Narrating Routines through Game Dynamics: Impact of a Gamified Routine Management App for Autistic Individuals

Bogoan Kim
Hanyang University
Seoul, Republic of Korea
bogoankim@hanyang.ac.kr

Dayoung Jeong
Hanyang University
Seoul, Republic of Korea
dayoungjeong@hanyang.ac.kr

Hwajung Hong
KAIST
Daejeon, Republic of Korea
hwajung@kaist.ac.kr

Kyungsik Han*
Hanyang University
Seoul, Republic of Korea
kyungsikhan@hanyang.ac.kr

ABSTRACT

Maintaining a daily routine has profound implications for physical, emotional, and social well-being. Autistic individuals may experience various challenges in establishing and maintaining a healthy daily routine due to their tendency to be inactive in daily life combined with their characteristics and preferences. Previous studies employing mobile technology to support autistic individuals have primarily focused on self-help functions, with limited exploration into the detailed needs of these individuals to develop and maintain personalized routines. In this study, we conducted a nine-week field study with 18 autistic individuals using *RoutineAid*, a gamified app designed to support key routines of autistic individuals (i.e., physical activity, diet, mindfulness, and sleep). Our analysis incorporated five measures of self-evaluation on daily life, app usage logs, Fitbit physical activity data, and interviews. Our findings demonstrate the effectiveness of *RoutineAid* and highlight its two primary affordances for autistic individuals: (1) promoting self-efficacy and embedding health behavior and (2) refining daily routines for healthier outcomes. We discuss salient design insights for developing daily routine management systems for autistic individuals.

CCS CONCEPTS

• **Human-centered computing → Accessibility**.

KEYWORDS

Autism, Daily Routine Management, Gamified App, Field Study

ACM Reference Format:
Bogoan Kim, Dayoung Jeong, Hwajung Hong, and Kyungsik Han. 2024. Narrating Routines through Game Dynamics: Impact of a Gamified Routine Management App for Autistic Individuals. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3613904.3642357

*Corresponding author

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642357

1 INTRODUCTION

Maintaining a daily routine is essential for everyone, as it can affect physical, emotional, and social well-being. For autistic individuals [1], their unique characteristics and preferences (e.g., specific behavior patterns, special interests) [3] can sometimes make it challenging to establish and maintain a healthy daily routine [7]. Many studies have highlighted that autistic individuals are particularly prone to hypoactive conditions, such as obesity, hyperlipidemia, and hypertension, due to their high levels of sedentary behavior [18, 23, 64]. These characteristics tend to intensify with age [58], and the recent pandemic has further exacerbated them [21]. Physical activity (PA), while forming a healthy routine, influences other essential daily routine components (e.g., eating habits, mindfulness, and sleep). Therefore, it is important to support environments in which autistic individuals can explore and establish these other routine components while improving their inactivity.

Smartphone-based programs have emerged as an effective tool for promoting a healthy lifestyle because they allow users to set behavioral goals, monitor their own progress, and quickly plan their actions in an accessible, time-efficient, and cost-effective manner [47, 49]. Several studies have explored the potential of smartphone technology to support improvements in the daily routines of autistic individuals [30, 60]. However, these studies have primarily focused on improving basic self-help functions for autistic children (e.g., washing the face, brushing the teeth). Moreover, the apps for routine management in commercial stores (e.g., Google Play Store, Apple App Store) often lack clear action items and come with a high degree of freedom, making the onboarding process and sustained use challenging and potentially burdensome for autistic individuals [37, 55].

To address this, we developed a gamified app, *RoutineAid*, to support the self-directed exploration and establishment of four key daily routine components for autistic individuals: (1) PA, (2) eating habits, (3) mindfulness, and (4) sleep. The design of *RoutineAid* carefully considered the characteristics and preferences of autistic individuals (e.g., appreciation of visual reasoning processes, desire for predictable interfaces [37, 46]). Throughout the design process, we gathered feedback from autism stakeholders (e.g., autism professionals involved in diagnosis advice and support, parents of autistic individuals, and autistic individuals), which helped us

[1]We use identity-first language due to a reported preference of autistic individuals [40] and recent movement in academia [9].

identify and incorporate recommended health behaviors. Users are introduced to these behaviors and their importance through in-game quests and visual narratives. Completing these quests influences gamified elements (e.g., rankings, scores) and rewards users with visual incentives (e.g., building upgrades, village character unlocks). Furthermore, the app provides a *routine diary*, allowing users to self-reflect on their accomplished quests. Such features increase autistic individuals' engagement with the app and support the exploration and establishment of healthy daily routines.

Design validation is a key process to ensure that user characteristics and requirements are appropriately reflected in an app [29]. However, analyzing user logs and feedback as they interact with the app in real-world settings is also important [25]. This approach facilitates a holistic understanding of end-user behavior patterns and habits. By analyzing this data, we can assess the ongoing effectiveness of the app and identify areas for improvement [53]. Furthermore, according to the recent autism research priorities [4], there is a critical need for more effective interventions over autism-related health conditions and an understanding of the long-term daily living challenges of autistic individuals [31]. While many systems designed for autistic individuals have been validated for their design's novelty and originality, a somewhat limited body of research extends beyond design validation to examine long-term usage patterns, interactions, and effectiveness in real-world settings [4, 65]. Furthermore, autistic individuals exhibit a wide range of challenges in communication, social interaction, and sensory experiences [3]. Particularly for routine management apps, given the vast differences in individual requirements for managing daily routines, it becomes essential to investigate the diverse needs and challenges that emerge during sustained use.

To this end, in this paper, we conducted a nine-week field study with 18 autistic individuals to investigate (1) the effectiveness of the routine management app, (2) the correlation between gamified performance and actual user behaviors, and (3) the changes in the four routine components (PA, eating habits, mindfulness, and sleep) that the participants exhibited over the course of the study. Our investigation used a mixed-methods approach that included the following: (1) five measures (i.e., General Self-efficacy Scale [61], International Physical Activity Questionnaires-Short Form [22], Eating Attitudes Test-26 [28], Perceived Stress Scale [19], and Pittsburgh Sleep Quality Index [15]) to assess the changes experienced by participants while using the app, (2) user interaction logs on *RoutineAid*, (3) Fitbit PA data (i.e., steps and intensity-based PA durations), and (4) interview analysis. We also conducted additional interviews with three family medicine physicians to determine the meaning and implications of the observed changes in autistic individuals from a general health management perspective. Based on the study findings, we highlighted two distinct affordances of *RoutineAid*: (1) promoting self-efficacy and embedding health behavior and (2) refining daily routines for healthier outcomes. Through extensive analysis of the multimodal data, we gained a comprehensive understanding of participants' behaviors and patterns. This insight enabled us to derive design guidelines to further improve routine management tools for autistic individuals. In particular, our findings emphasized the profound influence of the app's narrative on daily health behaviors, as well as the tangible benefits observed in areas such as sleep patterns and eating habits.

To summarize, this paper makes the following contributions:

- We demonstrate the effectiveness of *RoutineAid* in the sustained management of routines by autistic individuals from a nine-week field study, operationalized by the following two affordances of *RoutineAid*: promoting self-efficacy and embedding health behaviors, and refining daily routines for healthier outcomes.
- We conducted an analysis of measures of self-evaluation on daily life, app usage logs, Fitbit logs, and interview responses to track health behavior changes while using *RoutineAid* and to examine user interaction patterns within the app.
- We present five salient design implications for developing a routine management system that empowers autistic individuals to establish independent and healthy routines: (1) leveraging the interconnectivity of routine components for increased user engagement, (2) calibrating routine quests for improved sustainability, (3) strengthening social connections through emotional sharing in the mind diary, (4) navigating health literacy, and (5) providing more immersive storylines for sustained engagement.

2 RELATED WORK

2.1 Computer-assisted systems for autistic individuals

Many computer-assisted systems have been designed to assist autistic individuals and their families by suggesting behavioral interventions, addressing specific challenges (e.g., job interview training), and improving their well-being and daily functioning (e.g., communication, social, and problem-solving skills) [41, 42].

Over the past decade, there have been numerous attempts [41] to leverage mobile technology in the fields of education [10, 33] and health [6]. For example, Escobedo et al. [26] developed the "MOSOCO" to improve autistic children's social skills. MOSOCO integrates augmented reality (AR) and visual support, and was based on the Social Compass Curriculum (SCC) [11]. Building on the theme of social connections, Hong et al. [35] designed "SocialMirror," which connects autistic adolescents to their social networks. This work validated the potential of social media to help autistic adolescents transition to adulthood. There have also been mobile systems designed to support the health of autistic individuals. Jo et al. [39] introduced a mobile data collection tool, "GeniAuti," to reduce caregivers' stress when addressing specific behaviors of concern from autistic children. The app allowed caregivers to record these behaviors and associated situations, providing valuable insights into the autistic child's behaviors. After a 26-day study, the app demonstrated its effectiveness in providing insights into the behaviors of autistic children. Antle et al. [2] developed a gamified mobile app to help autistic children manage their anxiety. Over the course of six sessions, the children participated in a neurofeedback game. Following these sessions, teachers and parents observed a noticeable improvement in the children's ability to manage anxiety at school and at home.

More recently, virtual reality (VR) technology has emerged as an effective tool in various fields, allowing users to immerse themselves in realistic environments [12, 48]. Lyu et al. [56] developed a gamified mobile augmented reality app to enrich the social and attentional skills of autistic children. In a three-week field study with five autistic children and their caregivers, the app demonstrated its feasibility in cultivating these skills. Kim et al. [43] developed a VR system to promote the practice of social skills in autistic individuals. In their study of 10 participants, they identified sensor data that accurately captured characteristics of autism, such as repetitive behavior in highly complex tasks. These findings have broadened our quantitative understanding of autistic individuals. Burke et al. [14] developed a VR-based interview training system to mitigate anxiety during job interviews. The system showed positive results, improving job interview outcomes and increasing the self-efficacy of the autistic participants.

Taken together, these studies contribute significantly to the well-being of autistic individuals from multiple perspectives. However, there remains a research gap dedicated to the development of mobile systems that empower autistic individuals to structure their daily lives. *RoutineAid* is designed with gamified elements and visual narratives that enable autistic individuals to comprehensively understand and independently explore and establish healthy daily routines that encompass four routine components: PAs, eating habits, mindfulness, and sleep.

2.2 Mobile technology as a tool for managing daily routines in autistic individuals

Maintaining a consistent daily routine is essential for one's physical, emotional, and social well-being [36]. Research has shown that a healthy routine requires regular PAs and a constant consumption of balanced meals [66]. Regular physical exercise has been proven to improve cognitive functions and executive capabilities [20] and also significantly benefits sleep quality and mental health [50]. On the other hand, an unbalanced diet can intensify feelings of anxiety and increase the risk of depression [27]. In addition, irregular or inadequate sleep can negatively influence emotional and mental health [1]. A deficiency in any of these four components can lead to a cascade of imbalances.

Mobile technology has been widely used as a tool for managing daily routines due to its effectiveness in allowing users to set goals, monitor progress, and plan actions in a time and cost efficient manner. Such capabilities have been extended to autistic individuals. For instance, Lee et al. [54] developed a gamified mobile app to promote PA among autistic individuals. In a four-week field study, participants appreciated the gamified elements of the app and showed increased activity levels and walking time. Bittner et al. [8] proposed an app to encourage PA to support healthy metabolic functions in autistic children. Over a course of four weeks, these children participated in diverse movements (e.g., galloping as fast as possible from one cone to the other) and object control exercises. The researchers tracked metrics such as energy expenditure and heart rate, and concluded that the use of the app may lead to improved metabolic activity. Furthermore, Jo et al. [38] developed a mobile app that monitors meal times to help autistic children establish healthy eating habits. The three-week study results indicated that the app positively influenced autistic children's autonomy and concentration during mealtimes. Carlier et al. [16] designed an app to support the mental health of autistic children and their parents. During a two-week usability study, despite the potential distractions posed by smartphones, the app effectively reduced anxiety and stress levels for both autistic children and their parents.

As such, mobile apps designed to support daily routines (e.g., PAs, dietary habits) have been demonstrated to be effective in helping autistic individuals build healthy and independent lives. However, many of these tools address only a single routine component that makes up a day. Although focusing on a key routine component yields valuable research insights for understanding autistic individuals, there is a critical need for a more comprehensive approach, given the interconnectedness among each routine component. Moreover, the demands of routine management can substantially vary from individual to individual. Thus, longitudinal log collection and analysis can be important to effectively study the different usage patterns and challenges that arise for each user. *RoutineAid* can fill this research gap by analyzing the interaction patterns of autistic participants from four primary routines and long-term use of the app.

3 DAILY ROUTINE MANAGEMENT APP: ROUTINEAID

The validity of *RoutineAid* design was evaluated in our previous research [45]. In summary, *RoutineAid* is designed to support the exploration and formation of four routine components: (1) PA, (2) eating habits, (3) mindfulness, and (4) sleep. To understand the personal challenges and societal or health issues that may impede daily routine regulation, and to identify the primary design objectives of a routine management system, we conducted a formative study with 15 autism stakeholders – five autism professionals, five parents of autistic individuals, and five autistic adults. This formative study was structured based on the Reflective Agile Iterative Design (RAID) approach [34], with the app being redesigned after each iteration until the identified outcomes were fully incorporated. Each iteration involved the autism stakeholders who participated in the formative study. To select these four routine components, we synthesized insights from existing literature and iterative discussions with autism stakeholders. We focused on the essential aspects of a healthy daily routine, emphasizing the role of regular physical activity and a balanced diet in improving cognitive function, sleep quality, and mental health, as highlighted in previous research [20, 51, 67]. Recognizing the intricate relationship between these elements and their collective impact on overall wellness, especially in the context of the challenges posed by the COVID-19 pandemic, we carefully selected these routines to form the core of *RoutineAid*. They were conceptualized as distinct "villages" within the app: PA Village, Eat Village, Mindfulness Village, and Sleep Village. Each village represents a critical area of daily life.

The key design elements of the app incorporate a visual narrative and gamification strategies (e.g., rewards, scores). The primary mechanism of this app involves upgrading the buildings of four routine villages (up to Lv.5) and unlocking village characters (Figures 1 and 2). To achieve this, users are presented with a total of 16 routine quests (four routine quests per village). Users can also add up to two user-generated secondary routine quests per village by themselves (Figure 2-(c)). We provided detailed explanations for each routine

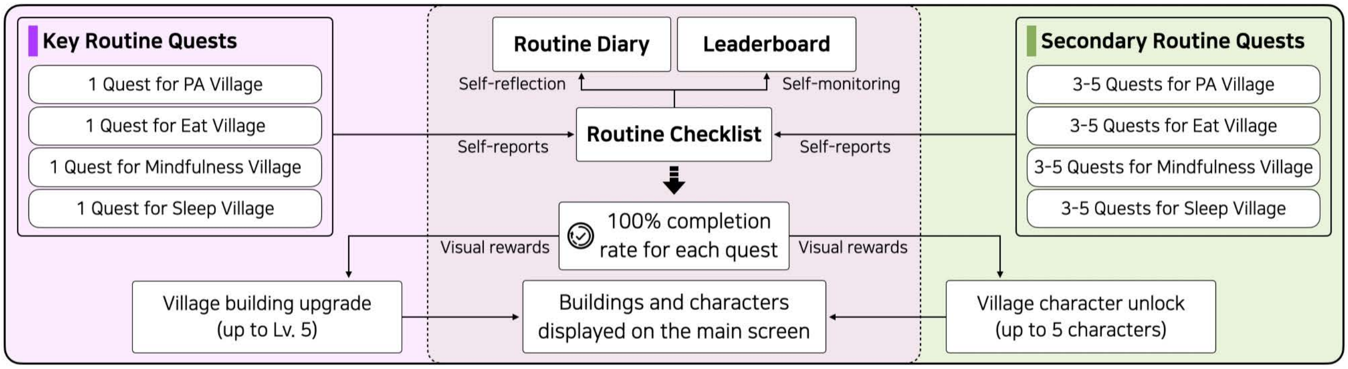

Figure 1: User flow diagram of *RoutineAid*. Users can self-report quest completion in the "routine checklist" once per day for each quest. Achievements are reflected in real-time within the "routine diary" and the leaderboard. The routine diary allows users to self-reflect on logs, while the leaderboard facilitates self-monitoring via scores and rankings. Completing key routine quests leads to building upgrades (up to Lv. 5) while finishing secondary ones results in character unlocks.

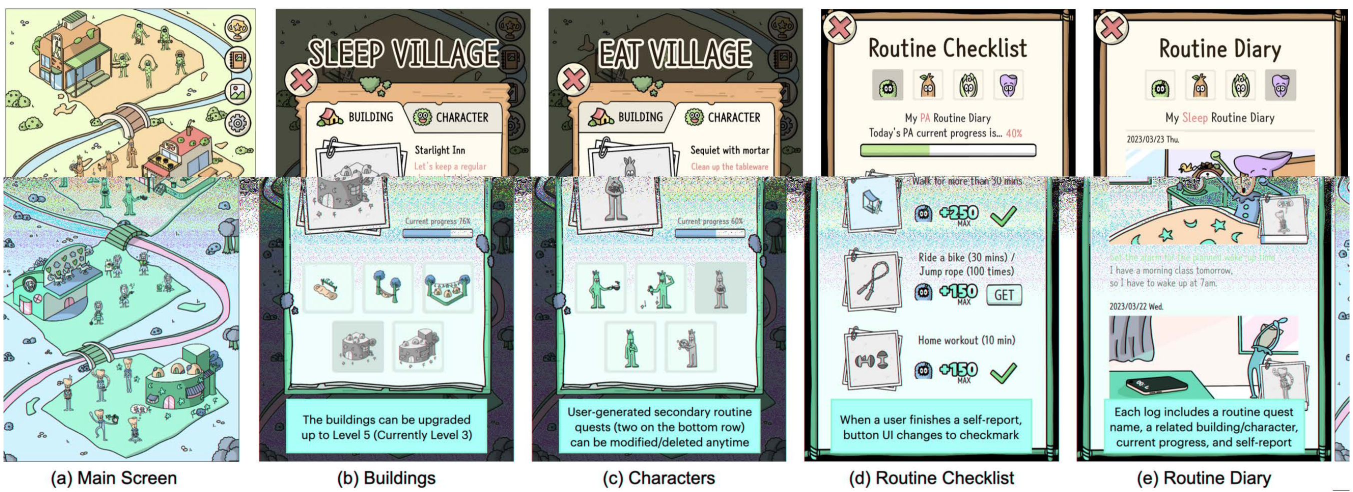

Figure 2: Screenshots of *RoutineAid*. In (a), the main screen displays the completed phase, with each village building fully upgraded to level 5 and all village characters unlocked. In (b) and (c), each village building and village character is presented with the corresponding image and progress bar. Note that there is a daily quota for both key and secondary routines, thus users need to consistently complete quests (encouraged daily) to reach 100% completion in order to upgrade village buildings and unlock village characters. In (d), the users can move to four routine villages by tapping the icon above. In (e), the users can upload their own images related to each routine quest.

quest (e.g., walk for more than 30 minutes), considering their difficulty in elaborating tasks for problem-solving [44]. Each village has a "key routine quest," and the user can upgrade a "building" when the key routine quest is completed. The remaining routine quests were set as "secondary routine quests," mapped to the "village character unlocks." Table 1 shows the key/secondary routine quests for each routine village, and Figure 3 shows a user scenario of "P7," illustrating an example of interaction with *RoutineAid*'s features.

Users can self-report their routine quest achievements in the "routine checklist" via a notification they receive at 9 PM or by accessing the app at their preferred time. After discussions with autism stakeholders, it was determined that adding reminders at specific times would be a good strategy to keep autistic users engaged in the app. Therefore, additional notifications were sent at 2 PM and 7 PM, times chosen based on these discussions to align with the post-lunch and post-dinner periods. These notifications provide brief information about the app or share detailed health facts in a quotable snippet (e.g., walking for 10 minutes after a meal can lower average blood sugar levels by 12%). In addition, users can self-reflect on their achieved routine quest records through the "routine diary" and check their scores and rankings through the leaderboard. In summary, the app uses visual cues and gamified elements to encourage users to complete as many routine quests as possible, with completed quests leading to progressively evolving UI and increased in-game points.

4 METHOD

We used a mixed methods approach, integrating measures, app usage log, Fitbit log, and interview analysis, to understand the use of *RoutineAid* among autistic individuals. Moreover, through

Table 1: Each village features a key routine quest and three secondary routine quests ("Key" refers to key routine quests, "Secondary" to secondary routine quests). By accomplishing the key quest in a village, users can upgrade a building. Meanwhile, the accomplishment of secondary quests in a village allows users to unlock village characters.

Village Name	Routine Type	Game Element	Routine Quest Description
PA Village	Key	PA Building	Walk for more than 30 minutes
	Secondary	PA character 1	Ride a bike for 30 minutes or jump rope 100 times
		PA character 2	10-minute home workout
		PA character 3	Run outdoors for five minutes
		PA character(s) 4-5	User-generated routine quest(s)
Eat Village	Key	Eat Building	Regular intake of meals
	Secondary	Eat character 1	Calculate meal calories
		Eat character 2	Calculate snack calories
		Eat character 3	Clean up the tableware
		Eat character(s) 4-5	User-generated routine quest(s)
Mindfulness Village	Key	Mindfulness Building	Write a mind diary
	Secondary	Mindfulness character 1	Do a simple household chore
		Mindfulness character 2	Obey the screen time limit on the app use plan
		Mindfulness character 3	Thorough personal hygiene
		Mindfulness character(s) 4-5	User-generated routine quest(s)
Sleep Village	Key	Sleep Building	Keep a regular sleep schedule
	Secondary	Sleep character 1	Plan your day for tomorrow
		Sleep charcater 2	Set the alarm for the planned wake-up time
		Sleep character 3	Put the phone out of reach before going to bed
		Sleep character(s) 4-5	User-generated routine quest(s)

User Scenario (P7)

P7 began using TheApp for a period of 8 weeks, focusing particularly on the 'Regular intake of meals' routine quest **to form healthy eating habits** among four routine components of *TheApp*. Initially, P7 struggled to maintain regular meal times, but P7 soon adapted successfully.

By consistently completing the 'Calculate meal calories' quest made P7 realize that a preference for frozen and junk food was **a major obstacle to forming healthy habits**. P7 **decided to change and began preparing healthy lunch boxes** with the help of P7's parents. Particularly, the visual rewards offered by *TheApp* for completing routine quests (i.e., building upgrades and item acquisitions) greatly motivated P7. Over time, P7 **successfully maintained a balanced diet** and even occasionally prepared P7's own healthy meals, which **gave P7 great satisfaction**.

Figure 3: An example user scenario (P7) of *RoutineAid*, highlighting the goals, challenges, strategies to overcome the challenges, and the positive outcomes achieved.

interviews with three family medicine physicians, we assess the importance and relevance of health behaviors related to changes in daily routines of autistic individuals during the field study.

Our research questions are as follows:

- RQ1: How effectively does *RoutineAid* support participants' self-efficacy in their daily routine management?
- RQ2: How do participants' health behaviors manifest in their use of *RoutineAid*?
- RQ3: What changes in health behaviors were observed in participants for each routine component?

4.1 Research procedure

We conducted a nine-week study with 18 autistic participants to evaluate the effectiveness of *RoutineAid* and to investigate their engagement [2] patterns. The procedure included (1) a 60-minute introductory session, (2) a nine-week field study, and (3) a 40-minute follow-up interview (Figure 4). During the introductory session, we provided detailed information about the types of data that would be collected and how they would be managed. We then asked participants to complete an app use plan table to set goals and strategies for four routine components. We also guided participants through a separate communication channel to ask questions or express concerns during the field study, providing an accessible and immediate support system outside of the app environment. This was critical for addressing technical issues, clarifying study procedures, and increasing participant engagement by ensuring ongoing support and feedback. We started a nine-week field study with a baseline phase during the first week. In this phase, we aimed to collect participants' PA data prior to introducing *RoutineAid*. Throughout the nine-week field study, participants wore a Fitbit for continuous PA data collection. At the end of the baseline phase, they received a link to download the app through the separate channel for use in the subsequent intervention weeks. After participants completed the field study, we conducted semi-structured one-on-one interviews to understand users' detailed experiences of routine management using *RoutineAid*. We additionally interviewed three family medicine physicians (5-12 years of experience) to understand the significance of the changes in daily routines demonstrated by autistic individuals in terms of overall health. Our study was approved by the IRB at the author's institution (B-2202-736-301), and informed consent was obtained from each participant. To ensure participants' privacy,

[2] We define engagement as user experience engagement [57], highlighting aspects such as presence, focused attention, involvement.

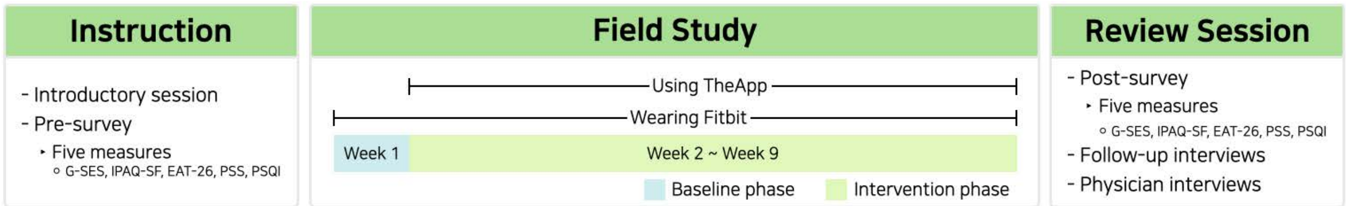

Figure 4: Overall research procedure. After completing an introductory session and a pre-survey, participants took part in a nine-week field study. "Week 1" was the baseline phase, where they only wore a Fitbit. From "week 2" to "week 9" (intervention phase), they also used *RoutineAid*. We then asked participants to complete a post-study survey and attend follow-up interviews. Finally, we conducted interviews with physicians to collect feedback on *RoutineAid* from a health behavior expert's perspective.

we anonymized the collected data by removing all personally identifiable information and assigning code names before conducting the analysis.

4.2 Recruitment

Over seven months (from July 2022 to February 2023), we recruited 20 autistic individuals (male = 16, female = 4). Our recruitment efforts included distributing leaflets to several public healthcare organizations, conducting information sessions with autistic individuals, and posting a recruitment notice on an online autism community website. The inclusion criteria for participation were as follows: (1) age range of 18-40 years old, (2) diagnosis of autism by medical experts, and (3) ability to understand the purpose of our field study and to participate independently without any assistance from parents or caregivers. To ensure stability in the participants' daily environments, our screening process also included a thorough assessment of any major life changes planned during the study period, such as relocation or changes in educational or welfare settings. We set the age range at 18-40 based on feedback from autism stakeholders involved in the design process of the app, and considering the increased need for routine management among young adults. The participants were aged between 19 and 38 years (M_{age} = 28.2, SD_{age} = 5.3). Two of the 20 participants withdrew from the study due to mental health issues and personal reasons, resulting in 18 participants completing all study procedures. Each of these participants received a compensation of $50.

4.3 Data collected

4.3.1 Measures of self-evaluation on daily life. All scales were assessed before (pre) and after (post) a nine-week field study. These measures were chosen because they reflect well-being and perceptions of daily life and represent the routines considered in *RoutineAid*.

General Self-efficacy Scale (G-SES). To evaluate the impact of the app on self-efficacy, we used the G-SES [61], designed to measure one's belief in one's ability to cope with various challenging demands in life. The G-SES comprises ten statements with a 4-point Likert scale. Higher scores signify a stronger belief in personal capability to overcome difficulties.

International Physical Activity Questionnaires-Short Form (IPAQ-SF). To assess the impact of the app on PA, we employed the IPAQ-SF [22]. The IPAQ-SF captures various intensities of activity, from walking to vigorous exercises. Participants report the frequency and duration of these activities, enabling the calculation of total weekly PA in metabolic equivalent of task (MET) minutes. This accumulated score classifies the user's activity level.

Eating Attitudes Test-26 (EAT-26). To determine the potential need for professional attention regarding eating habits, we adopted the EAT-26 [28], which consists of 26 items evaluated on a six-point scale from 0 to 5. Scores of 20 and above suggest a consultation with a medical professional.

Perceived Stress Scale (PSS). To evaluate the app's influence on perceived daily stress, we employed the PSS [19], a ten-item scale with a 5-point Likert response. Higher aggregate scores on this scale signify greater perceived stress.

Pittsburgh Sleep Quality Index (PSQI). We employed PSQI [15] to assess sleep patterns. The questionnaire encompasses seven sleep components: sleep quality, latency, duration, efficiency, medication necessity, sleep disturbances, and daytime sleep-related issues. Using ten questions, it produces a score between 0 and 21, categorizing sleep quality as 0 to 4 (typical), 5 to 10 (insufficient depth), and 11 to 21 (disorder).

4.3.2 App usage log. *RoutineAid* reflected user input, such as self-reports and game performance metrics (e.g., total score, rankings, and village development status), in real-time. This data was stored via Google Firebase [3]. For this study, we analyzed these metrics to understand user engagement with the system.

4.3.3 Fitbit log. To collect more accurate and consistent PA information, we used the Fitbit Luxe [4]. Participants could initially choose their Fitbit device from three colors: black, white, and pink. One participant found the default band material uncomfortable and was promptly provided with a band of her preferred material and design. As users' Fitbit devices can store daily totals for up to 30 days, we stored the data three times during the study period. We calculated users' daily wear time based on the heart rate endpoint, and only data collected for more than 15 hours were used for analysis. In addition, any participant whose data was not collected consecutively for more than 48 hours was excluded from the analysis. Two out of 18 participants were excluded because they had lost their Fitbit devices, leaving data from only 16 participants for Fitbit data analysis. The 15-hour threshold is more conservative than that used in previous studies using the Fitbit for PA analysis [5]. Moreover, to ensure consistent data collection, the researcher daily checked the dashboard and notified users via the separate channel if (1) the

[3]https://firebase.google.com/
[4]http://bit.ly/3qXG7QI

Table 2: Interview topics and sample questions for participants and family medicine physicians.

Interviewee	Topic	Sample Question
Participants	*RoutineAid*'s impact on daily life	What changes have you noticed in your daily life after using *RoutineAid*?
	Routine quests completion	What was the most difficult routine quest to complete, and why was it challenging?
	Long-term management insights	During your long-term use of the app, what new perspectives or methods have you discovered in managing your routines?
	Encountered challenges	What was the biggest challenge you have faced while using the app? How did you overcome it?
Family Medicine Physicians	Perceptions of study results	What do you think about the novelty and significance of our primary results?
	Views on participants' challenges	What do you think are the biggest challenges in using *RoutineAid*?
	Expectations of long-term health impact	What are your expectations for the longer-term impact of using *RoutineAid* on health behavior change?

battery level dropped below 20%, (2) the previous day's wear time was below 15 hours, or (3) there had been no synchronization for more than two days.

4.3.4 Follow-up interview. The interviews with the 18 autistic participants were conducted by the clinician on our research team and lasted an average of 47 minutes. We mainly asked the participants to share their thoughts and experiences on the following topics: (1) the overall impact of *RoutineAid* on their daily life, (2) which routine quests they usually accomplished or did not accomplish, (3) lessons learned from long-term routine management, and (4) challenges they encountered.

Subsequently, we interviewed three family medicine physicians. Prior to these interviews, we provided them with a summary of the field study results and asked for their feedback, with a particular emphasis on the following topics: (1) their perceptions of the novelty and significance of the primary findings, (2) their views on the main challenges faced by the participants in using *RoutineAid*, and (3) their expectations regarding the longer-term impact of using *RoutineAid* on health behavior changes. Each interview lasted an average of 43 minutes. Our primary objective was to understand and validate the broader impact of *RoutineAid* in the context of general health behaviors, rather than to assess the participants' specific patterns of use. Table 2 outlines the interview topics and sample questions for both autistic participants and family medicine physicians.

4.4 Analysis

We conducted paired t-tests to examine any significant differences between the pre- and post-evaluation results of the collected measures, except for IPAQ-SF, which categorizes PA levels rather than providing specific scores. We also employed Cohen's d as a measure of the effect size to examine the strength of the differences between two groups, where 0.20, 0.50 and 0.80 indicate small, medium and large effect sizes, respectively [17]. The analysis of the app usage logs was divided into two parts: the game performance analysis (e.g., quest completion rate) and the sentiment analysis of the "write a mind diary" quest (the key routine quest of Mindfulness Village), based on the BERT-based model [24], which shows highly accurate classification performance (87.07% for negative, neutral, positive emotion labels). We applied the model to classify the content of the *mind diary* into three emotional classes: negative, neutral, and positive. Additionally, two authors manually reviewed the classification results to ensure their validity. We next analyzed the Fitbit log to investigate the changes in steps and PA for each intensity (light, moderate, vigorous) of the participants. The significance of changes in participants' PA over the study period was analyzed using a one-way repeated measures ANOVA (RM-ANOVA), assuming that the sphericity condition was met. Tukey's HSD was employed for post-hoc analysis to provide further understanding through pairwise comparisons.

The interviews were recorded and transcribed, providing a detailed basis for our analysis. To analyze the interviews, three authors of this paper individually open-coded the interview analyses [13]. We then moved on to axial coding, where we organized these initial codes into 23 categories by finding connections among them. Finally, through selective coding, we narrowed these categories down to six key themes that captured the core insights about user experiences and interactions with the app. We resolved conflicts between coders' results through iterative discussions by comparing their coded r. The Cohen's Kappa measure [52] indicated scores higher than 0.72 for each category, with a maximum of 0.83 and an average of 0.77.

We denote participant quotes using "PX," which refers to "Participant Number X." For three family medicine physicians, we denote their quotes using "PhysX," which refers to "Physician Number X."

5 RESULTS

In this section, we present the findings from five measures, analyses of the app and the Fitbit logs, and the thematic analysis of the interviews, addressing each RQ. Sections 5.1.1, 5.1.2, and 5.2 address RQ1, RQ2, and RQ3, respectively.

5.1 Promoting self-efficacy and embedding health behavior

5.1.1 Enhanced self-efficacy through self-awareness of routine management. While using *RoutineAid*, participants learned the priorities for building a healthy daily routine and sought personalized approaches to accomplishing these tasks effectively. One participant stated, "*RoutineAid provided clear guidance on priorities, though tasks like 'walking for 30 minutes' initially seemed challenging. To adapt, I started getting off earlier on my commute to walk the required distance*" (P11).

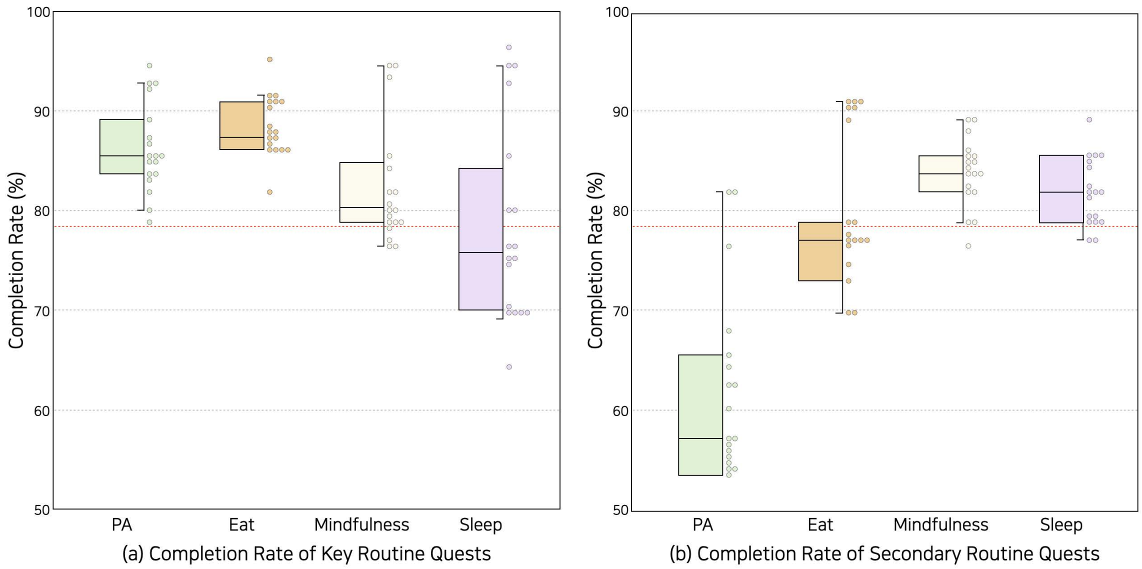

Figure 5: Box plots displaying the completion rates for key and secondary routine quests across four villages. (a) presents the completion rate for key routine quests, while (b) illustrates that for secondary routine quests. Participants were asked to self-report daily on the quests. Each quest can only be completed once per day, thus the high average daily quest completion rate indicates highly sustained participation during the study. Each dot aligned with the whiskers represents each participant's completion rate for each routine quest. The red dashed line indicates the threshold to experience all visual rewards from the app (e.g., village building upgrades, village character unlocks) during the study.

Although they recognized the importance of routine quests for their healthy daily lives, some participants found certain quests challenging to complete on their own and therefore asked for assistance from their parents or caregivers. For instance, P3 mentioned that "*I've struggled with this ['Regular intake of meals'] quest in the past and always failed. So, I asked my parents to help by ordering healthy ready-to-eat meals for me.*" Another participant also said, "*In the past, my parents helped by taking away my phone during study time. [...] While using RoutineAid, I gave it back to them again to better control my usage, which has helped me sleep earlier and wake up easier*" (P8).

As such, participants varied in the amount of time they needed to complete certain routine quests independently. However, they gradually increased their ability to address challenges autonomously in pursuit of these goals. This increased awareness and proactive change towards routines were consistently observed throughout the study as a result of the participants' sustained use of *RoutineAid*. All participants achieved more than 70% of the key routine quests, as shown in Figure 5-(a).

Beyond simply sustainably completing given routine quests, the use of *RoutineAid* had a profound effect on the participants' mindset and attitude towards their overall health. As a result, they actively sought out information, engaged in activities tailored to their health conditions, and endeavored to establish better lifestyle habits. "*I've always eaten regularly, but it was only when I started logging my food in the app that I became aware of how unhealthy my usual choices were. Once I realized that, my wife helped by packing balanced lunches for me every day. [...] I've done some PA quests to burn off calories and didn't want to undermine these efforts with poor eating choices*" (P14).

Some participants were also motivated to make healthy changes to their routines by the detailed health facts provided through the app's notifications about how to make their routines healthier. "*The health facts from the app that 'walking 10 minutes after meals lowers blood sugar by 12%' was a game changer for me and led to regular practice*" (P11). A physician commented on these findings: "*This demonstrates how specific health information from the app improves routine management and increases health literacy, which is particularly effective for autistic individuals*" (Phys2). Supporting these qualitative observations, there was a statistically significant increase in the participants' sense of self-efficacy after the study ($t(17) = 2.67$, $d = 0.63$, $p < 0.05$).

5.1.2 App narrative reflecting health behaviors in daily life. We observed that the gamified performance metrics in the app (e.g., points, village development status) effectively mirrored participants' real-world actions. As a manifestation of these metrics, a village development status, which included upgraded buildings and unlocked characters, was easily monitored on the main screen of the app. One participant mentioned, "*I struggled to complete these ['Run outdoors for five minutes' and 'Put the phone out of reach before going to bed'] quests, making it obvious on the app every time I looked at the screen*" (P4). On the other hand, another participant highlighted her success, saying, "*I loved the 'Eat Village' color palette and completed many of its quests as they aligned with my routine*" (P12). In this way, key and secondary routine quests were well reflected on the main screen of the app, reflecting the participants' actual behaviors and capturing their challenges and successes.

In addition, it appears that participants were immersed in both the narrative and story displayed on the main screen of the app, actively recording their actual behaviors. One participant mentioned,

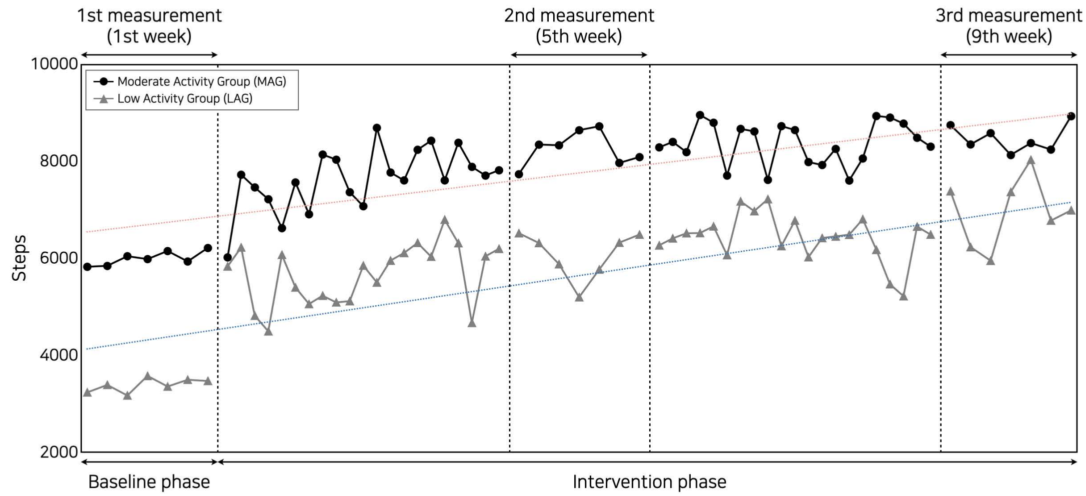

Figure 6: Step count trend during the field study. Black dots represent data from the "Moderate Activity Group (MAG)," while gray triangles represent the "Low Activity Group (LAG)." The red and blue dashed lines illustrate the linear regression trendlines for each group. Average PA data from week 1, week 5, and week 9 were used for the RM-ANOVA analysis.

"*Feeling a sense of responsibility towards the lonely [PA village] character in the PA village, I was driven to log more time on my stationary bike to ensure I could add another companion for him*" (P13). Such immersion in *RoutineAid*'s story encouraged participants to integrate *RoutineAid* into their daily lives and subsequently focus more on managing their routines. A physician shared her perspective on this deep immersion demonstrated by the participants, stating, "*The high level of engagement displayed by the participants may be related to their unique individual characteristics. While it's uncommon for the general population to be deeply influenced by such visual and narrative elements to the point of altering behaviors, these participants showed a particularly strong connection*" (Phys3).

However, while deeply engaged in the narrative of *RoutineAid*, there were some negative aspects of *RoutineAid* use. Some participants felt an overwhelming sense of responsibility, particularly for quests that they perceived as having a high barrier to entry. As shown in Figure 5-(b), this sense of burden was notably pronounced for the PA secondary routine quests that required moderate to vigorous exercise. One participant reflected, "*I wanted to collect all the characters, but tasks like 'ride a bike for 30 minutes' were just too daunting for me. I gave up halfway through and it was discouraging every time I opened the app*" (P18). Physicians emphasized that tracking and evaluating healthy routines is sufficiently demonstrated by completing key routine quests. One physician commented, "*Completing all secondary quests is ideal, but missing some doesn't imply an unhealthy day unless you're skipping key quests*" (Phys3). Another physician added, "*Many participants had disrupted routines. Through the app, they want to move from an unhealthy lifestyle to a healthier one. At this stage, it was enough to consistently do well in the key routine quests*" (Phys1).

5.2 Refining daily routines for healthier outcomes

5.2.1 Walking-oriented activities: initial boost and sustained engagement. During the study period, we analyzed changes in participants' steps and intensity-based PAs collected using Fitbit, categorized by group. Participants were grouped based on the evaluation rubric of the IPAQ-SF assessed prior to the study: "the low activity group (LAG)" and "the moderate activity group (MAG)." Patterns of movement and engagement in PA routine quests may differ when using the app, depending on PA level. Such categorization is commonly used in comparative analysis studies based on PA levels [22, 32, 63]. In line with this categorization, both groups displayed a notable increase in steps by the second week of *RoutineAid* use, and this increase in steps tended to persist until the end of the study (Figure 6). A linear regression analysis showed significant trends in the increase of steps for both groups: the adjusted R^2 = 0.55 for LAG ($F(1, 61)$ = 75.758, p <0.001) and R^2 = 0.56 for MAG ($F(1, 61)$ = 80.745, p <0.001). In other words, of the total variability in step increase, 55% can be explained by the study duration included in the model for LAG, and 56% for MAG. To extend this trend, and to investigate the impact of *RoutineAid* on changes in activity levels, we used data from the Fitbit API [5], which categorizes activities into three types of intensity-based PA: (1) light physical activity (LPA), (2) moderate physical activity (MPA), and (3) vigorous physical activity (VPA). These activity levels, measured in minutes, were assessed during the baseline phase (week 1) and then again at weeks 5 and 9 of the intervention phase.

Within the LAG, our analysis revealed significant variations in LPA levels ($F(2, 14)$ = 57.248). Using Tukey's HSD as a post-hoc test (Table 3), a pairwise comparison was performed and indicated significant increases in PA from the baseline to week 5 (p <0.001, 95% C.I. = [50.876, 108.364]) and from the baseline to week 9 (p <0.001, 95% C.I. = [59.098, 116.586]). For MPA, we observed $F(2, 14)$ = 22.619. The post-hoc test revealed notable differences between the baseline and both week 5 (p <0.05, 95% C.I. = [0.335, 25.246]) and week 9 (p <0.05, 95% C.I. = [0.958, 25.869]). VPA levels ($F(2, 14)$ = 11.442) displayed significant improvements, especially from the baseline to week 5 (p <0.05, 95% C.I. = [0.487, 7.328]). For the MAG,

[5]https://dev.fitbit.com/

Table 3: RM-ANOVA Tukey HSD post hoc comparisons for PA intensities (time units in minutes). The upper table represents the results from the Low Activity Group (LAG), while the lower table represents that from the Moderate Activity Group (MAG). Comparisons involved three measurement viewpoints: Baseline week, week 5, and week 9 ("Baseline" indicates week 1.). Statistically significant results are highlighted in bold and marked as $^{*}p$ <0.05, $^{**}p$ <0.01, $^{***}p$ <0.001.

Low Activity Group (LAG)						
PA Intensity	**Pair**	**Mean Difference**	**Std. Error**	**Adjusted p-value**	**95% Confidence Interval**	
					Lower Bound	**Upper Bound**
LPA	Baseline vs. Week 5	**79.620*****	12.189	0.000	50.876	108.364
	Baseline vs. Week 9	**87.842*****	7.994	0.000	59.098	116.586
	Week 5 vs. Week 9	8.221	13.329	0.754	-20.522	36.965
MPA	Baseline vs. Week 5	**12.791***	4.417	0.044	0.335	25.246
	Baseline vs. Week 9	**13.413***	4.427	0.033	0.958	25.869
	Week 5 vs. Week 9	0.623	5.844	0.991	-11.833	13.078
VPA	Baseline vs. Week 5	**3.908***	1.177	0.023	0.487	7.328
	Baseline vs. Week 9	3.198	1.209	0.070	-0.222	6.619
	Week 5 vs. Week 9	-0.710	1.637	0.861	-4.130	2.711

Moderate Activity Group (MAG)						
PA Intensity	**Pair**	**Mean Difference**	**Std. Error**	**Adjusted p-value**	**95% Confidence Interval**	
					Lower Bound	**Upper Bound**
LPA	Baseline vs. Week 5	**62.280****	18.666	0.007	16.704	107.856
	Baseline vs. Week 9	**63.630****	18.755	0.006	18.054	109.206
	Week 5 vs. Week 9	1.350	16.753	0.997	-44.226	46.927
MPA	Baseline vs. Week 5	10.382	4.684	0.089	-1.336	22.101
	Baseline vs. Week 9	10.110	4.442	0.099	-1.609	21.829
	Week 5 vs. Week 9	-0.272	4.814	0.998	-11.991	11.446
VPA	Baseline vs. Week 5	4.213	2.086	0.125	-0.970	9.395
	Baseline vs. Week 9	2.979	1.821	0.335	-2.203	8.162
	Week 5 vs. Week 9	-1.233	2.239	0.822	-6.415	3.949

LPA levels demonstrated $F(2, 14) = 62.469$. Post-hoc comparisons showed significant growth from the baseline at both week 5 (p <0.01, 95% C.I. = [16.704, 107.856]) and week 9 (p <0.01, 95% C.I. = [18.054, 109.206]). MPA exhibited $F(2, 14) = 16.547$, indicating overall group differences. However, specific pairwise comparisons in the post-hoc tests revealed no significant differences between individual groups. Similarly, for VPA, $F(2, 14) = 16.002$ was observed without significant contrasts in the pairwise comparisons.

Regarding MPA and VPA, both groups exhibited fluctuations in their activity levels during the intervention phase. For the LAG group, a significant difference in MPA was observed between the baseline and both weeks 5 and 9. In contrast, for VPA, significant differences were found only between the baseline and week 5. The absence of significant differences between the baseline and week 9 suggests a decrease in the upward trend observed at week 5. However, for the MAG, neither MPA nor VPA showed significant differences throughout the intervention. These variations were consistent with the feedback from participants, which showed that they felt most overwhelmed by the MPA/VPA quests (i.e., ride a bike for 30 minutes or jump rope 100 times, run outside for five minutes) of all the routine quests. One participant stated, "*MPA/VPA quests like jumping rope or running were particularly challenging for me, more so than other tasks, without offering greater rewards. In games, harder missions usually have better rewards*" (P17). From a professional perspective, one physician noted that LPA alone could be sufficient as a form of exercise to maintain health. "*For less active individuals, starting with low-intensity activities like walking is often enough for health maintenance*" (Phys1).

5.2.2 Improved nutritional changes and mealtime behaviors. While the app did not induce significant changes in the participants' primary dietary habits, and the pre- and post-assessments of the EAT-26 showed no significant differences ($t(17) = -1.84$, $d = -0.43$, $p = 0.08$), there were notable improvements in secondary but important aspects of their eating behaviors. For instance, one participant stated, "*After seeing the high calorie count in my coffee drinks in the app, I halved my intake and noticed improved sleep. I didn't expect such a change from reducing these drinks*" (P7).

Furthermore, some participants reported improvements in dining etiquette and behavior as a result of *RoutineAid*. One participant noted, "*While I was working on the 'clean up the tableware' quest, I was also cleaning up the dining area where I ate. [...] I used to have these little arguments with my mom about cleaning the dining area, but those disagreements have decreased significantly now*" (P14). P15 further mentioned, "*I've always eaten quickly because I have sensitive taste buds. When I was trying to complete the quest, I realized I spilled a lot when I cleared my seat. [...] I still eat at the same speed, but now I'm more careful not to spill my food.*" In the light of these behavioral changes, a physician stated, "*These changes, like P15's experience, show how self-awareness through the app's quests can lead to conscious behavior modification*" (Phys2).

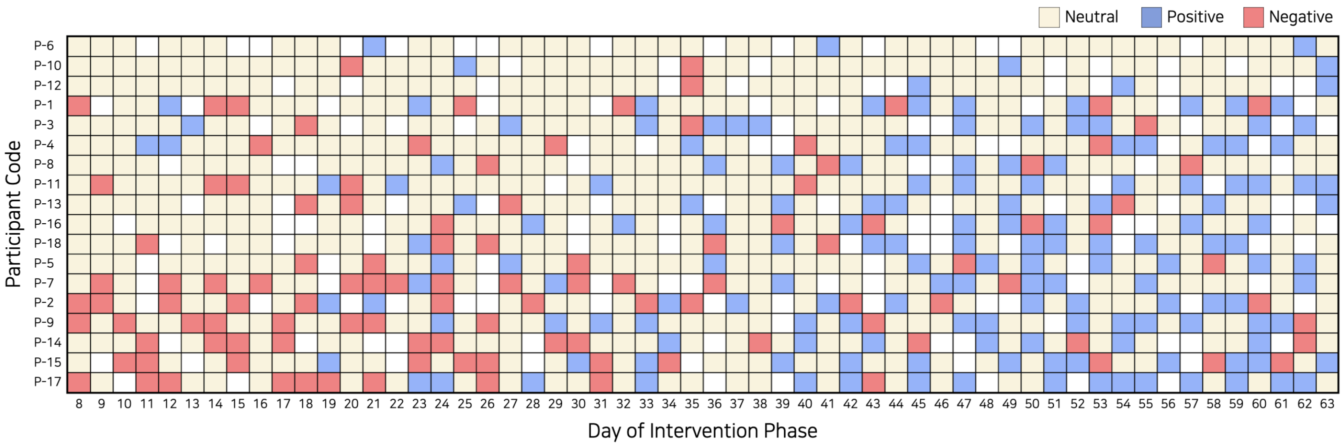

Figure 7: Visualization of sentiment analysis results for daily mind diary content. The content is classified into three emotions: "Neutral," "Positive," and "Negative" using a BERT-based model [24]. White blocks indicate days when participants did not write in their mind diary. The numbers on the x-axis (Day of Intervention Phase) indicate the respective days of the study, ranging from "Day 8" to "Day 63."

5.2.3 Enhancing emotional expression and positive transition through micro-achievements. Participants actively managed and expressed their emotions in relation to their routines by "writing a mind diary" (the key routine quest of Mind Village). At the beginning of the study, we observed that many participants reported their daily routines without emotional expression. However, as shown in Figure 7, they actively began to express their emotions from the third week. One participant commented on this trend: "*Initially, I simply reported my day briefly. As I began writing more, my thoughts and feelings became clearer, encouraging me to write more each day*" (P16). Furthermore, as participants began to express their emotions, the length of the content also increased. On average, participants wrote five words (SD = 1.39) for neutral content, 24 words (SD = 5.60) for negative content, and 11 words (SD = 7.45) for positive content in their emotional diaries. The results indicate that individuals often provide more detailed descriptions for their negative emotions, while positive emotions tend to be expressed more directly, although there were occasionally longer descriptions with large SD. Reflecting on this, a physician commented, "*The fact that participants are documenting their emotions in RoutineAid means their growing trust in it, demonstrating the importance of self-awareness and expression for healthy behavior*" (Phys2).

The majority of participants (15 out of 18; 83%) experienced a positive emotional change as a result of *RoutineAid*. This positive change was also confirmed in a pre-post comparison using the PSS, which showed a statistically significant improvement ($t(17)$ = -2.59, d = -0.61, p <0.05). This indicates that the emotional changes achieved through *RoutineAid* were not only evident in the participants' experience, but also through objective metrics. We identified three primary patterns in the participants' emotional transition towards a more positive state. The first pattern was an increase in self-esteem through the completion of routine quests, predominantly associated with PA-related quests. By engaging in the proposed PAs, participants gained confidence and noticed changes in their appearance, leading to self-satisfaction. One participant stated, "*Initially, even a '30-minute walk' was a challenge for me, but after a month, a 30-minute walk did not feel so hard. [...] As a result, I lost about 17 pounds*" (P7). The second pattern was emotional improvement through compliments from others. One participant mentioned, "*Recently, my sister said that I seem to have lost some weight. I felt really strange. We used to argue all the time because she always made fun of my appearance*" (P1). Finally, there was a positive emotional change due to the deepening of family ties among the younger participants. Some participants performed routine quests with family members, which increased family communication. One participant mentioned, "*I used to go for walks with my mom, but about a month ago, my dad started joining us 3-4 times a week. [...] Just last week, for the first time, we all went for a picnic in the park together*" (P17).

5.2.4 Sleep quality improvements driven by enhanced daytime physical activity. The results of the PSQI analysis showed a significant improvement in the participants' sleep quality ($t(17)$ = -4.08, d = -0.96, p <0.001), especially in sleep latency and total sleep time. Participants' average sleep latency decreased from 65 minutes (SD = 14.28) to 39 minutes (SD = 11.54), and the total sleep time increased from an average of 6.2 hours (SD = 1.35) to 7.3 hours (SD = 1.25). The benefits of increased PA during the day became increasingly apparent as participants moved more and slept better. Physicians affirmed the positive link between increased PA and improved sleep quality. They mentioned, "*The calories burned, the sense of accomplishment from achieving goals, and the build-up of fatigue – all of these factors, especially when moving from little or no exercise, create a conducive environment for better sleep*" (Phys1). The majority of participants reported an improvement in sleep quality over the course of the study. One participant said, "*Initially, I was skeptical about this ['put the phone out of reach before going to bed'] quest. However, as I increased my PA and followed this quest, I noticed a significant improvement in falling asleep faster and experiencing fewer disturbances during the night*" (P3).

However, some participants noted challenges alongside the reported improvements in sleep quality with increased PA. These were mainly related to personal or environmental sensitivities. One participant mentioned, "*My sleep quality has somewhat improved,* *but noise, especially like the volume fluctuations during channel changes at night, still disrupts my sleep*" (P13). Another participant stated, "*Increased daytime activity improved my sleep, but noise from living near a busy road and stress from job applications still affect it*" (P2). In response to these concerns, a physician emphasized, "*Daytime PA improves sleep quality, but environmental and social factors like noise, light, and stress are also crucial and can often offset the positive effects of increased activity*" (Phys3).

6 DISCUSSION

6.1 Leveraging interconnectivity of routine components for increased user engagement

We confirmed, through *RoutineAid*, that the four proposed routine components were not just independent, but were interactively connected (Section 5.2.2 – 5.2.4). This interconnectedness was particularly evident around the PA routine quests (e.g., improvements in dietary habits and sleep quality due to increased daytime PA). As the participants consistently documented their health status in *RoutineAid*, they recognized the relationships between these interconnected routines, leading to persistent efforts to maintain and improve these connections. This observation highlights the importance of acknowledging the interrelationships between routine components to increase user engagement with the app and tailor daily structures to individual needs.

From these findings, two design implications can be suggested that aim to emphasize the interconnectedness of routine components. First, an extension of the interactive map in *RoutineAid* is suggested. Although users can currently view the four routine villages on the main screen, each village stands alone. To further emphasize their interconnectedness, dynamic development paths between the villages should be better introduced as users maintain their routines. When activated, these paths could provide brief information highlighting how the routines are linked. This feature aligns with the visual processing strengths often observed in autistic individuals [37]. Second, it is recommended that transitional quests between villages be included in *RoutineAid*. Offering quests that connect each routine village to another can encourage users to explore and merge different routines, thereby deepening their app experiences. These design changes are expected to enrich users' understanding of the intertwined nature of the four routine components and encourage more proactive routine management.

6.2 Calibrating routine quests for improved sustainability

Throughout the study, we found that the participants were uncomfortable with being forced to report their daily completion status in *RoutineAid*, even for routine quests that they consistently failed to complete (Section 5.2.1). This was particularly evident in the PA Village, where secondary routine quests requiring MPA/VPA were most often left uncompleted. Many participants reported that the suggested exercises (i.e., cycling, outdoor running) were perceived as having a high barrier to entry, leading them to abandon these quests midway (see Figure 5-(b)). In response to these findings, the physicians involved in the interviews suggested that while completing the key routine quests was essential for building a healthy lifestyle, the secondary routine quests could be made optional, given the current state of the participants. Their perspective is noteworthy, especially considering that *RoutineAid* visually conveys a higher priority for the key routine quests than for the secondary ones (e.g., buildings vs. village characters), yet requires them to be completed equally. In other words, because *RoutineAid* requires the completion of both key and secondary routine quests without considering individual circumstances and resources, some participants felt constantly overwhelmed. While the discomfort expressed by participants should be alleviated for long-term use, there is also a need to encourage them to engage as much as possible with the key and secondary routine quests, which have been defined through repeated discussions with autism stakeholders.

To address these challenges, we could rethink the design and presentation of routine quests. By retaining the essence of each routine quest and adaptively adjusting the difficulty and goals based on the user's abilities and resources, *RoutineAid* ensures that users can still achieve a significant portion of the health goals defined in the original version of each quest. In addition, if a user routinely achieves the goals of a customized quest and desires greater challenges, the system can facilitate self-directed progression, allowing the user to progress through routine quests and ultimately engage with those in their original form.

6.3 Strengthening social connections through emotional sharing in mind diary

Through *RoutineAid*'s *mind diary*, participants actively shared their emotions, encompassing both the highs and lows of their emotional spectrum. Whether positive or negative, such transparent sharing of feelings improves self-awareness and expressiveness. This openness and willingness to communicate any emotional state appears to be crucial to cultivating healthy behavioral habits, as evidenced by the physicians' feedback. Previous research has also shown that articulating their profound feelings and experiences in writing has a positive impact on their mental and physical health [59]. Furthermore, based on their efforts to establish and maintain a self-directed and healthy routine, participants gradually experienced a shift toward positive emotions (Figure 7). Notably, the three main patterns discussed in Section 5.2.3 – an increase in self-esteem, emotional improvement through compliments from others, and strengthened family ties – suggest a clear potential for the app to increase an individual's social connectedness. Such an increase in the social connectedness of autistic individuals may give greater meaning to their daily activities and social participation, potentially helping them lead a richer and more meaningful life and ultimately contributing to an overall improved quality of life.

Two primary design directions can be suggested to improve *RoutineAid* for greater emotional transparency and positive transformation. First, providing users with encouraging pop-up messages and visual tools (e.g., an emotion progression tracker) can help them gauge how consistently they express their personal feelings. Such tools emphasize the value of sharing and articulating feelings and draw inspiration from B.F. Skinner's concept of positive feedback mechanisms [62], which advocates reinforcing desired behaviors through positive responses. Second, by incorporating a visualization feature (e.g., Figure 7), users can intuitively understand their emotion transition patterns. When combined with the *mind diary*, this feature encourages deeper self-reflection, enabling users to understand the underlying reasons for their emotional shifts.

6.4 Navigating health literacy: the essential role of environmental adjustments

Participants recognized personal challenges in performing specific routine quests and proactively created a supportive environment to increase their health literacy (Section 5.1.1). In addition, participants understood the holistic performance of routine quests and their relevance to daily life (e.g., mood improvement, energy boost). They also showed considerable interest in specific health information (e.g., app notifications for health facts). These changes empowered participants to develop and implement their actionable improvements in a self-directed manner. For instance, participants deliberately created an environment to carry out their routine quests (e.g., getting off public transportation early and walking home) or sought support from family members to supplement challenging routine components (e.g., asking family members to prepare healthy meals). Such efforts to maintain a healthy routine through environmental adaptations need to be better supported by the app.

To better support users who need environmental adaptations, the app can be designed to synchronize some features with parents' or caregivers' devices, enabling the delivery of notifications about the completion of routine quests or summaries of achievements (while maintaining the level of information detail, because autistic people do not want their detailed use of the app to be revealed to others). This integration can help parents or caregivers to be aware of the use and progress of the app and provide timely assistance where support is needed. While mobile apps are primarily designed for independent use, some users prioritize environmental adaptations to support a healthy routine, and the app should be flexible enough to operate according to the user's context. This can be an important design element to encourage autistic individuals to use *RoutineAid* to effectively achieve desired daily life outcomes.

6.5 Providing more immersive storylines for sustained engagement

We found that participants' immersion in the visual narratives provided by *RoutineAid* was a key factor in sustaining user engagement (Section 5.1.2). Participants often performed routine quests with a high personal barrier to entry in order to enrich the visuals of their routine villages (e.g., building upgrades, unlocking village characters). Based on the feedback from the interviews, physicians suggested that a preference for visual narratives could play an important role in shaping daily management habits. They also noted that while visual cues are universally appealing, the visual-centric design of the app may not consistently induce behavioral changes in all users. Our research confirmed the tendency for autistic individuals to strongly prefer visual reasoning processes [37, 46]. Furthermore, we discovered that these preferences, when carefully incorporated into the system's narrative, can be a powerful motivator for improving and maintaining healthy behaviors.

Therefore, it is important to capitalize on *RoutineAid*'s visual narrative to harness the strengths of autistic individuals and encourage sustained and voluntary user engagement. One way to do this is to introduce historical backgrounds to the routine villages and characters, and suggest additional health behavior quests linked to this context. This can motivate users to further immerse themselves in the story and to complete quests not only from routine components that they have consistently excelled at prior to using the app, but also from components that they have found challenging. Another strategy is to link the app's storyline to real-life scenarios. For instance, *RoutineAid* could mirror the user's real-time weather on its main screen and suggest related quests (e.g., sharing a YouTube link for indoor exercise on rainy days). This not only makes routine management more enjoyable for autistic individuals, but also provides a sense of legitimacy to the routine quests, enhanced by a cohesive narrative.

6.6 Limitations and future work

Although our nine-week field study confirmed the effectiveness of *RoutineAid* in supporting the daily routines of autistic individuals and offered many insights and potentials, we acknowledge two important limitations.

First, we observed that some participants had difficulty engaging in high-intensity PA (e.g., running outdoors for five minutes) and maintaining regular sleep patterns, as noted in Sections 5.1.2 and 5.2.4. In future research, we aim to explore ways to reduce the burden on participants by further segmenting or offering alternatives to increase the accessibility of high-intensity PA. Moreover, for users with underlying sleep challenges (e.g., sleep disorders), we consider incorporating features into the system that prioritize environmental adaptations, such as linking to medical services or family support interventions. Second, our participants had sufficient cognitive ability to understand our research and the key features of *RoutineAid*. Our findings and insights may not generalize to individuals who require significant caregiver support or who have more severe autism symptoms. While the current version of *RoutineAid* is tailored to individuals within a specific autism spectrum, the lessons learned may be applicable to the development of systems for a broader range of autistic individuals.

7 CONCLUSION

In this paper, we investigated how the participants use *RoutineAid*, a gamified smartphone app that supports the healthy routines of autistic individuals, and presented design suggestions for apps to support their independent and sustainable daily routine management. The results of the nine-week study highlighted the transformation of participants' health-centered actions into daily habits and the healthier outcomes that resulted from performing the given routine quests. These results further suggested salient design implications, including leveraging the interconnectivity of routine components for increased user engagement, calibrating routine quests for improved sustainability, strengthening social connections through emotional sharing in the mind diary, navigating health literacy, and providing more immersive storylines for sustained engagement. We hope that our findings and suggestions in this study will provide researchers, designers, and developers with insights into how to use and design technology to empower the characteristics and preferences of autistic individuals and improve their routines for their well-being.

ACKNOWLEDGMENTS

This research was supported by the National Research Foundation of Korea (NRF) (2021M3A9E4080780, RS-2023-00273552), and Institute of Information & Communications Technology Planning & Evaluation (IITP) (2020-0-01373).

REFERENCES

[1] Bruce M Altevogt, Harvey R Colten, et al. 2006. Sleep disorders and sleep deprivation: an unmet public health problem. (2006).

[2] Alissa N Antle, Elgin-Skye McLaren, Holly Fiedler, and Naomi Johnson. 2019. Evaluating the impact of a mobile neurofeedback app for young children at school and home. In *Proceedings of the 2019 CHI Conference on human factors in computing systems*. 1–13.

[3] American Psychiatric Association et al. 2013. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.

[4] Autism Speaks. 2013. Strategic Plan for Science 2018 - 2020. https://bit.ly/2kwA7vX. Last accessed 05 September 2019.

[5] Guilherme M Balbim, Isabela G Marques, David X Marquez, Darshilmukesh Patel, Lisa K Sharp, Spyros Kitsiou, and Sharmilee M Nyenhuis. 2021. Using Fitbit as an mHealth intervention tool to promote physical activity: potential challenges and solutions. *JMIR mHealth and uHealth* 9, 3 (2021), e25289.

[6] Abigail Bangerter, Nikolay V Manyakov, David Lewin, Matthew Boice, Andrew Skalkin, Shyla Jagannatha, Meenakshi Chatterjee, Geraldine Dawson, Matthew S Goodwin, Robert Hendren, et al. 2019. Caregiver daily reporting of symptoms in autism spectrum disorder: observational study using web and mobile apps. *JMIR mental health* 6, 3 (2019), e11365.

[7] Amanda E Bennett, Judith S Miller, Natalie Stollon, Raghuram Prasad, and Nathan J Blum. 2018. Autism Spectrum Disorder and Transition-Aged Youth. *Current psychiatry reports* 20, 11 (2018), 103.

[8] Melissa D Bittner, B Rhett Rigby, Lisa Silliman-French, David L Nichols, and Suzanna R Dillon. 2017. Use of technology to facilitate physical activity in children with autism spectrum disorders: A pilot study. *Physiology & behavior* 177 (2017), 242–246.

[9] Kristen Bottema-Beutel, Steven K Kapp, Jessica Nina Lester, Noah J Sasson, and Brittany N Hand. 2021. Avoiding ableist language: Suggestions for autism researchers. *Autism in Adulthood* 3, 1 (2021), 18–29.

[10] Louanne E Boyd, Kathryn E Ringland, Oliver L Haimson, Helen Fernandez, Maria Bistarkey, and Gillian R Hayes. 2015. Evaluating a collaborative iPad game's impact on social relationships for children with autism spectrum disorder. *ACM Transactions on Accessible Computing (TACCESS)* 7, 1 (2015), 1–18.

[11] LouAnne E Boyd and Deborah M Ward. 2013. Social Compass Curriculum. *SAGE Open* 3, 4 (2013), 2158244013507289–2158244013507289.

[12] Lal Bozgeyikli, Andrew Raij, Srinivas Katkoori, and Redwan Alqasemi. 2017. A survey on virtual reality for individuals with autism spectrum disorder: design considerations. *IEEE Transactions on Learning Technologies* 11, 2 (2017), 133–151.

[13] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[14] Shanna L Burke, Tan Li, Adrienne Grudzien, and Stephanie Garcia. 2021. Brief report: Improving employment interview self-efficacy among adults with autism and other developmental disabilities using virtual interactive training agents (ViTA). *Journal of Autism and Developmental Disorders* 51 (2021), 741–748.

[15] Daniel J Buysse, Charles F Reynolds III, Timothy H Monk, Susan R Berman, and David J Kupfer. 1989. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry research* 28, 2 (1989), 193–213.

[16] Stéphanie Carlier, Sara Van der Paelt, Femke Ongenae, Femke De Backere, and Filip De Turck. 2019. Using a serious game to reduce stress and anxiety in children with autism spectrum disorder. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 452–461.

[17] J Cohen. 1988. Statistical power analysis for the behavioral sciences. Lawrence Erlbaum. *Hillsdale, NJ* (1988), 75–108.

[18] Simonne Cohen, Russell Conduit, Steven W Lockley, Shantha MW Rajaratnam, and Kim M Cornish. 2014. The relationship between sleep and behavior in autism spectrum disorder (ASD): a review. *Journal of neurodevelopmental disorders* 6, 1 (2014), 44.

[19] Sheldon Cohen, Tom Kamarck, and Robin Mermelstein. 1983. A global measure of perceived stress. *Journal of health and social behavior* (1983), 385–396.

[20] Stanley J Colcombe, Kirk I Erickson, Paige E Scalf, Jenny S Kim, Ruchika Prakash, Edward McAuley, Steriani Elavsky, David X Marquez, Liang Hu, and Arthur F Kramer. 2006. Aerobic exercise training increases brain volume in aging humans. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 61, 11 (2006), 1166–1170.

[21] Marco Colizzi, Elena Sironi, Federico Antonini, Marco Luigi Ciceri, Chiara Bovo, and Leonardo Zoccante. 2020. Psychosocial and behavioral impact of COVID-19 in autism spectrum disorder: An online parent survey. *Brain sciences* 10, 6 (2020), 341.

[22] Cora L Craig, Alison L Marshall, Michael Sjöström, Adrian E Bauman, Michael L Booth, Barbara E Ainsworth, Michael Pratt, ULF Ekelund, Agneta Yngve, James F Sallis, et al. 2003. International physical activity questionnaire: 12-country reliability and validity. *Medicine & science in sports & exercise* 35, 8 (2003), 1381–1395.

[23] LA Croen, O Zerbo, Y Qian, and ML Massolo. 2014. Psychiatric and medical conditions among adults with ASD. In *International Meeting for Autism Research. Atlanta, GA*.

[24] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* (2018).

[25] Paul Dourish. 2006. Implications for design. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. 541–550.

[26] Lizbeth Escobedo, David H Nguyen, LouAnne Boyd, Sen Hirano, Alejandro Rangel, Daniel Garcia-Rosas, Monica Tentori, and Gillian Hayes. 2012. MOSOCO: a mobile assistive tool to support children with autism practicing social skills in real-life situations. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 2589–2598.

[27] Joseph Firth, James E Gangwisch, Alessandra Borsini, Robyn E Wootton, and Emeran A Mayer. 2020. Food and mood: how do diet and nutrition affect mental wellbeing? *Bmj* 369 (2020).

[28] David M Garner, Marion P Olmsted, Yvonne Bohr, and Paul E Garfinkel. 1982. The eating attitudes test: psychometric features and clinical correlates. *Psychological medicine* 12, 4 (1982), 871–878.

[29] Joseph Giacomin. 2014. What is human centred design? *The design journal* 17, 4 (2014), 606–623.

[30] Good Karma Applications, Inc. 2010. First Then Visual Schedule. https://apps.apple.com/us/app/first-then-visual-schedule/id355527801

[31] Katherine Gotham, Alison R Marvin, Julie Lounds Taylor, Zachary Warren, Connie M Anderson, Paul A Law, Jessica K Law, and Paul H Lipkin. 2015. Characterizing the daily life, needs, and priorities of adults with autism spectrum disorder from Interactive Autism Network data. *Autism* 19, 7 (2015), 794–804.

[32] Pedro C Hallal, Lars Bo Andersen, Fiona C Bull, Regina Guthold, William Haskell, and Ulf Ekelund. 2012. Global physical activity levels: surveillance progress, pitfalls, and prospects. *The lancet* 380, 9838 (2012), 247–257.

[33] Gillian R Hayes and Stephen W Hosaflook. 2013. HygieneHelper: promoting awareness and teaching life skills to youth with autism spectrum disorder. In *Proceedings of the 12th international conference on interaction design and children*. 539–542.

[34] Clint Heyer and Margot Brereton. 2008. Reflective agile interative design. *Proceedings of SIMTECH, Social Interaction with Mundane Technologies* (2008), 1–4.

[35] Hwajung Hong, Jennifer G Kim, Gregory D Abowd, and Rosa I Arriaga. 2012. SocialMirror: motivating young adults with autism to practice life skills in a social world. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work Companion*. 41–42.

[36] Wai Kai Hou, Francisco TT Lai, Menachem Ben-Ezra, and Robin Goodwin. 2020. Regularizing daily routines for mental health during and after the COVID-19 pandemic. *Journal of global health* 10, 2 (2020).

[37] Juan Pablo Hourcade, Stacy R Williams, Ellen A Miller, Kelsey E Huebner, and Lucas J Liang. 2013. Evaluation of tablet apps to encourage social interaction in children with autism spectrum disorders. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 3197–3206.

[38] Eunkyung Jo, Hyeonseok Bang, Myeonghan Ryu, Eun Jee Sung, Sungmook Leem, and Hwajung Hong. 2020. MAMAS: supporting parent–child mealtime interactions using automated tracking and speech recognition. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (2020), 1–32.

[39] Eunkyung Jo, Seora Park, Hyeonseok Bang, Youngeun Hong, Yeni Kim, Jungwon Choi, Bung Nyun Kim, Daniel A Epstein, and Hwajung Hong. 2022. GeniAuti: Toward Data-Driven Interventions to Challenging Behaviors of Autistic Children through Caregivers' Tracking. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–27.

[40] Lorcan Kenny, Caroline Hattersley, Bonnie Molins, Carole Buckley, Carol Povey, and Elizabeth Pellicano. 2016. Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism* 20, 4 (2016), 442–462.

[41] Julie A Kientz, Matthew S Goodwin, Gillian Rachael Hayes, and Gregory D Abowd. 2014. Interactive technologies for autism. (2014).

[42] Bogoan Kim. 2023. Supporting Independence of Autistic Adults through Mobile and Virtual Reality Technologies. In *Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–5.

[43] Bogoan Kim, Dayoung Jeong, Mingon Jeong, Taehyung Noh, Sung-In Kim, Taewan Kim, So-Youn Jang, Hee Jeong Yoo, Jennifer Kim, Hwajung Hong, et al. 2022. VISTA: User-centered VR Training System for Effectively Deriving Characteristics of People with Autism Spectrum Disorder. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*. 1–12.

[44] Bogoan Kim, Dayoung Jeong, Jennifer G Kim, Hwajung Hong, and Kyungsik Han. 2023. V-DAT (Virtual Reality Data Analysis Tool): Supporting Self-Awareness for Autistic People from Multimodal VR Sensor Data. In *36th ACM Symposium on User Interface Software and Technology*. 1–13.

[45] Bogoan Kim, Sung-In Kim, Sangwon Park, Hee Jeong Yoo, Hwajung Hong, and Kyungsik Han. 2023. RoutineAid: Externalizing Key Design Elements to Support Daily Routines of Individuals with Autism. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–18.

[46] Bogoan Kim, Daehyoung Lee, Aehong Min, Seungwon Paik, Georgia Frey, Scott Bellini, Kyungsik Han, and Patrick C Shih. 2020. PuzzleWalk: A theory-driven iterative design inquiry of a mobile game for promoting physical activity in adults with autism spectrum disorder. *Plos one* 15, 9 (2020), e0237966.

[47] Bogoan Kim, Seok-Won Lee, Hwajung Hong, and Kyungsik Han. 2019. Automated time manager: effectiveness of self-regulation on time management through a smartphone application. *IEEE Access* 7 (2019), 90891–90903.

[48] Sung-In Kim, So-youn Jang, Taewan Kim, Bogoan Kim, Dayoung Jeong, Taehyung Noh, Mingon Jeong, Kaely Hall, Meelim Kim, Hee Jeong Yoo, et al. 2024. Promoting Self-Efficacy of Individuals With Autism in Practicing Social Skills in the Workplace Using Virtual Reality and Physiological Sensors: Mixed Methods Study. *JMIR Formative Research* 8 (2024), e52157.

[49] Sung-In Kim, Eunkyung Jo, Myeonghan Ryu, Inha Cha, Young-Ho Kim, Heejung Yoo, and Hwajung Hong. 2019. Toward becoming a better self: Understanding self-tracking experiences of adolescents with autism spectrum disorder using custom trackers. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 169–178.

[50] Yiannis Koumpouros and Theodoros Kafazis. 2019. Wearables and mobile technologies in Autism Spectrum Disorder interventions: A systematic literature review. *Research in Autism Spectrum Disorders* 66 (2019), 101405.

[51] M Alexandra Kredlow, Michelle C Capozzoli, Bridget A Hearon, Amanda W Calkins, and Michael W Otto. 2015. The effects of physical activity on sleep: a meta-analytic review. *Journal of behavioral medicine* 38, 3 (2015), 427–449.

[52] J Richard Landis and Gary G Koch. 1977. The measurement of observer agreement for categorical data. *biometrics* (1977), 159–174.

[53] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. 2017. *Research methods in human-computer interaction*. Morgan Kaufmann.

[54] Daehyoung Lee, Georgia C Frey, Aehong Min, Bogoan Kim, Donetta J Cothran, Scott Bellini, Kyungsik Han, and Patrick C Shih. 2020. Usability inquiry of a gamified behavior change app for increasing physical activity and reducing sedentary behavior in adults with and without autism spectrum disorder. *Health informatics journal* 26, 4 (2020), 2992–3008.

[55] Susan R Leekam, Margot R Prior, and Mirko Uljarevic. 2011. Restricted and repetitive behaviors in autism spectrum disorders: a review of research in the last decade. *Psychological bulletin* 137, 4 (2011), 562.

[56] Yue Lyu, Pengcheng An, Yage Xiao, Zibo Zhang, Huan Zhang, Keiko Katsuragawa, and Jian Zhao. 2023. Eggly: Designing Mobile Augmented Reality Neurofeedback Training Games for Children with Autism Spectrum Disorder. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 2 (2023), 1–29.

[57] Heather L O'Brien and Elaine G Toms. 2008. What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American society for Information Science and Technology* 59, 6 (2008), 938–955.

[58] Chien-Yu Pan and Georgia C Frey. 2006. Physical activity patterns in youth with autism spectrum disorders. *Journal of autism and developmental disorders* 36 (2006), 597–606.

[59] James W Pennebaker and Joshua M Smyth. 2016. *Opening up by writing it down: How expressive writing improves health and eases emotional pain*. Guilford Publications.

[60] Phaneronsoft. 2020. Fun Routine - Visual Schedules. https://bit.ly/48e0REE

[61] Ralf Schwarzer and Matthias Jerusalem. 1995. Generalized self-efficacy scale. *J. Weinman, S. Wright, & M. Johnston, Measures in health psychology: A user's portfolio. Causal and control beliefs* 35 (1995), 37.

[62] Burrhus Frederic Skinner. 2014. *Contingencies of reinforcement: A theoretical analysis*. Vol. 3. BF Skinner Foundation.

[63] Jared M Tucker, Gregory J Welk, and Nicholas K Beyler. 2011. Physical activity in US adults: compliance with the physical activity guidelines for Americans. *American journal of preventive medicine* 40, 4 (2011), 454–461.

[64] Carl V Tyler, Sarah C Schramm, Matthew Karafa, Anne S Tang, and Anil K Jain. 2011. Chronic disease risks in young adults with autism spectrum disorder: forewarned is forearmed. *American journal on intellectual and developmental disabilities* 116, 5 (2011), 371–380.

[65] Roma A Vasa, Amy Keefer, Judy Reaven, Mikle South, and Susan W White. 2018. Priorities for advancing research on youth with autism spectrum disorder and co-occurring anxiety. *Journal of autism and developmental disorders* 48, 3 (2018), 925–934.

[66] Christiaan H Vinkers, Therese van Amelsvoort, Jonathan I Bisson, Igor Branchi, John F Cryan, Katharina Domschke, Oliver D Howes, Mirko Manchia, Luisa Pinto, Dominique de Quervain, et al. 2020. Stress resilience during the coronavirus pandemic. *European Neuropsychopharmacology* 35 (2020), 12–16.

[67] Holly R Wyatt, Gary K Grunwald, Cecilia L Mosca, Mary L Klem, Rena R Wing, and James O Hill. 2002. Long-term weight loss and breakfast in subjects in the National Weight Control Registry. *Obesity research* 10, 2 (2002), 78–82.