International Journal of Applied Positive Psychology (2024) 9:435–454
https://doi.org/10.1007/s41042-023-00133-2

RESEARCH PAPER

Evaluation of an Artificial Intelligence Enhanced Application for Student Wellbeing: Pilot Randomised Trial of the Mind Tutor

Christian Ehrlich[1] · Sarah E. Hennelly[2] · Natalie Wilde[2] · Oliver Lennon[3] · Alan Beck[3] · Hazel Messenger[4] · Kat Sergiou[4] · Emma L. Davies[2]

Accepted: 9 October 2023 / Published online: 23 October 2023

Abstract

There has been an increase in the number of UK university students disclosing mental health conditions in recent years. This paper describes the evaluation of the Mind Tutor app, an artificial intelligence based wellbeing app specifically designed for first year undergraduate students, which included a chatbot function that guided students to relevant wellbeing content. The content of the app was developed based on data about mental health and wellbeing issues reported by students and focussed on anxiety, low mood, academic study, transition to university and relationships. Two randomised controlled evaluation studies were conducted with $N = 177$ and $N = 240$ first year undergraduate students from two UK universities (the second due to delays in development work and difficulties with recruitment in the first trial). The Mind Tutor had no significant impact on student wellbeing. The study suffered from poor recruitment and retention rates. However, further research is warranted to understand factors that may increase engagement and acceptability of app based tools to increase student wellbeing.

Keywords Chatbot, evaluation study · Evaluation Study, First Year Undergraduate Students · Mind Tutor · Student Wellbeing

✉ Christian Ehrlich
cehrlich@brookes.ac.uk

[1] Oxford Brookes Business School, Oxford Brookes University, Oxford, England

[2] Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, England

[3] Syndeo Ltd., Belfast, Northern Ireland

[4] London Metropolitan University, London, UK

1 Introduction

1.1 The Need for a Student Wellbeing App

University students appear to experience more mental health problems than their non-student counterparts (Brown, 2016; Pereira et al., 2019). There has been a large increase in the number of UK university students disclosing a mental health problem in the last ten years (Johnston & Lis, 2022; Thorley, 2017). Some of the main causes of mental health difficulties in the student population are related to finances, social and academic pressures, with recent studies suggesting that around 37% of students felt their mental health had deteriorated since commencing their studies (Neves & Hillman, 2017; cf. Neves & Hewitt, 2021).

Furthermore, according to the latest figures provided by the Higher Education Statistics Agency (HESA, 2022), 5.3% of a total of 329,315 students do not proceed to the second year of their university education (17,453 students!) which is often associated with the variety of demands required for a successful transition to university life (Pennington et al., 2018). Consequently, they leave university education after their first year of study (Bermingham, 2022) resulting in significant psychological and economical losses for the students, universities and wider society. The high drop-out rate may be related to students' mental health and an overwhelming increase in demand for counselling services with long waiting times for help (Ehrnstrom et al., 2022; Thorley, 2017). Thus, one of the biggest challenges universities face, exacerbated through COVID-19 (Frampton & Smithies, 2021; Savage et al., 2020) is to find new ways to fulfil their duty in cultivating student wellbeing, particularly among first year students.

To address this mental health crisis universities have attempted to respond, amongst other measures, with recourse to a range of digital wellbeing tools (Lattie et al., 2022). However, many of these tools lack a detailed understanding of the effective elements of the digital health interventions (Lattie et al., 2019), and are not grounded in the needs of students themselves (resulting in low usage; Lattie et al., 2022), let alone specifically tailored to the needs of first year undergraduate (UG) students who face the additional challenge of transitioning successfully into university life. There is also a lack evidence for the effectiveness of student wellbeing apps in targeting and improving student wellbeing (Wicks & Chiauzzi, 2015). A further challenge is that fear of stigma about mental health may reduce help-seeking among students (Aguirre Velasco et al., 2020; Gulliver et al., 2010). Evidence suggests that the use of a specific artificial intelligence (AI) wellbeing application, mostly based on a chatbot function that delivers personalised wellbeing content, may be able reduce feelings of stigma relating to help-seeking, as chatbots are perceived as non-judgemental and neutral (Lattie et al., 2022; Lovejoy, 2019).

In the context of this mental health crisis, there is evidence that chatbots may be an appropriate way to deliver mental health interventions (Abd-alrazaq et al., 2019; Perski et al., 2017) – some of which are aimed at students (but not for first Year UG students specifically). One study of a chatbot enhanced intervention for mental health (the 'Woebot') delivered over 2–3 weeks found that university students engaged with the chatbot on average 12.14 times (Fitzpatrick et al., 2017) compared to a control

group who received information only. Those in the Woebot group had significantly lower depression scores, although a similar reduction in anxiety scores was observed in both groups (Fitzpatrick et al., 2017). A further study with university students, a chatbot ('Tess') was developed to deliver therapies such as Cognitive Behavioral Therapy (CBT) and motivational interviewing to students with symptoms of depression and anxiety (Fulmer et al., 2018). When compared to a control group, participants who interacted with Tess had significantly lower depressions, low mood and anxiety scores at two and four week follow up (Fulmer et al., 2018).

Based on the arguments presented above it can be concluded that the development of an AI enhanced wellbeing app specifically designed for first year UG students may have the potential to alleviate some of the challenges faced by universities with regards to the wellbeing and mental health of students.

1.2 The Mind Tutor

The current paper presents an evaluation of the Mind Tutor app. The Mind Tutor app integrates academic support with wellbeing support. It uses an Artificial Intelligence (AI) tool, whereby users are guided to various wellbeing content mostly through a chatbot function. The Mind Tutor was developed to address five key areas relating to student wellbeing and attainment which are: dealing with anxiety, helping with low mood, managing academic work, transitions, and relationships. The five key areas were identified in consultation with students and wellbeing services, a focus group study, and consultation with students in a lecture. Full details about the Mind Tutor development process are reported elsewhere (Davies et al., 2022).

The five topics are delivered via five specific intervention sections which are (1) general information on the five key topics; (2) relevant goal-setting exercises; (3) mindfulness exercises; (4) suggestions around specific skills and actions students could engage with to increase their wellbeing and, (5) how to positively reframe the ways students felt about anxiety, low mood etc. The general information section contained a short, text-based introduction about the importance of the five key well-being topics for students. The goal setting interventions were based on established goal setting research or techniques and include aspects like developing implementation intentions (Gollwitzer, 1999) in relation to one's goals or translating goals into learning goals (Grant & Dweck, 2003). The students were also offered the opportunity to reflect and modify their goal-striving reasons based on the goal-striving reasons framework (Ehrlich & Milston, 2022; Ehrlich, 2012; Ehrlich & Bipp, 2016). Goal-striving reasons are defined as the reasons why individuals pursue their idiosyncratic goals and differences in people's goal-striving reasons have been linked to changes in people's wellbeing. Overall, the goal striving reasons framework distinguishes between four important reasons. Of those four reasons, the following three reasons have been used within the Mind Tutor app. These were pursuing goals because of positive emotions associated with it (pleasure reasons), pursuing goals because the goal helps to make the world a better place (altruistic reasons) and pursuing goals to avoid any loss of self-esteem (self-esteem reasons). The goal-striving reasons framework has been translated into a Positive Psychology Intervention (Ehrlich & Milston, 2022) and therefore offers relevant exercises to modify one's goal-striving reasons

to pursue goals more out of pleasure reasons, more for altruistic reasons and less for self-esteem reasons. Selected exercises alongside the relevant background information on goal-striving reasons have been integrated into the Mind Tutor app. Thus, exercises on how to modify the three relevant goal-striving reasons provided students with the opportunity to think about how to modify their goals in order to make their pursuit more enjoyable, or to reduce the self-validating aspects of their goals, thus lowering the potential threat to their self-esteem.

Mindfulness content was based on the widely cited operational definitions of mindfulness as "the awareness that emerges through paying attention on purpose, in the present moment, and non-judgementally to the unfolding of experience moment by moment" (Kabat-Zinn, 2003, p. 145). 'Self-help' as well as instructor-led mindfulness based interventions (MBIs) have the potential to reduce university students' state anxiety and distress, including anxiety and depression, and to improve their wellbeing (Dawson et al., 2020). MBIs include core attention-focused practices, alongside exercises that target particular problems (Kabat-Zinn, 2003), in this case students' anxiety, mood, academic work, transitions, and relationships.

The mindfulness content therefore aimed to increase users' psychological capability to tolerate and reduce reactions to academic and personal stressors; focus their attention on academic tasks; enhance acceptance of themselves and others; and improve their emotional state using guided meditations, psychoeducation, breath work, and movement. Users' skills were developed using foundational practices, including breath counting and paying attention to physical sensations, before offering exercises to increase self-compassion and compassion to others, and noticing how thoughts and emotions manifest in the body. A number of brief core mindfulness exercises, including a three step 'breathing space', aimed to quickly settle users' mood and attention.

Some of the mindfulness exercises were also specifically designed to complement the goal setting interventions. For example, mindfulness content was created which specifically helped students to become more mindful in relation to having learning goals or in relation to having a positive view about themselves and practice self-compassion as a positive self-esteem enhancing intervention.

The sections on actions and skills focussed on giving students specific practical things to do in order to improve their well-being such as eating well or exercising regularly or the importance of sleep. The section on reframing aimed to normalise some of the negative feelings students are likely to experience due to adapting to a new situation. For example, feeling anxious is not always bad, in fact it is natural and appropriate to feel anxious sometimes.

Within the Mind Tutor app, participants interacted with a chatbot, which identified which of the five listed topics they needed help with. This was based on a data model and an algorithm which accepted the user input (any user interaction), which was defined as an "intent". The algorithm is used to 'understand' this intent (the intent could be free text or the user selecting a visual prompt) based on the data trained in the data model. Once the relevant topic is identified, the Mind Tutor directed the participant to receive one of the five interventions. Participants may complete only one or all possible interventions within that topic. They may also go back to the main menu and start looking at another topic. The Mind Tutor app was also integrated into

the students' virtual learning environment (VLE – in this case Moodle). This allowed the app to push out notifications to the students when new teaching materials were posted on the VLE. This integration aimed to facilitate the perception of students that the app was directly linked to their learning at university. Integration was developed with Moodle using exposed Application Programming Interfaces (APIs). The Mind Tutor app notified the user based on the number of "unseen activities"[1] since they last accessed their course on Moodle. To obtain this information the following APIs were used: identification of which course the student is taking; identification of the activities the user has already completed; and identification of the activities that were visible to the student. Based on this information the Mind Tutor app kept an aggregate of how many activities were unseen at the end of each day to identify if a user was active on Moodle or not. It also alerted the students if a new activity had been posted but they had not yet completed it.

The Mind Tutor app was built utilizing the pre-existing Syndeo Conversation AI platform. At its core, the platform facilitates the design and build of chatbot functionality. Interaction flows can be created and inputs captured from end-users. User inputs are captured in multiple ways – free-form text or selecting options using graphical elements such as carousels or reply buttons. The Syndeo Conversational AI platform facilitates the design and construction of these interaction flows as well as the interpreting of the user responses using natural language processing (NLP). Based on the interpretation of the user input, the platform then determines the next best action on dealing with that user prompt. The Mind Tutor app is therefore a multiplatform (covering Android and iOS) mobile app which can communicate with the Syndeo Conversational AI platform. The app was built using the Flutter framework. The key content-based features of the app included:

- A chat-based interface enabling the student to converse with the app in a conversational manner.
- Student engagement features using a number of story scenarios related to the five key topics (managing anxiety, understanding low mood, transitioning to university etc.)
- The provision of content including audio and micro articles to assist with the story scenarios.
- Functionality to assist with recording and monitoring of goal-setting activities specific to the individual student. This allowed students to record their personal goals and, as they engaged in different goal-setting activities, to revisit their goals and change/update them based on what they had learned as they continued to use the app.
- Tools to assist with breathing exercises and focus timers specifically the incorporation of a 'Pomodoro Timer' (a tool to assist with focus by selecting a task and setting a timer to focus on that task). The breathing exercise app provided visual stimulation to assist with breathing exercises.

[1] This could be announcements or any other documents posted on Moodle by the teaching team that have not been looked at by the student.

- A library functionality of content the student had engaged with while using the app, to allow them to revisit this content.

The Mind Tutor app employed an AI data model to define the intents and entities extracted from user conversations – either with the free text entered or when the student selected rich media items (such as a button, a piece of audio etc.). Tracking this data using the AI algorithm also enabled the system to track usage by the student and then provide suggested story flows (interventions) for the student to help them address one of the five key topics.

1.3 Study Aim and Objectives

The aim of the current study was to evaluate the Mind Tutor app. The study protocol was registered on the Open Science Framework (Davies et al., 2022). The primary objectives of this study were:

a) To determine the impact of the Mind Tutor on student wellbeing over a 6 week period in comparison to an inactive control group and;
b) To determine the overall feasibility of delivering a 6 week RCT to assess the effectiveness of the Mind Tutor on student wellbeing in university students.

The secondary objectives of the study were:

a) explore the impact of the Mind Tutor over a 6 week period on four additional measures which represent a more detailed view on student well-being with (a) Life satisfaction; (b) Affect (positive and negative affect); (c) Mindfulness and (d) Self-efficacy and;
b) To explore the feasibility of a 6 week RCT of the Mind Tutor in terms of.

    i) Recruitment – what proportion of students invited to the trial will consent?
    ii) Retention – what proportion of students who consent to take part will complete i) baseline and ii) follow up measures.
    iii) Engagement with the app - what is the pattern of student engagement with the app over a period of six weeks measured by (daily) interactions with app.
    iv) Acceptability of the Mind Tutor as measured by qualitative questions within follow up measures mostly around the perceived usefulness of the app by the students.

2 Methods

Trial Design A two arm, randomised, controlled trial.

Participants and Setting To be eligible to take part in the trial, potential participants had to be current first year undergraduate students enrolled on participating modules

at the two institutions. There were no other criteria for exclusion from the study. First year UG students were recruited in four first year undergraduate modules at two institutions (in Business, Technology, and Psychology). Members of the research team attended lectures/seminars/classes in person to talk about the study and invite students to take part. They were then sent a message via their module leader/class tutor with a link to the participant information sheet and the baseline survey. Participants who completed both surveys were eligible to enter a prize draw for £100 Amazon vouchers.

Ethics and Consent to Participate Prior to completing the pre-measures participants completed a participation information sheet as well as a consent form. Thus, written informed consent from study participants was obtained. The study also obtained ethic clearance from the relevant ethics committees of the participating universities.

Intervention The Mind Tutor app (content outlined above and see supplementary materials).

Control Group The control group were also completing the pre-and post-trial measures. They were offered the chance to download the Mind Tutor at end of the study period.

2.1 Outcomes

2.1.1 Primary Outcome Measures

Student wellbeing was measured with the Short Warwick Edinburgh Mental Wellbeing Scale (SWEMWBS) which enables the monitoring of mental wellbeing in a general population including first year UG students. The short form uses seven items which need to be answered on a five-point Likert scale ranging from (1) none of the time to (5) all of the time and are reported in the literature with high internal consistency (Stewart-Brown et al., 2009). Examples of items are 'I am feeling useful' and 'I've been dealing with problems well'.

2.1.2 Secondary Outcomes

The Satisfaction with Life Scale (SWLS) by Diener et al., (1985) consists of five items with strong internal validity and is widely used to measure life satisfaction. Participants were asked to rate each of the five items on a scale from (1) 'strongly disagree' to (7) 'strongly agree'. Example items are: "In most ways my life is close to ideal" or "I am satisfied with my life".

The Positive and Negative Affect Schedule (PANAS-SF; Watson et al., 1988) assessed mood. Participants answered to which degree they felt each of the described things on a scale from (1) "very slightly or not at all" to (5) "extremely" over the last two weeks. Scale contains 10 positive affect items (e.g. interested, excited, enthusiastic) and 10 negative affect items (e.g. distressed, irritable, ashamed).

Mindfulness was measured using the Cognitive and Affective Mindfulness Scale – Revised (CAMS-R; Feldman et al., 2006). Example items included "I can accept things I cannot change" and "I try to notice my thoughts without judging them".

Finally, self-efficacy was measured using the General Self-Efficacy Scale (Schwarzer & Jerusalem, 1995). This is a measure containing 10 items on a four point scale ranging from "not true at all" to "exactly true". Examples of items are: "I can always manage to solve difficult problems if I try hard enough" or "If someone opposes me, I can find the means and ways to get what I want."

2.1.3 Feasibility Outcomes

Recruitment % in relation to eligible N; Retention in terms of % lost to follow up; Engagement as measured by daily interactions with the app as well as additional feedback on the app from the post-intervention questionnaire. Here, two questions asked participants to agree or disagree whether the Mind Tutor helped them to become happier at university and how much it helped them to achieve better grades. Both questions had to be answered on the seven-point Likert scale ranging from 1 (do not agree at all) to 7 (completely agree). Another set of questions asked about the usefulness of each of the five central topics of the app (Academic study goals, relationships, Low mood/depression, transitions, and worry/anxiety). Again questions had to be answered on a five-point Likert scale ranging from (1) "not useful at all" to (5) "very useful".

Sample size We aimed to recruit a minimum of 400 participants into the study; 200 from each institution and we aimed to retain at least 50% into the follow up survey at time two – six weeks later. This sample size was based on a sample size calculation conducted in GPower for a linear multiple regression analysis to address differences between the intervention and control group on the primary outcome measure with a small to medium effect size (F2 = 0.1), 95% power and an alpha level of p = .001. This is for a model to include up to six predictors to allow for the time one score on the primary outcome measure, group (intervention/ control) and up to four other co-variates to be entered into the model (gender/institution/degree subject/age). This also allowed for incomplete cases to be dropped from the analysis if needed.[2]

Randomisation The participants who consented to take part were then allocated to the intervention group or the control group via Qualtrics survey software. The researchers were thus blinded to the condition to which the participant was allocated.

Procedures The control group were directed to a webpage thanking them for completing the measures and letting them know that they did not need to do anything further until they were sent the follow up survey at the end of the trial. The intervention group were sent to a page with instructions about how to download the app. During

[2] Please note, this was the original study protocol registered on the Open Science Framework (Davies et al., 2022). Due to a lower sample size and fewer predictors included in the analysis the power calculations have changed.

the study period, each week, a reminder to engage with the app was sent to the intervention group. At the end of the study period, an email with a link to the follow up measures was sent to all participants.

Statistical Methods To assess for differences between the primary outcome measure (SWEMWBS) between the intervention and control group a multiple linear regression model was used. Time two SWEMWBS was entered as the outcome variable, with time one SWEMWBS and group (intervention or control) entered as co-variates.

2.2 Secondary Outcomes

Similar regression models were used to assess differences between the intervention and control groups on the secondary outcome measures. Feasibility outcomes were explored using descriptive statistics.

2.3 Changes to Study from Registered Protocol

Due to delays with software development, we initially ran a six week trial of the Mind Tutor in February 2022. The study began in week four of the semester. There were a number of challenges relating to the software and recruitment of students. Therefore, it was decided that further improvements would be made to the app during summer 2022 and a second trial would be run in September 2022, targeting students at the start of their studies as originally planned. This also resulted in the possibility to allow the second evaluation study to continue over a period of eight weeks.

Feedback from the first trial of the Mind Tutor app indicated that the app needed to be revised to increase user uptake. Thus, changes were applied to the structure of the interaction flows for the second trial. This included a rework of the delivery of content (reframing, low mood, relationships and transition). These changes focused on making the engagement snappier conversations and consumption of the content in more "bite-sized" elements, adjusting the micro articles with conversational-based storytelling. The recording and changing of the goals functionality were amended to make it less repetitive, easier to access and review goals and adjust some elements such as focusing on learning and self-esteem goals. Also, a notification framework was added allowing users to receive notifications, either notifications they configure personally or notifications initiated by the system. From a more technical point, a number of key framework changes were applied to leverage some of the AI features of the Syndeo platform (invoke and unwind segue) and the gateway services framework API. Finally, throughout the whole process Syndeo implemented a number of defect fixes.

The second trial did not differ from the first with regards to research design specified in the protocol, other than the follow-up period was eight weeks instead of six weeks allowing the students to use the app for a longer period of time. In our analyses we planned to control for institution and degree subject, but due to low numbers from one institution and wide variability in degree courses we remove these as co-variates.

3 Results

3.1 Participants and Recruitment

Trial 1 The pool of potential participants was 1057 students with an even split between the two universities. Of the 1057 students, the study successfully recruited 177 students who completed the baseline questionnaire prior to the six weeks trial period. The sample consisted of 83 male participants, 92 female participants and two non-binary participants. The average age was 20.80 years ($SD = 5.18$). The split between the two UK universities was quite unbalanced with 156 students from one of the two universities. Of the 177 students, 85 were allocated to the control group and 92 were allocated to the intervention group. The numbers were uneven due to duplicate responses, which were filtered out.

Trial 2 Of the 854 potential students, the second evaluation study recruited 250 students who were willing to take part in the trial. The sample consisted of 103 men and 145 women and 1 non-binary student as well as one person who did not declare their

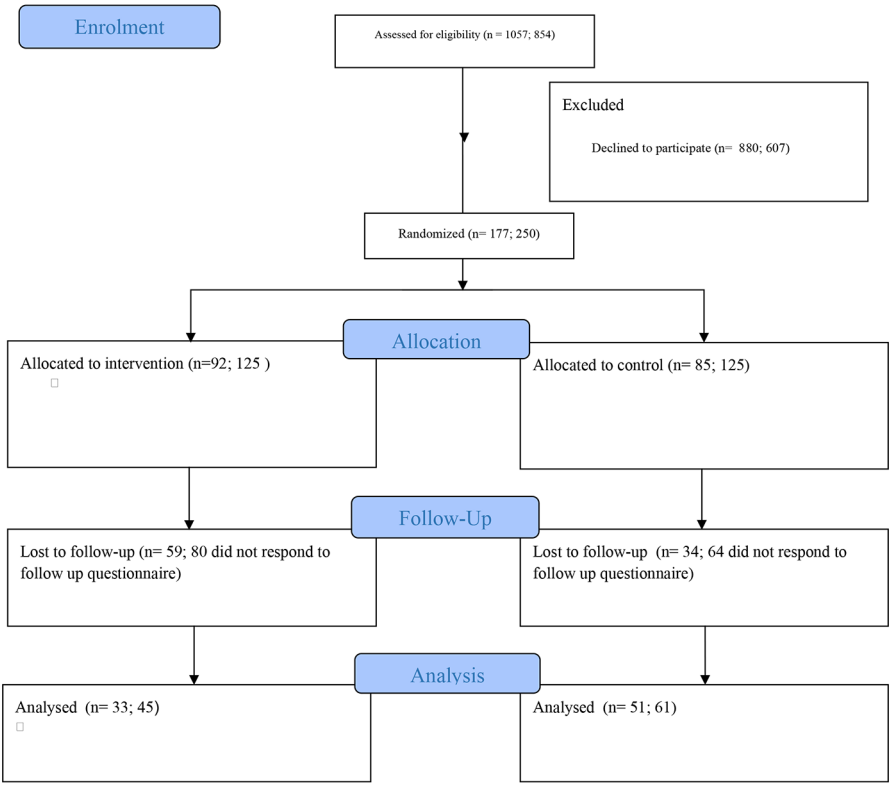

Fig. 1 Consort 2010 Flow Diagram for Trial 1 and Trial 2

gender. The average age of the sample was 20.23 years (SD = 4.96). The split between universities of the 250 recruited students was again uneven with 203 students from one university and 47 students from the second university. Of the 250 students 125 were randomly allocated to the control group and 125 were allocated to the intervention group.

3.2 Baseline Data

Baseline and follow-up measures for both trials are displayed in Table 1 along with Cronbach's alpha values for the scales and a bivariate comparisons of pre and post measures. Descriptive statistics are based on participants who completed both questionnaires (pre- and post-questionnaire) as otherwise a comparison of pre-and post-scores within each trial cannot be interpreted meaningfully. Also, no imputation of missing data was performed resulting in marginally lower numbers in the post measures (Trial 1: SWEMWBS n = 82; SWLS n = 83; PA = n = 81; NA n = 81; GSE n = 80. Trial 2: PA n = 105, all other variables n = 106). The descriptive statistics of the main study variables show that in both trials the students reported moderate levels of well-being. All mean scores are slightly above the mid-point of the scales used which – at first sight - indicates that participants in both trials reported moderate levels of wellbeing. However, well-being measures tend to be negatively skewed with averages above the mid-point of the scales (Harzer & Ehrlich, 2016). Thus, participants on both trials were overall reporting rather low levels of wellbeing as reported in the literature. The findings also show that the control and intervention group largely did not differ significantly before the trial in most study variables indicating a successful randomisation of participants. However, at trial 2 the control group scored significantly higher on mindfulness and self-efficacy before and after the trial as well as on SWEMWBS before the trial (see Table 1)[3].

3.3 Outcomes

Results of multiple regression analyses for trial 1 are shown in Table 2, with those for trial 2 in Table 3. In both trials, there were no significant differences between the Mind Tutor group and the control group on the primary outcome or secondary outcome measures when controlling for age, gender and the relevant time 1 measure of the dependent variables in question. In all cases, the only relevant significant predictor was the corresponding time 1 measure. Thus, student well-being was best (and only) predicted by their prior well-being rather than the fact that they have used the Mind Tutor app or not.

[3] Please note that in the following multiple regression analyses differences in well-being at time 1 have been controlled for as those time1 well-being measures have been included as predictors.

Table 1 Mean, standard deviation and Cronbach's alpha scores for baseline and follow up measures and t-tests comparing pre and post measures

	Baseline Trial 1				Follow up Trial 1				Baseline Trial 2				Follow up Trial 2			
	α	Mind Tutor	Control	p	α	Mind Tutor	Control	p	α	Mind Tutor	Control	p	α	Mind Tutor	Control	p
SWEMWBS	0.78	3.32 (0.65)	3.24 (0.61)	0.52	0.85	3.31 (0.62)	3.26 (0.72)	0.75	0.79	3.12 (0.66)	3.38 (0.51)	0.03*	0.84	3.16 (0.60)	3.28 (0.70)	0.34
Life Satisfaction	0.80	4.81 (1.12)	4.53 (1.16)	0.28	0.84	4.73 (1.11)	4.56 (1.25)	0.54	0.81	4.69 (1.23)	4.66 (0.94)	0.90	0.83	4.55 (1.17)	4.59 (1.04)	0.85
Positive Affect	0.87	3.33 (0.69)	3.03 (0.079)	0.08	0.82	3.19 (0.65)	2.98 (0.87)	0.26	0.86	3.33 (0.70)	3.50 (0.61)	0.19	0.84	3.01 (0.66)	3.23 (0.61)	0.09
Negative Affect	0.86	2.31 (0.85)	2.32 (0.80)	0.95	0.90	2.39 (0.97)	2.37 (0.77)	0.92	0.86	2.54 (0.83)	2.44 (0.72)	0.52	0.82	2.41 (0.67)	2.34 (0.68)	0.64
CAMS_R	0.68	2.81 (0.42)	2.87 (0.44)	0.51	0.77	2.83 (0.46)	2.86 (0.48)	0.83	0.79	2.36 (0.46)	2.51 (0.52)	0.04*	0.81	2.24 (0.51)	2.50 (0.50)	0.01*
Self-efficacy	0.86	2.86 (0.42)	2.93 (0.45)	0.46	0.90	2.92 (0.49)	2.91 (0.49)	0.97	0.84	2.82 (0.43)	3.03 (0.41)	0.01*	0.86	2.85 (0.33)	3.09 (0.46)	0.01*

Note. SWEMWBS = Short Warwick Edinburgh Mental Wellbeing Scale. Reliabilities were similar between intervention and control groups for all variables in both trials

3.4 Feasibility

Recruitment 16.7% of those eligible in trial 1 and 29.2% of those eligible in trial 2 were recruited.

Retention In trial 1, 84 students were retained at time two which equates to a retention rate of 47.7% overall. A split into intervention and control group shows that the retention rate was higher for the control group ($n=51$, 60%) compared to the intervention group ($n=33$; 35.8%). In trial 2, 106 students were retained. Of those 106, 61 were from the control group and 45 from the intervention group. Thus, as for trial one, the retention rate for the control group (49%) was higher compared to the intervention group (27%).

3.4.1 Engagement: Trial 1 and Engagement Trial 2

Engagement data for trial 1 shows that 60 users downloaded the app and produced 114 interactions with the app whereby an interaction is defined as a continuous engagement with the chatbot. Engagement data by days shows that there was a consistently low engagement with the app on most days ranging between two to five daily interactions. On six days the daily interactions were above five but not higher than seven. From an individual user perspective 53% only had one single interaction with the app, whereas 42% had between 2 and 4 interactions and the remaining 5% had more than four interactions. The university with the higher recruitment rate also reported more interactions from their students compared to the second university of this study which struggled to recruit students.

Engagement data for trial 2 shows that 84 students downloaded the app and produced 187 interactions with the app. The engagement data by date shows that engagement in the first 14 days was reasonably high with an average of 9.7 interactions a day. Engagement after the first 14 days dropped significantly with an average of 1.8 interactions a day. The engagement data from an individual user perspective was slightly better for trial 2 with 56% only had one single interaction with the app, 26% had between 2 and 4 interactions but 18% reported more than four interactions with the app. Total engagement separated by the two universities who take part in the study resembles the recruitment figures. Students from the university from which most students were recruited were reported with 139 interactions in total whereas only 24 interactions were reported from the second university (24 further interactions were from individuals who could not be allocated to one of the two universities).

3.4.2 Acceptability

The results from the general questions from those students who experienced the app revealed the following results. At trial 1, the students rated the app's helpfulness to become happier with a mean score of 3.31 ($SD=1.90$) and the degree to which the app helped them to achieve better grades with a mean score of 2.87 ($SD=1.77$). Those rating are in both cases below the mid-point of the seven point Likert scale.

Table 2 Results of six multiple regression models comparing the Mind Tutor and control group on the study outcome measures for trial 1

	SWEMWBS β	SWLS β	PA β	NA β	CAMS β	Self-efficacy β
Age	0.01	−0.08	−0.07	−0.06	−0.01	−0.04
Gender	−0.07	0.06	−0.06	0.04	−0.06	−0.06
Time 1 measure	0.73**	0.77**	0.73**	0.68**	0.71**	0.71**
Group	0.01	-. 04	0.05	0.02	0.05	0.08
$R^2(adjustedR^2)$	0.54 (0.53)**	0.61 (0.59)**	0.53 (0.51)**	0.48 (0.46)**	0.52 (49)**	0.51 (48)**

Note. SWEMWBS = Short Warwick Edinburg Mental Wellbeing Scale, SWLS = Satisfaction with Life Scale; PA = Positive Affect; NA = Negative Affect; CAMS = Cognitive and Affective Mindfulness Scale – Revised, SE = Self-Efficacy.

Table 3 *Results of six multiple regression models comparing the Mind Tutor and control group on the study outcome measures for trial 2*

	SWEMWBS β	SWLS β	PA β	NA β	CAMS β	Self-efficacy β
Age	0.01	−0.08	−0.01	0.01	0.18**	0.04
Gender	−0.02	0.11	−0.07	−0.07	0.01	0.05
Time 1 measure	0.68**	0.65**	0.63**	0.61**	0.71**	0.68**
Group	0.04	-. 05	−0.07	0.02	−0.08	−0.11
$R^2(adjustedR^2)$	0.37 (0.35)**	0.46 (0.44)**	0.42 (0.40)**	0.36 (0.34)**	0.60 (59)**	0.53 (0.51)**

Note. SWEMWBS = Short Warwick Edinburg Mental Wellbeing Scale, SWLS = Satisfaction with Life Scale; PA = Positive Affect; NA = Negative Affect; CAMS = Cognitive and Affective Mindfulness Scale – Revised, SE = Self-Efficacy,

Some of the analytics also showed that only one person took advantage of the possibility to integrate or synchronise the Mind Tutor app with Moodle. With regards to the usefulness of the five key topics of the app the student rated all five topics equally important (on the five point Likert scale) with an average score of 3.14 (SD = 1.20) for academic study goals, 3.07 (SD = 1.56) for relationships, 3.33 (SD = 1.51) for low mood, 3.75 (SD = 1.71) for transitions and, 3.68 (SD = 1.65) for worry.

The results for trial 2 are very similar. With regard to the question whether the app helped students to become happier at university the students rated the helpfulness of the app with an average of 2.88 (*SD* = 1.60). In relation to how the students perceived the app to achieve better grades at uni the students rated the app with an average of 2.77 (*SD* = 1.57). Thus, in both cases the app was rated (similarly to trial 1) below the mid-point of the scale. With regards to the usefulness of the five key topics of the app the student rated (again) all five topics equally important, with an average score of 3.74 (SD = 1.48) for academic study goals, 3.55 (SD = 1.75) for relationships, 3.69 (SD = 1.64) for low mood, 3.76 (SD = 1.62) for transitions and, 3.57 (SD = 1.56) for worry.

4 Discussion

4.1 Main Outcomes

The aim of this study was to evaluate the Mind Tutor, an AI enhanced app that was developed to improve student wellbeing. In addition, the study aimed to provide insights into aspects around the feasibility of a wellbeing app in the student context, particularly for first year UG students. With regards to the effectiveness of the Mind Tutor app, the findings show that there was no significant difference in subjective wellbeing between students who used the Mind Tutor and the control group with regards to the primary outcome variable. There were also no significant differences in secondary outcome measures between the two groups. While it may be concluded that the Mind Tutor app did not have its intended impact, it is important to note that both studies were underpowered, having failed to recruit sufficient numbers of participants. This did not allow to test for the effectiveness of the app with all relevant control variables as originally planned (for example including institution and type of degree). However, the less robust analyses with only four predictors and 80% power with a 5% error rate were sensitive enough, according to a sensitivity gpower analysis, to detect small effects (i.e. Beta weights between 0.10 and 0.29 within multiple regression analysis)[4]. Thus, the sample size was overall, for the analyses conducted, sensitive enough to detect any (small) effects due to the app.

Following on from these observations, there are several aspects that serve as an explanation as to why the Mind Tutor app did not improve student wellbeing. Like any other wellbeing intervention, using a wellbeing app requires intentional effort (Lyubomirsky et al., 2011; Perski et al., 2017; Sheldon & Lyubomirsky, 2019) and therefore requires a certain level of motivation and commitment. It should be noted that the study required students to engage with the app at a time when they needed to transition into university life – a time which for many students is highly stressful (Pennington et al., 2018). Consequently, students might have lacked resources or motivation (Antezana et al., 2022) to engage with the app alongside all the challenges that come with transitioning into university. Thus, there is a danger that students feel overwhelmed to use a stand-alone wellbeing app, to invest time into another aspect of their life – rather than focusing predominantly on how to be successful at university. This clearly highlights the fact that the Mind Tutor app needs to be positioned by universities as a means to reduce (perceived) student workload and help students reduce their stress levels rather than adding additional work. This aspect needs to be highlighted to students at an early stage and potentially through means outside of the app. Potential routes are to integrate usage of the app into the module content to give students the perception that working on their wellbeing is directly contributing to their academic work and therefore their academic success.

Another reason for the ineffectiveness of the Mind Tutor app could have been due to the design features of the app itself (Antezana et al., 2022). The Mind Tutor app offered individualised wellbeing content to students depending on their needs.

[4] A power analyses of the type "sensitivity" instead of "post-hoc" has been employed because the latter is biased due to the unknown true population effect (Yuan & Maxwell, 2005).

The Mind Tutor was also integrated into the students' learning platform but both of those features could potentially be build out further. As a result the students were not guided "instantly" to the most relevant content for them – giving the impression that they have to trawl through some less relevant content which is time consuming and might have led to low engagement with the app. Equally, due to low level of integration into the learning platform used by the students the app potentially appeared to the students as a stand-alone offer and was not sufficiently perceived as an app that helps them with their learning as well, i.e. helping them dealing better with stress due to heavy workload or assignment deadlines.

4.2 Feasibility Outcomes

Recruitment into the study was much lower than anticipated despite the known challenges with recruitment of students (Khatamian Far, 2018). In trial 1 this may have been explained by the late launch of the study – during the fourth week of the second semester rather than the first week of the first semester. The late launch was one reason why the trial was run for a second time. However, although trial 2 recruited more students, it was still underpowered for the originally planned analyses as stated in the study protocol. This prompts the question of how and when to introduce the Mind Tutor app to students in order to achieve the highest level of recruitment and retention. An interesting finding in this regard was that in both trials, retention was higher in the control groups. One explanation for this could be that the students in the intervention group who did not use the app much may not have viewed the follow up survey as relevant.

The low retention rate within the intervention group also provides some important findings relating to future research on student wellbeing. One of the key implications of this study are around the need of an integration of the Mind Tutor app into the students overall sociotechnical ecosystem (Lattie et al., 2022) which in a student context means into the students teaching and learning arrangements as well as the overall wellbeing approach within universities. In particular, this may ensure a larger uptake and a higher motivation to engage with the app, a problem that is also reported by other wellbeing apps in the student context (Lattie et al., 2022). Based on these related findings in the literature in relation to other digital well-being interventions (Lattie et al., 2022) it therefore seems reasonable to assume that the app seems to be less used if it is offered as a stand-alone feature for the students to use. Thus, further studies need to include considerations how the app can be best integrated into the learning platform of the students which should include further information on relevant studies highlighting the clear link between student's mental health and their academic performance (Bostani et al., 2014). This is to ensure that students see a more direct link between usage of the app and their academic development. Examples of a better integration could be an introductory lecture on "student wellbeing and academic achievement" as well as given the students greater insights into the relevance and importance of the five key elements of the app and how and when those key elements can be utilised within the app.

4.3 Engagement and Acceptability

With regards to engagement with the app the findings show that the improved Mind Tutor version used in trial 2 lead to a reasonably high use of the app within the first two weeks, and then usage flattened out. This is in line with previous studies which tested the effectiveness of other chatbot based wellbeing apps within a time frame of two to four weeks (Fitzpatrick et al., 2017; Fulmer et al., 2018). Thus, engagement data from the second trial suggests that an eight week long exposure of the app might be too long to keep students continuously engaged with the app. Hence, a shorter exposure period of the app of about two to three weeks might be more appropriate for the Mind Tutor app. This would have the additional benefit of exposing the students to the app at a time where they are less overwhelmed by the demands associated with settling at university. The findings from both trials suggest that approaching students in week four might be too late as lecture and seminar attendance tends to drop around this point (Doggrell, 2021; Neri & Meloche, 2007). Equally, an exposure to the app in week one might be too early as students at this point in time are overloaded with information and therefore find it difficult to engage with another wellbeing app. Bearing this in mind, the ideal time to expose students to the Mind Tutor app can be assumed to be in week three of the semester where student attendance is relatively high and the students are likely to have settled in more.

With regards to acceptability of the of the app the findings of both trials suggest that students did not perceive the app as overly helpful in relation to them being happier at university or achieving better grades. At the same time, the students rated the usefulness of the five key areas in both trials as high. This suggests that the lack of acceptability of the app might be more an issue around the delivery of the content rather than the content itself. According to Perski et al. (2017) the following features typically increase acceptability of digital-based interventions: design features, challenge, complexity, control features, credibility features, ease of use, familiarity, guidance, interactivity, message tone, novelty, narrative personalisation, and professional support features. Whilst the Mind Tutor app delivered on some of those features (pleasing design, ease of use, message tone, narrative personalisation) the integration of some of the other features mentioned above might increase acceptability of the app further.

4.4 Limitations and Future Research

The study had several limitations. As stated, both trials were underpowered with regards to the more robust research design which included originally up to six predictors and a power of 0.95 and a 1% error rate. Given the slightly less strict power settings used (especially the error rate of 5%) the likelihood for a type II error is slightly increased. However, given that both studies still had reasonably statistical power, the given sample sizes in both trials allowed for small effect sizes to be detected with still a small error rate of 5%. Thus, the likelihood that the non-significant results are due to a type II errors appear very small.

Furthermore, the analytics regarding the usage of the Mind Tutor app did not allow for any individualised analysis of the usage of the app. Thus, it was not possible to match any data about individuals and their specific use of the app with the measures used in the evaluation study. Thus future studies are needed to link changes in wellbe-

ing measures to specific patterns of usage of the app. This will give important insights into further relevant app developments to tailor the features of the app to those specific patterns of usage. One method of understanding how users interact with digital apps for health and wellbeing is the Think Aloud method, where participants are prompted to voice their thoughts while engaging with apps (Davies et al., 2017).

5 Conclusions

The Mind Tutor, a chatbot enhanced wellbeing app, had no significant impact on student wellbeing in a study that suffered from poor recruitment and retention rates. However, the two evaluation studies provide some important insights into feasibility, engagement and acceptability aspects of a digital wellbeing app in a first year UG student context.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s41042-023-00133-2.

Funding This study was funded by Innovate UK (Award number: 10003836).

Declarations

Competing Interests The authors declare the following financial interests which may be considered as potential competing interests: Syndeo Ltd has a financial interest in the commercialisation of the Mind Tutor App.

References

Abd-alrazaq, A. A., Alajlani, M., Alalwan, A. A., Bewick, B. M., Gardner, P., & Househ, M. (2019). An overview of the features of chatbots in mental health: A scoping review. *International Journal of Medical Informatics*, *132*, 103978. https://doi.org/10.1016/j.ijmedinf.2019.103978

Aguirre Velasco, A., Cruz, I. S. S., Billings, J., Jimenez, M., & Rowe, S. (2020). What are the barriers, facilitators and interventions targeting help-seeking behaviours for common mental health problems in adolescents? A systematic review. *Bmc Psychiatry*, *20*(1), 293. https://doi.org/10.1186/s12888-020-02659-0

Antezana, G., Venning, A., Smith, D., & Bidargaddi, N. (2022). Understanding what we know so far about young people's engagement with wellbeing apps. A scoping review and narrative synthesis. *Digital Health*, *8*, 20552076221144104. https://doi.org/10.1177/20552076221144104

Bermingham, J. (2022). Non-continuation summary: UK Performance Indicators. Retrieved from https://www.hesa.ac.uk/data-and-analysis/performance-indicators/non-continuation-summary#full-time. Retrieved 27.01.2023, from Higher Education Statistics Agency https://www.hesa.ac.uk/data-and-analysis/performance-indicators/non-continuation-summary#full-time.

Bostani, M., Nadri, A., & Nasab, A. (2014). A study of the relation between Mental health and academic performance of students of the Islamic Azad University Ahvaz Branch. *Procedia - Social and Behavioral Sciences*, *116*, 163–165. https://doi.org/10.1016/j.sbspro.2014.01.186

Brown, J. A. O. (2016). Student mental health: Some answers and more questions. *Journal of Mental Health*, *27*(3), 1360–0567. https://doi.org/10.1080/09638237.2018.1470319

Davies, E. L., Lonsdale, A. J., Hennelly, S. E., Winstock, A. R., & Foxcroft, D. R. (2017). Personalized Digital interventions showed no impact on Risky drinking in young adults: A pilot randomized controlled trial. *Alcohol and Alcoholism*, 1–6. https://doi.org/10.1093/alcalc/agx051

Davies, E. L., Hennelly, S., & Ehrlich, C. (2022). Feasibility and effectiveness of an artificial intelligence enhanced application for student wellbeing: Pilot trial of the Mind Tutor. doi:https://doi.org/10.17605/OSF.IO/Z9GNX

Dawson, A. F., Brown, W. W., Anderson, J., Datta, B., Donald, J. N., Hong, K., & Galante, J. (2020). Mindfulness-based interventions for University students: A systematic review and Meta-analysis of Randomised controlled trials. *Applied Psychology: Health and Well-Being*, *12*(2), 384–410. https://doi.org/10.1111/aphw.12188

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with Life Scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Doggrell, S. A. (2021). Quantitative study of lecture attendance and the association between this attendance and academic outcomes for nursing and nonnursing students in an introductory pathophysiology course. *Advances in Physiology Education*, *45*(4), 651–660. https://doi.org/10.1152/advan.00037.2021

Ehrlich, C. (2012). Be careful what you wish for but also why you wish for it – goal-striving reasons and subjective well-being. *The Journal of Positive Psychology*, *7*(6), 493–503. https://doi.org/10.1080/17439760.2012.721382

Ehrlich, C., & Bipp, T. (2016). Goals and subjective well-being: Further evidence for goal-striving reasons as an additional level of goal analysis. *Personality and Individual Differences*, *89*, 92–99. https://doi.org/10.1016/j.paid.2015.10.001

Ehrlich, C., & Milston, S. (2022). *Happiness through goal setting: A practical guide to reflect on and change the reasons why you pursue your most important goals in life* (1 ed.). Routledge.

Ehrnstrom, C., Blankenheim, A., Bottiglio, M., Tomatz, M., West, A., Bryant, T., & Gerberich, S. (2022). Embedding counselors within departments across campus can meet increased demand. *Student Affairs Today*, *25*(3), 6–7. https://doi.org/10.1002/say.31083

Feldman, G., Hayes, A., Kumar, S., Greeson, J., & Laurenceau, J. P. (2006). Mindfulness and emotion regulation: The Development and initial validation of the cognitive and affective mindfulness scale-revised (CAMS-R). *Journal of Psychopathology and Behavioral Assessment*, *29*(3), 177. https://doi.org/10.1007/s10862-006-9035-8

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Ment Health*, *4*(2), e19. https://doi.org/10.2196/mental.7785

Frampton, N., & Smithies, D. (2021). *Life in a pandemic*. Retrieved from https://www.studentminds.org.uk/lifeinapandemic.html

Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2018). Using psychological Artificial Intelligence (Tess) to relieve symptoms of depression and anxiety: Randomized Controlled Trial. *JMIR Ment Health*, *5*(4), e64. https://doi.org/10.2196/mental.9782

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*(7), 493–503. https://doi.org/10.1037/0003-066X.54.7.493

Grant, H., & Dweck, C. S. (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology*, *85*(3), 541–553. https://doi.org/10.1037/0022-3514.85.3.541

Gulliver, A., Griffiths, K. M., & Christensen, H. (2010). Perceived barriers and facilitators to mental health help-seeking in young people: A systematic review. *Bmc Psychiatry*, *10*(1), 113. https://doi.org/10.1186/1471-244X-10-113

Harzer, C., & Ehrlich, C. (2016). Different forms of life satisfaction and their relation to affectivity. *Edorium J Psychol*, *2*, 8–13.

HESA (2022, March, 17). Non-contiuation: UK Performance Indicators. https://www.hesa.ac.uk/data-and-analysis/performance-indicators/non-continuation

Johnston, C., & Lis, R. (2022). Coronavirus and higher Education students; 25 February to 7 March 2022. Retrieved from https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandwellbeing/bulletins/coronavirusandhighereducationstudents/25februaryto7march2022

Kabat-Zinn, J. (2003). Mindfulness-based interventions in Context: Past, Present, and Future. *Clinical Psychology: Science and Practice*, *10*(2), 144–156. https://doi.org/10.1093/clipsy.bpg016

Khatamian Far, P. (2018). Challenges of Recruitment and Retention of University Students as Research Participants: Lessons learned from a pilot study. *Journal of the Australian Library and Information Association*, *67*(3), 278–292. https://doi.org/10.1080/24750158.2018.1500436

Lattie, E. G., Adkins, E. C., Winquist, N., Stiles-Shields, C., Wafford, Q. E., & Graham, A. K. (2019). Digital Mental Health Interventions for Depression, anxiety, and enhancement of Psychological Well-being among College students: Systematic review. *Journal of Medical Internet Research*, *21*(7), e12869. https://doi.org/10.2196/12869

Lattie, E. G., Cohen, K. A., Hersch, E., Williams, K. D. A., Kruzan, K. P., MacIver, C., & Mohr, D. C. (2022). Uptake and effectiveness of a self-guided mobile app platform for college student mental health. *Internet Interventions*, *27*, 100493. https://doi.org/10.1016/j.invent.2021.100493

Lovejoy (2019). The UX of AI Retrieved from. Available at: https://design.google/library/ux-ai/

Lyubomirsky, S., Dickerhoof, R., Boehm, J. K., & Sheldon, K. M. (2011). Becoming happier takes both a will and a proper way: An experimental longitudinal intervention to boost well-being. *Emotion (Washington D C)*, *11*(2), 391–402. https://doi.org/10.1037/a0022575

Neri, F., & Meloche, Y. (2007). The impact of lecture attendance on academic performance in a large First Year Economics Course. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.975573

Neves, J., & Hewitt, R. (2021). *Student Academic Experience Survey*. Higher Education Policy Institute. Retrieved from https://www.hepi.ac.uk/wp-content/uploads/2021/06/SAES_2021_FINAL.pdf

Neves, J., & Hillman, N. (2017). *Student Academic Experience Survey*. Higher Education Policy Institute. Retrieved from https://www.hepi.ac.uk/wp-content/uploads/2017/06/2017-Student-Academic-Experience-Survey-Final-Report.pdf

Pennington, C. R., Bates, E. A., Kaye, L. K., & Bolam, L. T. (2018). Transitioning in higher education: An exploration of psychological and contextual factors affecting student satisfaction. *Journal of Further and Higher Education*, *42*(5), 596–607. https://doi.org/10.1080/0309877X.2017.1302563

Pereira, S., Reay, K., Bottell, J., Walker, L., Dzikiti, C., Platt, C., & Goodrham, C. (2019). *University Student Mental Health Survey 2018*. Retrieved from.

Perski, O. A., Blandford, A., West, R., & Michie, S. (2017). Conceptualising engagement with digital behaviour change interventions: A systematic review using principles from critical interpretive synthesis. *Transl Behav Med*, *7*, 1613–9860. https://doi.org/10.1007/s13142-016-0453-1

Savage, M. J., James, R., Magistro, D., Donaldson, J., Healy, L. C., Nevill, M., & Hennis, P. J. (2020). Mental health and movement behaviour during the COVID-19 pandemic in UK university students: Prospective cohort study. *Mental Health and Physical Activity*, *19*, 100357. https://doi.org/10.1016/j.mhpa.2020.100357

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

Sheldon, K. M., & Lyubomirsky, S. (2019). Revisiting the sustainable happiness Model and Pie Chart: Can Happiness be successfully pursued? *The Journal of Positive Psychology*, *16*(2), 145–154. https://doi.org/10.1080/17439760.2019.1689421

Stewart-Brown, S., Tennant, A., Tennant, R., Platt, S., Parkinson, J., & Weich, S. (2009). Internal construct validity of the Warwick-Edinburgh Mental Well-being scale (WEMWBS): A Rasch analysis using data from the Scottish Health Education Population Survey. *Health and Quality of Life Outcomes*, *7*(1), 15. https://doi.org/10.1186/1477-7525-7-15

Thorley, C. (2017). *Not by degrees: Improving student mental health in the UK's universities* Retrieved from https://www.ippr.org/publications/not-by-degrees

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *J Pers Soc Psychol*, *54*(6), 1063–1070. https://doi.org/10.1037//0022-3514.54.6.1063

Wicks, P., & Chiauzzi, E. (2015). Trust but verify' – five approaches to ensure safe medical apps. *BMC Medicine*, *13*(1), 1–5. https://doi.org/10.1186/s12916-015-0451-z

Yuan, K. H., & Maxwell, S. (2005). On the Post Hoc Power in Testing Mean differences. *Journal of Educational and Behavioral Statistics*, *30*(2), 141–167. https://doi.org/10.3102/10769986030002141