apps based on their BCS. The grouping bins of the histogram show the distribution of the behaviour change score of all 70 apps. They show a normal distribution of the scores, where the mean is 24.7, and the median and mode are 24. The figure shows that the peak is between 22 and 27, where 17 out of the 70 apps fall in the peak area. It also shows that the $BCS_{app}$ spread from 7 to 52. Three apps were found in the lowest cluster (7 to 12), and two apps were found on the other end of the histogram (47 to 52). This widespread indicates that some apps operationalize a limited number of behaviour change strategies compared to others. Also, we notice that the figure is slightly right-skewed, which means that more apps belong to the figure's left side (i.e., lower BCS). This distribution shows that some apps operationalize fewer behaviour change strategies than other apps. Based on these results and given that all apps considered in this study are highly rated, we can say that mobile health apps gain high users' acceptance and ratings, even if they deploy a limited number of behaviour change strategies.

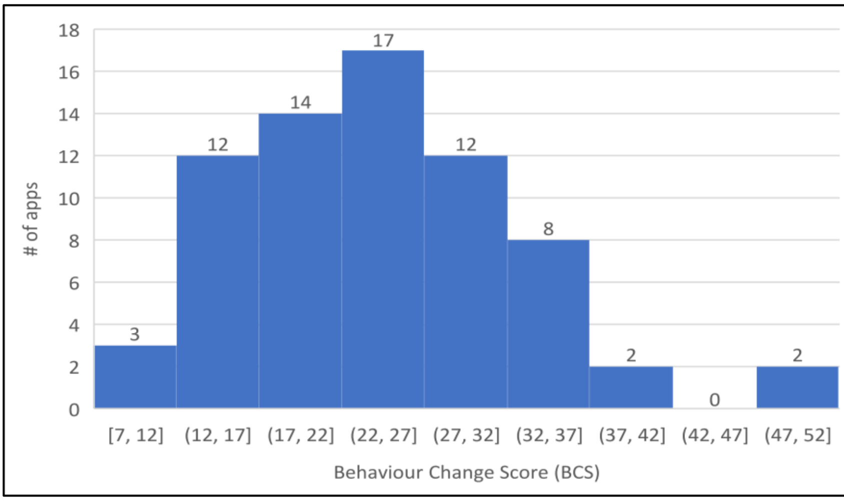


Figure 11. Distribution of apps based on BCS

Figure 12 depicts the average BCS for each app category. As the figure indicates, *Health and Wellbeing* category has the highest behaviour change score (*BCS* = 43). *Habit Forming* come next (*BCS* = 38), followed by *Mental Health* (*BCS* = 33.6) and *Food track* (*BCS*=33). On the other side, Health Track has the lowest behaviour change score (BCS=18), while *Productivity Track* and *Journal* have the second and third least scores (BCS = 19 and 19.1), respectively.

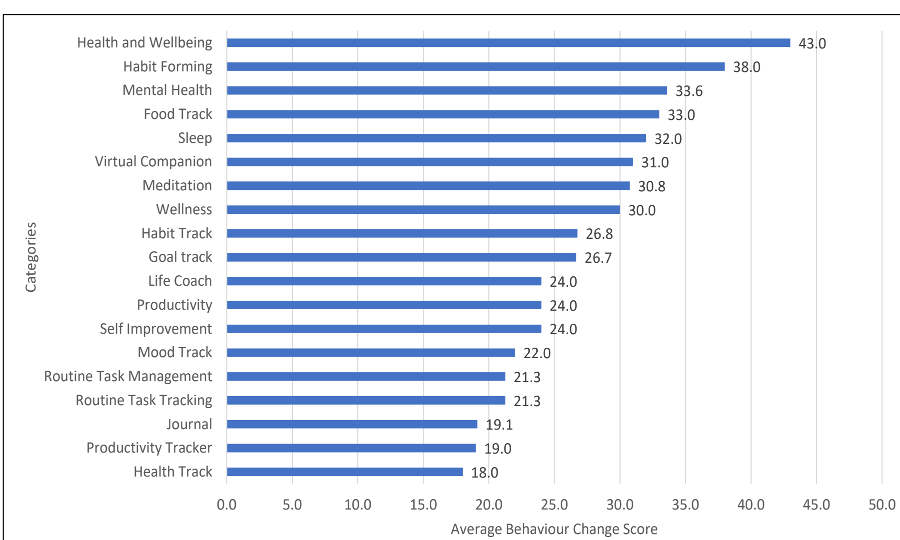


Figure 12. Average Behaviour Change Score for each category

Regarding the number of operationalized strategies, we found that the apps implement 12 strategies on average, and median and mode are both (12). The histogram presented in Figure 13 depicts the statistics about the number of behaviour change strategies implemented in the apps. It shows that the minimum number of strategies is 5, and the maximum number is 26. The figure shows that 19 out of the 70 apps fall in the peak area (between 11 to 14 strategies). It also shows a right skewness, which indicates that more apps implement fewer strategies. Particularly, most apps implement a low to a moderate number of strategies. (49 apps implement less than 14 strategies, while two apps implement more than twenty). This wide spread indicates that some apps operationalize a limited number of behaviour change strategies compared to others.

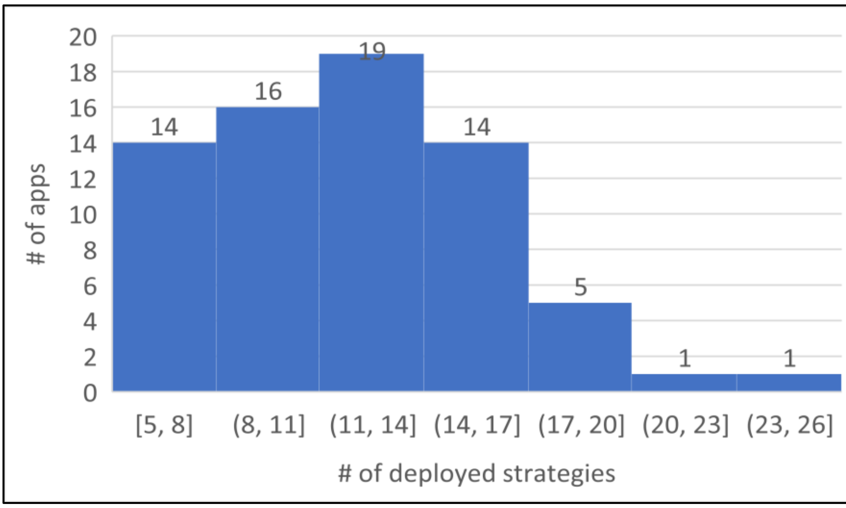


Figure 13. Distribution of apps based on deployed strategies

To get more insights on the top apps in terms of BCS, we analyzed the data about the highest 20 apps (i.e., apps with the highest BCS scores). All these apps except one have versions for both systems (Android and iOS). We noticed that the average rating for these apps is very high; (4.56) in the Google play store and (4.67) in the Apple store. Regarding the deployed behavioural change strategies, Figure 14 depicts the popularity of each strategy in the top 20 apps. It shows that all 20 apps deployed the following five strategies: *Reduction*, *Reminders and Prompts*, *Self-Monitoring*, *Instruction on how to perform the behaviour*, and *Customize and Personalize Features*. On the other hand, the least operationalized strategies are *Distraction or Avoidance*, *Opportunity To Plan For Barrieres*, *Willingness for Behaviour Change*, and *Understand the Difference Between Current Action and Future Goals*.

### 4.6. Correlations Analysis

This section concerns the correlation between the behaviour change score (BCS) and apps ranks (based on users' ratings and the number of installations), as described in Section 3. Pearson correlation was used to test the correlation, and the significance level was set to ($P \leq 0.05$). Table 4 summarizes the results. As the table shows, the correlation between Apps' ranks and the BCS of most of the strategies (25 out of 28 strategies) is positive. These positive correlations indicate that apps' ranks and BCS share some relationships. That is, apps that implement more strategies tend to have a higher rank, which implies higher user ratings and more installations.

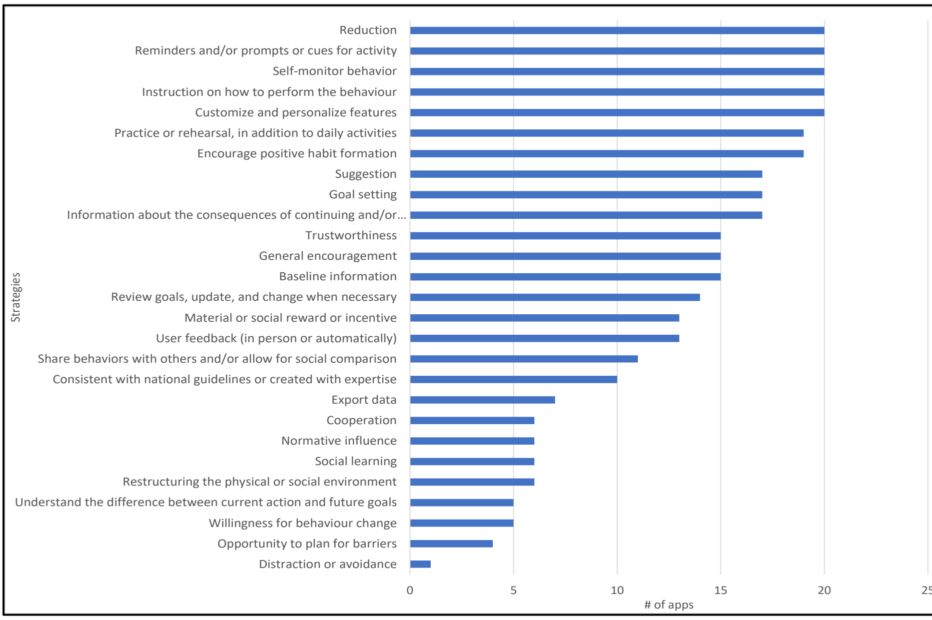


Figure 14. The popularity of behaviour changes strategies among the top 20 apps

Most of these positive correlations range from weak to moderate correlations, except three strategies (*Opportunity to Plan for Barriers*, *Restructuring the Physical or Social Environment*, and *Review goals, update, and change when necessary*) have a strong correlation. Regarding the significance level, out of these 25 positive correlations, eight were found to be statistically significant ($p < 0.05$). We used bold text to easily recognize these values in Table 4.

On the other hand, the results show a negative correlation with three strategies: *Distraction or Avoidance*, *Social Learning*, and *Normative Influence*. Specifically, the results show a moderate negative correlation with *Normative Influence* and a weak negative correlation with the other two strategies. These three correlations were found to be statistically insignificant.

## 5. Discussion and design recommendations

Our results show that mobile health apps are in high demand. There are over 98 million installs for the 70 apps evaluated in this study. In addition, these apps capture users' interest, and they gain a high rating from users based on how useful they find them. The high number of installations along with high ratings indicate that people accept these apps and show an intention to use them to manage their own health. Therefore, our first conclusion is that mobile apps are a large and promising domain that has gained credibility and viability in recent years. Nonetheless, the field of mobile apps for health and wellness is not in its ideal shape yet, and there are many opportunities for improvement. Thus, we invite researchers and practitioners to put more effort toward advancing working in the area of mobile apps for health and wellness. The following points summarize our findings, insights, and recommendations based on the findings from the review and app analysis.

Table 4. Correlation between BCS score and the behaviour change strategies. Bolded values are statistically significant.

| **Strategy** | **Pearson Correlation (*r*)** | ***p*-value** | **N** |
|---|---|---|---|
| Opportunity to plan for barriers | **0.819** | **.046** | 6 |
| Restructuring the physical or social environment | **0.668** | **.049** | 9 |
| Review goals, update, and change when necessary | **0.503** | **.003** | 32 |
| Export data | **0.460** | **.004** | 38 |
| Information about the consequences of continuing and/or discontinuing behaviour | **0.424** | **.020** | 30 |
| Reduction | **0.264** | **.038** | 62 |
| Customize and personalize features | **0.245** | **.046** | 67 |
| Self-monitor behaviour | **0.238** | **.050** | 68 |
| Willingness for behaviour change | 0.641 | .170 | 6 |
| Understand the difference between current action and future goals | 0.592 | .093 | 9 |
| Consistent with national guidelines or created with expertise | 0.376 | .125 | 18 |
| Baseline information | 0.341 | .076 | 28 |
| Instruction on how to perform the behaviour | 0.252 | .075 | 51 |
| User feedback (in person or automatically) | 0.430 | .075 | 18 |
| Share behaviors with others and/or allow for social comparison | 0.302 | .083 | 34 |
| Cooperation | 0.301 | .241 | 17 |
| Practice or rehearsal, in addition to daily activities | 0.240 | .064 | 60 |
| Reminders and/or prompts or cues for activity | 0.237 | .056 | 66 |
| Trustworthiness | 0.234 | .122 | 45 |
| Suggestion | 0.207 | .219 | 37 |
| Goal setting | 0.199 | .213 | 41 |
| Material or social reward or incentive | 0.165 | .421 | 26 |
| Encourage positive habit formation | 0.171 | .306 | 38 |
| General encouragement | 0.112 | .529 | 34 |
| Distraction or avoidance | -0.204 | .742 | 5 |
| Social learning | -0.281 | .403 | 11 |
| Normative influence | -0.528 | .144 | 9 |

- In regard to the app's categories, we found that *Journalling* is the most popular category. Journaling apps have shown success in the health and wellness domain [34]. Several benefits of journaling have been identified in the literature, which include the release of pent-up emotions, lowered blood pressure, enhanced immune function, and decreased depressive symptoms [36, 37, 38]. Accordingly, we can say that journaling is an effective application. Therefore, **we recommend using journaling as a general strategy to promote healthy behaviours**.

- The literature introduces several behaviour change strategies. Our results indicate that the considered mobile apps operationalize 12 strategies on average. A previous study found that mHealth apps in physical activity, diet, drug and alcohol use, and mental health deploy less than 12 strategies, where the mean is 5. The literature in behaviour change interventions indicates that designers tend to operationalize multiple strategies in a single app with the hope of making the app convincing for a broader range of users. The use of multiple strategies, however, may cause cognitive overload because the more the strategies, the more likely the app becomes too complex [39]. According to Kaptein [40], "*in some situations, using multiple strategies can be detrimental as compared to the presentation of a single, correct strategy*." Also, using multiple strategies in a single app makes it difficult to evaluate which strategy works better (for which audience) and why they work, making it difficult to adapt the findings from successful behaviour change apps to other apps [41]. For instance, in a comparison between two versions of an app, Orji et al. [41] demonstrated that using a single appropriate strategy in behaviour change games leads to an effective game. These observations and results indicate a lack of clear guidelines for mHealth apps' development. Whether or not using multiple strategies will increase the effectiveness of persuasive interventions for change is still unclear. Therefore, further studies are required to explore the reasons behind these differences and identify best practices regarding the number of employed strategies. This suggests that **combining multiple effective strategies in behaviour change apps may not necessarily enhance the app's performance** with respect to promoting desired behavior change. Therefore, **we recommend designers choose appropriate strategies cautiously, especially if a user model is not available**.
- The results presented in Section 4.6 show significant positive correlations between apps' ranks and eight strategies (as depicted in Table 4). These significant positive correlations indicate that operationalizing the eight strategies helps design a successful mHealth intervention. Accordingly, **we recommend designers to consider the following points when developing mHealth apps**: 1) encourage users to think about potential barriers and identify ways of overcoming them, 2) prompt users to alter the environment to make it more supportive of the target behaviour, 3) help users review and monitor how they progress in achieving their previously set behavioural goals, 4) allow for the export of information and progress to an external user, 5) emphasize the consequences of behaviours, 6) help users perform the target behaviour by reducing and dividing the tasks into simple tasks, 7) deliver services in a personalized manner and allows users to customize the app's features, and 8) allows for regular monitoring of the activity, such as logging of daily cigarettes. For detailed results, we recommend referring to Table 4 as a reference before implementing strategies.
- By comparing the number of apps (Section 4.2) and the total number of installations (Section 4.3) for each category, we notice that app categories with a higher number of apps do not necessarily have a higher number of installations. For instance, although *the Meditation category has fewer apps than* the *Habit Track*, *Mental Health*, and *Routine Task Tracking* categories, it has a higher number of installations. Specifically, it has the second-highest number of installations. Also, apps that belong to *Health and Wellbeing* (only two apps) are the third most installed apps. These numbers indicate that categories such as *Meditation*, and *Health and Wellbeing*, are highly sought after by users, and they capture users' interest. However, only a few apps are available. Therefore, **we recommend that decision-makers and investors consider developing more apps**

**focused on general health and wellbeing and meditation**. For the same reason, we also **recommend that decision-makers consider apps that help users track their progress toward achieving their goals and change users' habits**.

- Section 4.3 shows the total number of installations for Android apps compared to iOS apps. The first observation here is that the *Meditation* category is among the top three categories in both domains, which means that it captures the interest of users of both systems. Therefore, we suggest that this category is a good target for future apps and research. The second observation is that the total number of installations for Android systems is different compared to iOS systems. For instance, in Android, the highest three categories in terms of installations are *Journaling*, *Meditation*, and *Health and Wellbeing* categories. In iOS, on the other hand, the most installed categories are *Sleep*, *Virtual Companion*, and *Meditation*. These results show different interests for Android users compared to iOS users. Previous research began investigating the factors potentially related to the differences between Android and iOS users, such as users' personalities [42], demographics [43], privacy concerns [44], and more. Other research described the general distribution of Android and iOS users. For instance, previous studies found that iOS users are more likely to be female, in their mid-30s, have a graduate degree and a higher income group, and spend more time using apps than Android users [43]. Therefore, **we recommend considering target users' differences when engineering mHealth apps**. Despite the existence of research on the differences between Android and iOS users, the field is not been fully explored yet. Therefore, **we invite researchers to investigate the exact shape and characteristics of these differences between users of different platforms and to understand the correlation between users of each platform and different apps' categories**. By identifying these differences, mHealth interventions can be designed and developed to be more adaptive and personalized.
- The results also show that *the most employed strategies are self-monitoring, Customize and Personalize Features, and Reminder*s. Previous studies have confirmed these results to a high extent [2, 21]. So, **we recommend implementing these strategies as essential components for future implementations of mobile health applications**. On the other hand, *Recognition*, *Distraction or Avoidance*, and *Willingness for Behaviour Change* are the least employed strategies. The actual reason behind the less popularity of the aforementioned strategies is not clear yet. It could be related to the strategies' effectiveness, but it could be related to other causes. Therefore, **we recommend systems designers and HCI researchers investigate the actual effectiveness of these less popular strategies and investigate the reasons behind not using them by most apps**.
- Our findings show that the most commonly employed strategies (*Self Monitoring*, and *Customize and Personalize Content*) are also significantly correlated (in terms of the BCS) with apps' rank. These results indicate that *Self Monitoring*, and *Customize and Personalize Content* are common and efficient. Thus, **we strongly recommend designing mHealth apps such that they help and motivate users to log and track their target behaviour and that give users control over the apps' features and the services provided**. On the other hand, the two strategies (*Opportunity to Plan for Barriers,* and *Restructuring the Physical or Social Environment*) are among the least common strategies. However, the BCS of both strategies is significantly positively correlated with apps ranking. These results indicate that (*Opportunity to Plan for Barriers,*

and *Restructuring the Physical or Social Environment*) are underrepresented strategies in the apps. **Therefore, we confirm the need for a more comprehensive study to understand the actual effectiveness of each strategy and reveal the reasons why some strategies are rarely used although they have potential effectiveness**.

- Mobile health apps differ in their targeted health behaviour, targeted patients, and behaviour change strategies deployed. Also, health behaviours can be divided into several types, such as physical activity, diet, drug use, and mental health [2]. Thus, **designers should pay special attention to these differences when deciding on the strategies to be deployed**. This is also a promising research direction where **researchers should consider the differences between mHealth apps and study the relationship between all factors that affect selecting the most appropriate strategy**. These factors include application domain, health behaviour to be changed, targeted users, etc.
- As mentioned in Section 3, our app selection was focused on apps that adopt AI-based solutions. The selected apps used a wide range of AI techniques to help with various requirements. For instance, "Replika: My AI Friend[1]", and "Wysa"[2] uses AI techniques to implement virtual companion that can chat like a human, while "Youper Therapy[3]" use AI technique for mood tracking. Other apps, like "Sleep Cycle: Sleep Tracker[4]", use machine learning algorithms to enhance sleep by monitoring light and deep sleep patterns and analyzing users' sounds. "Headspace", on the other hand, incorporate multiple AI-based features, including semantic embeddings for user search terms and personalized exercise content based on users' biometrics. These are just examples of how the considered apps adopted AI-based techniques. We invite researchers to consider investigating AI techniques used in various mHealth apps domains.
- Finally, Section 4.5 discusses the behaviour change potential of the considered app based on the Behaviour Change Score. It is worth mentioning that this score represents the apps' effectiveness in terms of implementation of the behaviour change strategies considered in the study rather than the actual effectiveness in changing users' behaviour. The actual effectiveness of an app in changing behaviour could not be measured without an in-depth user study. So, longitudinal and in-depth user studies are needed to understand the actual effectiveness of each app in its particular domain.

### 5.1. Limitations

This paper revealed trends in the behaviour change app for the health and wellness domain. Despite the results and insights presented in this paper, the work has some limitations. As mentioned in Section 3, some apps' features were not investigated practically by the authors because they are only available for premium accounts. Therefore, the team relied on the description of these apps and their features as presented in the app stores and other resources. Nonetheless, these descriptions and resources are enriched with screenshots and video demos that clearly describe these paid features, and we believe they could be sufficient. Also, the study focused on popular applications only (selected using a combination of app downloads and user rating) with the aim of uncovering popular trends in the highly rated and adopted apps that will

1 https://play.google.com/store/apps/details?id=ai.replika.app&hl=en&gl=us

2 https://play.google.com/store/apps/details?id=bot.touchkin&hl=en&gl=us

3 https://play.google.com/store/apps/details?id=br.com.youper&hl=en&gl=us

4 https://play.google.com/store/apps/details?id=com.northcube.sleepcycle&hl=en&gl=us

inform future app development in this area. We recognize that this may not be represented in all the health and wellness apps; hence, our findings may not generalize. Finally, it is worth mentioning that the ratings of an application in app stores show the average rating across all app versions. That is, if the latest version has improved, there is no way to separate ratings according to different versions of the apps. So, low ratings based on previous versions will still affect the overall rating. Nonetheless, to the best of our knowledge, there is no way to separate ratings according to different versions of the apps.

## 6. Conclusion and Future Work

This paper explores the behaviour change mobile apps for health and wellness. It investigates 70 apps available in the most common mobile app stores, Google Play Store, and Apple's App Store. The study aims to understand trends in these apps, the most common behaviour changes strategies used, and to what extent these apps implement those strategies. The apps were evaluated against 28 behaviour change strategies selected from two common frameworks, ABACUS and PSD. Our results show that (1) Journaling is the most common apps category, (2) *Self-monitoring* was the highest employed strategy, (3) Recognition strategy was not used in any of the apps, (4) the apps implement 12 strategies on average, and (5) apps' user ratings are positively correlated to most of the strategies employed.

In future work, we need to cross-examine (theoretical versus practical) the best-performing strategies. That is, we will evaluate which strategies have the most theoretical and evidence-based potential for behaviour change and see if that maps onto the most highly ranked and low-ranked apps. Also, we are planning to study the opposite side of the current work. Specifically, we need to investigate low-rated apps and see which strategies are implemented in these apps compared to the highly ranked apps. To examine why behaviour does not occur (or change), we will rely on frameworks, such as Fogg's Behavioural Model (FBM) [45]. Finally, we will investigate the AI techniques and approaches used in mHealth apps.

## ACKNOWLEDGEMENT

This research was undertaken, in part, thanks to funding from the Canada Research Chairs Program. We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) through the Discovery Grant and Mitacs through the Mitacs Accelerate program.

## REFERENCES

[1] T. N. Stephens, A. Joerin, M. Rauws, and L. N. Werk, "Feasibility of pediatric obesity and prediabetes treatment support through Tess, the AI behavioral coaching chatbot," *Transl. Behav. Med.*, vol. 9, no. 3, pp. 440–447, May 2019, doi: 10.1093/TBM/IBZ043.

[2] F. H. Mckay, A. Wright, J. Shill, B. Hons, and M. Uccellini, "Using Health and Well-Being Apps for Behavior Change : A Systematic Search and Rating of Apps Corresponding Author :," vol. 7, pp. 1–11, 2019, doi: 10.2196/11926.

[3] P. Parvin, "Real-time anomaly detection in elderly behavior," *Proc. ACM SIGCHI Symp. Eng. Interact. Comput. Syst. EICS 2018*, Jun. 2018, doi: 10.1145/3220134.3220145.

[4] A. Trifan, M. Oliveira, and J. L. Oliveira, "Passive sensing of health outcomes through smartphones: Systematic review of current solutions and possible limitations," *JMIR mHealth uHealth*, vol. 7, no. 8, 2019, doi: 10.2196/12649.

[5] D. Levine *et al.*, "Design and testing of a mobile health application rating tool," *npj Digit. Med.*, vol. 3, no. 1, pp. 1–7, 2020, doi: 10.1038/s41746-020-0268-9.

[6] M. L. Patel, C. Hopkins, M. Brooks, and G. G. Bennett, "Comparing Self-Monitoring Strategies for Weight Loss in a Smartphone App: Randomized Controlled Trial," *JMIR mHealth uHealth*, vol. 7, no. 1, pp. 1–18, 2019, doi: 10.2196/11130.

[7] H. Oinas-kukkonen, M. Harjumaa, and H. Oinas-kukkonen, "Persuasive Systems Design : Key Issues , Process Model , and System Features," vol. 24, 2009, doi: 10.17705/1CAIS.02428.

[8] M. Mitchell, E. Lau, L. White, and G. Faulkner, "Commercial app use linked with sustained physical activity in two Canadian provinces: A 12-month quasi-experimental study," *Int. J. Behav. Nutr. Phys. Act.*, vol. 17, no. 1, pp. 1–9, Feb. 2020, doi: 10.1186/S12966-020-00926-7/FIGURES/2.

[9] K. Oyibo, "Investigating the Key Persuasive Features for Fitness App Design and Extending the Persuasive System Design Model: A Qualitative Approach," 2021, doi: 10.1177/2327857921101022.

[10] T. Räisänen, T. Lehto, and H. Oinas-Kukkonen, "Practical findings from applying the PSD model for evaluating software design specifications," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6137 LNCS, pp. 185–192, 2010, doi: 10.1007/978-3-642-13226-1_19.

[11] F. M. Calisto, N. Nunes, and J. C. Nascimento, "BreastScreening: On the Use of Multi-Modality in Medical Imaging Diagnosis," *ACM Int. Conf. Proceeding Ser.*, Sep. 2020, doi: 10.1145/3399715.3399744.

[12] F. M. Calisto, C. Santiago, N. Nunes, and J. C. Nascimento, "Introduction of human-centric AI assistant to aid radiologists for multimodal breast image classification," *Int. J. Hum. Comput. Stud.*, vol. 150, p. 102607, Jun. 2021, doi: 10.1016/J.IJHCS.2021.102607.

[13] F. M. Calisto, J. C. Nascimento, A. Ferreira, and D. Gonalves, "Towards touch-based medical image diagnosis annotation," *Proc. 2017 ACM Int. Conf. Interact. Surfaces Spaces, ISS 2017*, pp. 390–395, Oct. 2017, doi: 10.1145/3132272.3134111.

[14] M. L. Patel, C. M. Hopkins, T. L. Brooks, and G. G. Bennett, "Comparing Self-Monitoring Strategies for Weight Loss in a Smartphone App: Randomized Controlled Trial," *JMIR Mhealth Uhealth 2019;7(2)e12209 https//mhealth.jmir.org/2019/2/e12209*, vol. 7, no. 2, p. e12209, Feb. 2019, doi: 10.2196/12209.

[15] J. G. Thomas and D. S. Bond, "Review of innovations in digital health technology to promote weight control," *Curr. Diab. Rep.*, vol. 14, no. 5, 2014, doi: 10.1007/S11892-014-0485-1.

[16] N. Reščič, E. Valenčič, E. Mlinarič, B. K. Seljak, and M. Luštrek, "Mobile nutrition monitoring for well-being," *UbiComp/ISWC 2019- - Adjun. Proc. 2019 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput. Proc. 2019 ACM Int. Symp. Wearable Comput.*, pp. 1194–1197, Sep. 2019, doi: 10.1145/3341162.3347076.

[17] B. Y. B. Kim and J. Lee, "Smart Devices for Older Adults Managing Chronic Disease: A Scoping Review," *JMIR mHealth uHealth*, vol. 5, no. 5, May 2017, doi: 10.2196/MHEALTH.7141.

[18] J. Geuens, T. W. Swinnen, R. Westhovens, K. de Vlam, L. Geurts, and V. Vanden Abeele, "A Review of Persuasive Principles in Mobile Apps for Chronic Arthritis Patients: Opportunities for Improvement," *JMIR mHealth uHealth*, vol. 4, no. 4, Oct. 2016, doi: 10.2196/MHEALTH.6286.

[19] N. Baghaei *et al.*, "Designing Individualised Virtual Reality Applications for Supporting Depression: A Feasibility Study," *EICS 2021 - Companion 2021 ACM SIGCHI Symp. Eng. Interact. Comput. Syst.*, pp. 6–11, Jun. 2021, doi: 10.1145/3459926.3464761.

[20] J. Matthews, K. Than Win, H. Oinas-Kukkonen, and M. Freeman, "Persuasive Technology in Mobile Applications Promoting Physical Activity: a Systematic Review," doi: 10.1007/s10916-015-0425-x.

[21] M. Alhasani, D. Mulchandani, O. Oyebode, and R. Orji, "A Systematic Review of Persuasive Strategies in Stress Management Apps," 2020.

[22] P. R. T. Ko, J. A. Kientz, E. K. Choe, M. Kay, C. A. Landis, and N. F. Watson, "Consumer Sleep Technologies: A Review of the Landscape," *J. Clin. Sleep Med.*, vol. 11, no. 12, pp. 1455–1461, Dec. 2015, doi: 10.5664/JCSM.5288.

[23] H. E. Payne, C. Lister, J. H. West, and J. M. Bernhardt, "Behavioral Functionality of Mobile Apps in Health Interventions : A Systematic Review of the Literature Corresponding Author :," vol. 3, 2015, doi: 10.2196/mhealth.3335.

[24] A. Monge Roffarello and L. De Russis, "The race towards digital wellbeing: Issues and opportunities," *Conf. Hum. Factors Comput. Syst. - Proc.*, May 2019, doi: 10.1145/3290605.3300616.

[25] M. Han and E. Lee, "Effectiveness of Mobile Health Application Use to Improve Health Behavior Changes : A Systematic Review of Randomized Controlled Trials," vol. 24, no. 3, pp. 207–226, 2018.

[26] F. H, A. Wright, J. Shill, H. Stephens, and M. Uccellini, "Using Health and Well-Being Apps for Behavior Change: A Systematic Search and Rating of Apps," *JMIR Mhealth Uhealth 2019;7(7)e11926 https//mhealth.jmir.org/2019/7/e11926*, vol. 7, no. 7, p. e11926, Jul. 2019, doi: 10.2196/11926.

[27] A. C. Powell *et al.*, "Interrater Reliability of mHealth App Rating Measures: Analysis of Top Depression and Smoking Cessation Apps," *JMIR Mhealth Uhealth 2016;4(1)e15*, vol. 4, no. 1, pp. 1–9, Feb. 2016, doi: 10.2196/MHEALTH.5176.

[28] F. H. McKay, C. Cheng, A. Wright, J. Shill, H. Stephens, and M. Uccellini, "Evaluating mobile phone applications for health behaviour change: A systematic review," *J. Telemed. Telecare*, vol. 24, no. 1, pp. 22–30, Jan. 2018, doi: 10.1177/1357633X16673538.

[29] S. Batra, R. A. Baker, T. Wang, F. Forma, F. DiBiasi, and T. Peters-Strickland, "Digital health technology for use in patients with serious mental illness: a systematic review of the literature," *Med. Devices (Auckl).*, vol. 10, pp. 237–251, Oct. 2017, doi: 10.2147/MDER.S144158.

[30] H. Þórarinsdóttir, L. V. Kessing, and M. Faurholt-Jepsen, "Smartphone-Based Self-Assessment of Stress in Healthy Adult Individuals: A Systematic Review," *J Med Internet Res 2017;19(2)e41 https//www.jmir.org/2017/2/e41*, vol. 19, no. 2, p. e6397, Feb. 2017, doi: 10.2196/JMIR.6397.

[31] H. Te Kao, S. Yan, H. Hosseinmardi, S. Narayanan, K. Lerman, and E. Ferrara, "User-Based Collaborative Filtering Mobile Health System," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 4, no. 4, p. 17, Dec. 2020, doi: 10.1145/3432703.

[32] M. Milne-Ives, C. LamMEng, C. de Cock, M. H. van Velthoven, and E. M. Ma, "Mobile apps for health behavior change in physical activity, diet, drug and alcohol use, and mental health: Systematic review," *JMIR mHealth uHealth*, vol. 8, no. 3, pp. 1–16, 2020, doi: 10.2196/17046.

[33] R. Zhang, K. E. Ringland, M. Paan, D. C. Mohr, and M. Reddy, "Designing for Emotional Well-being: Integrating Persuasion and Customization into Mental Health Technologies; Designing for Emotional Well-being: Integrating Persuasion and Customization into Mental Health Technologies," 2021, doi: 10.1145/3411764.3445771.

[34] L. J. Dimitroff, L. Sliwoski, S. O'brien, and L. W. Nichols, "Change your life through journaling-The benefits of journaling for registered nurses Integrating mindfulness into daily practice with undergraduate nursing students View project Change your life through journaling-The benefits of journaling for registered nurses," *Artic. J. Nurs. Educ. Pract.*, vol. 7, no. 2, 2016, doi: 10.5430/jnep.v7n2p90.

[35] R. Orji, R. Lomotey, K. Oyibo, F. Orji, J. Blustein, and S. Shahid, "Tracking feels oppressive and 'punishy': Exploring the costs and benefits of self-monitoring for health and wellness," *Digit. Heal.*, vol. 4, p. 205520761879755, Jan. 2018, doi: 10.1177/2055207618797554.

[36] C. L. Kuo, M. Turton, S. F. Cheng, and J. Lee-Hsieh, "Using clinical caring journaling: Nursing student and instructor experiences," *J. Nurs. Res.*, vol. 19, no. 2, pp. 141–148, 2011, doi: 10.1097/JNR.0B013E31821AA1A7.

[37] G. Parr, S. Haberstroh, and J. Kottler, "Interactive journal writing as an adjunct in group work," *http://dx.doi.org/10.1080/01933920008411464*, vol. 25, no. 3, pp. 229–242, Sep. 2008, doi: 10.1080/01933920008411464.

[38] E. M. Brady and H. Z. Sky, "JOURNAL WRITING AMONG OLDER LEARNERS," *http://dx.doi.org/10.1080/713844282*, vol. 29, no. 2, pp. 151–163, Feb. 2010, doi: 10.1080/713844282.

[39] K. Rilla, "Culturally-Relevant Persuasive Technology," *Dr. Thesis*, 2008, Accessed: Oct. 22, 2021. [Online]. Available: http://researcharchive.vuw.ac.nz/handle/10063/365.

[40] M. C. Kaptein, "Personalized persuasion in ambient intelligence," *Tech. Univ. Eindhoven*, 2012, doi: 10.6100/IR729200.

[41] R. Orji, R. L. Mandryk, and J. Vassileva, "Improving the Efficacy of Games for Change Using Personalization Models," *ACM Trans. Comput. Interact*, vol. 24, no. 5, 2017, doi: 10.1145/3119929.

[42] F. M. Götz, S. Stieger, and U.-D. Reips, "Users of the main smartphone operating systems (iOS, Android) differ only little in personality," *PLoS One*, vol. 12, no. 5, p. e0176921, May 2017, doi: 10.1371/JOURNAL.PONE.0176921.

[43] H. Kaur Ubhi, D. Kotz, S. Michie, O. C. P van Schayck, and R. West, "TBM A comparison of the characteristics of iOS and Android users of a smoking cessation app," doi: 10.1007/s13142-016-0455-z.

[44] Z. Benenson, F. Gassmann, and L. Reinfelder, "Android and iOS Users' Differences concerning Security and Privacy," *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2013-April, pp. 817–822, Apr. 2013, doi: 10.1145/2468356.2468502.

[45] B. Fogg, "A behavior model for persuasive design," *ACM Int. Conf. Proceeding Ser.*, vol. 350, 2009, doi: 10.1145/1541948.1541999.


Received February 2022; accepted April 2022