Behaviour & Information Technology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tbit20

Personalisation and Recommendation for Mental Health Apps: A Scoping Review

Paul Matthews & Clemence Rhodes-Maquaire

BEHAVIOUR & INFORMATION TECHNOLOGY
https://doi.org/10.1080/0144929X.2024.2356630

Personalisation and Recommendation for Mental Health Apps: A Scoping Review

Paul Matthews and Clemence Rhodes-Maquaire

School of Computing and Creative Technologies, University of the West of England, Bristol, UK

ABSTRACT

Personalisation, which tailors to individual preferences, is considered a possible route for improving engagement with digital mental health (DMH) products. Despite claims about the presence and importance of personalised features, evidence about their extent and impact is limited. The study objective was to review evidence from published research to investigate and characterise the contribution of personalisation to engagement and effectiveness.

Research papers were retrieved using keywords, with 139 papers being fully examined and 61 meeting our eligibility criteria. Most of the eligible articles reviewing DMH systems have weak to intermediate forms of personalisation (45). Only nine were coded as having strong, adaptive personalisation. Of the 40 articles which evaluated the personalisation effectiveness, 28 were qualitative indicating user preference for personalised features. The 16 controlled, quantitative designs lacked a non-personalised intervention, making it difficult to determine the added value. Effect sizes calculated from available data showed minimal differences in effectiveness compared to non-personalised apps.

Our review indicates mixed evidence of personalisation's performance in DMH interventions. Generally, there is a lack of good quality evidence to isolate specific contributions of personalisation. Opportunities were identified for improved evaluation, however caution is required when implementing more sophisticated methods of DMH personalisation.

ARTICLE HISTORY
Received 29 May 2023
Accepted 28 April 2024

KEYWORDS
Digital mental health; mental health; personalisation; recommender systems; applications; adaptive interventions

1. Introduction

1.1. Rationale

Many mobile apps exist to help those with general well-being or specific mental health issues and the consumer base is predicted to grow. Apps can offer a sense of anonymity and privacy for people in a society in which there is often still a stigma surrounding mental ill health (Lindow et al. 2020). Apps are fulling a growing need for people who are unable to access face-to-face support, subject to long service wait times, or whose symptoms are mild. But it remains unclear how much they really help over long periods. While a little more evidence of positive impact is now available, at least for certain conditions and groups (e.g. Lattie et al. 2020; Weisel 2019) it remains a little scattered and prone to publication bias.

Low engagement and attrition amongst users is an often cited problem in Digital Mental Health (DMH) products, with many people abandoning them before they can really help in improving mental wellness and outlook (Borghouts et al. 2021; Matthews, Topham, and Caleb-Solly 2018). High levels of abandonment have been associated with a low perceived personal relevance (Jarrahi, Gafinowitz, and Shin 2018). User reviews of mental health apps have indicated possible reasons for such low engagement, for instance stating that developers should better consider user experience and usability when designing apps to earn better adherence, and hence overall effectiveness (Alqahtani and Orji 2020). Personalising the services, features and touch points of an app is consequently usually assumed to lead to higher engagement, and thereafter to impact, mirroring the causal logic of engagement with face-to-face support programmes. This research aims to investigate these assumptions and is to our knowledge the first review to focus on personalisation in digital mental health design and evaluation where the products cover a broad spectrum of intended audiences and mental health conditions.

Before further outlining our research aims, we will define personalisation, its posited benefits and present evidence from other reviews as to its application and usefulness.

CONTACT Clemence Rhodes-Maquaire clemence2.rhodes-maquaire@live.uwe.ac.uk School of Computing and Creative Technologies, University of the West of England, Bristol, UK

1.2. Definitions and presumed benefits of personalisation

Hornstein et al. (2023) define personalisation as 'purposefully designed variation between individuals in an intervention's therapeutic elements or its structure'. They differentiate personalisation from customisation, usage, interactivity, and group-based adaptations and suggest four mechanisms of personalisation: user choice, provider choice, rule based and machine learning (ML) based. This characterisation is useful in distinguishing elements that are either under the control of the designer or the consumer, and where there may also be a trade-off (e.g. more provider direction may reduce user choice).

A useful typology of personalisation was developed by op den Akker, Jones, and Hermens (2014) and applied for digital fitness products by Monteiro-Guerra et al. (2020), helping to define the differences between what is often active, machine learning-based DMH (adaptation or self-learning), human-mediated, personalised support (interaction and feedback) and coarser, rule-based DMH (goal setting, user targeting). This taxonomy was used as a starting point in the present study and more information will be found in the coding and validation section.

Oinas-Kukkonen (2018) proposed that personalisation should be understood as a continuum from shallow or weak (tailoring and broad-brush filtering) to deep or strong (contextual understanding, continuous adaptation). We also adhere to these distinctions and see user targeting with an underlying simple rule-based model as weak to intermediate personalisation, and online learning with an underlying, continuously updated user model as strong. This conception is a little broader than that of Hornstein et al. (2023), in that 'weak personalisation' in our conception is often synonymous with segmentation at a group level.

In their recent work, Valentine, D'Alfonso, and Lederman (2023) have summarised some presumed benefits and ethical dilemmas associated with the integration of personalisation features in mental health applications. The authors note that the advantages of employing recommender systems and personalisation include the reduction of 'choice overload' for users, improving the therapeutic alliance, and supporting the user to self-manage. Nevertheless, this progression also gives rise to ethical concerns around privacy, a lack of transparency and control over a user's interaction record. Like many similar summaries, personalisation in Valentine et al. paper is often presented as an *a priori* benefit and some of their examples of possible personalised messaging based on captured contextual data may be perceived by users as intrusive.

In terms of people's perceptions, overt enthusiasm among current and potential DMH users as well as product providers for more personalised experiences is often reported. As Cabrita et al. (2018) note, 'each individual is unique, and dynamic, in a sense that a strategy that works for one, might not work for another' and so personalisation is hypothetically a pathway to improved engagement and enabling users to take control of their wellbeing journey. This somewhat parallels the current calls for more personalised treatment pathways in face-to-face support. And in the digital space there are a prefusion of predictive and persuasive analytical techniques that have the potential to support stronger personalisation, incorporating on-device contextual cues such as location, health sensor data and time of day alongside app interaction data (Oinas-Kukkonen 2018; Insel 2017).

1.3. Existing reviews of personalisation impact

Existing reviews have looked at specific MH conditions and contexts, research study adherence, or have addressed personalisation with a more face-to-face framing.

Hornstein et al. (2023), for instance, conducted a comprehensive systematic review focusing on personalisation in DMH interventions specifically for depression. They identified 139 eligible papers, of which 94 were related to digital interventions. The researchers categorised their analysis into three primary personalisation aspects: a typology of personalisation with terms like usage, customisation, interactivity and group-based adaptation; the dimensions of personalisation and the mechanisms for personalisation. These dimensions encompassed elements such as content, order, guidance and communication. Among the digital interventions, approximately 66% were found to incorporate some form of personalisation. Around 32% employed personalisation mechanisms, with personalised communication constituting 30% (25% involving type and 4% focusing on order). Interestingly, when considering interventions with personalisation mechanisms, 69% implemented just a single dimension of personalisation. The overarching conclusion drawn from this review was that the work reviewed merely scratched the surface of the potential scope for personalisation. The findings demonstrated the lack of multi-dimensional personalisation and reveal the need for increased incorporation of advanced machine learning techniques.

Moe-Byrne et al (2022) reviewed evidence for tailored DMH interventions in the workplace, finding seven studies where a control was compared to a digital intervention. Tailoring was achieved by user screening or in-app exercises (four studies), by user characteristics and goals (one study) or by sleep data (one study). Four of the counted studies did not show significant effects of the treatment in reduction of anxiety or depression. One showed long-term positive impact on depression, and one showed significant impacts on depression up to one month post-treatment. They concluded that tailored interventions *could* work for those with higher levels of psychological distress, but less clear impact was shown for the general working population.

Additional reviews of personalisation impact are less related to the present work but still of some relevance. For personalised *face-to-face* interventions, O'Cillin (2022) reviewed nine quantitative and twelve qualitative studies. They found strong engagement (low drop-out rates) for the personalised interventions, and significant treatment effectiveness in six of the nine quantitative studies. Participants in seven of the qualitative studies expressed subjective recovery and there were strong themes of the importance of therapeutic alliance/collaborative engagement. In four studies, participants expressed ideas for improving engagement.

A review of more general mhealth app evaluations by Jakob et al (2022) found that 16% of the 97 studies from 8 health domains (including mental health) reported that personalisation contributed to study adherence – though it was not reported if the apps in the remaining studies *did* have personalisation features but these did not contribute to adherence.

1.4. Objectives

The reviews above reveal considerable interest in, and implementation of, personalisation techniques of one form or another, and suggest the potential for improved retention, adherence and impact of more personalised digital services. But they also highlight potential gaps in our knowledge of what forms of personalisation are being used, and perhaps more importantly, how personalised interventions improve (or not) on more standardised treatments.

Given these uncertainties, it is important and timely to review recent evidence of the kinds of personalisation being designed and evaluated across the spectrum of mental health conditions, their relative effects on adoption and ongoing usage and any subsequent improvements in mental health among user populations. The findings from this scoping review are intended to expand the knowledge base and identify research and development opportunities in digital mental health (DMH) field for researchers, app and service providers.

The research questions that framed this review were:

1. What different approaches to personalisation and recommendation have been used in DMH research and research-based products?
2. Does personalisation and recommendation in MH apps lead to better user engagement and outcomes?
3. What gaps and opportunities exist for better understanding and promoting effective use of personalisation and recommendation?

2. Methods

This study used the PRISMA framework (Tricco et al. 2018) as a guideline for the scoping review.

2.1. Search strategy

Research papers were retrieved between August 2022 and August 2023 with the articles sourced from Scopus, ACM Digital Library, Google Scholar and IEEE Explore databases. The articles from Google Scholar were frequently found to either be irrelevant or duplicates of ones found in the other databases. The titles and abstracts of the articles were read, and relevant cases were added to the collection to identify duplicates from the other author's search. This search strategy is shown in Figure 1.

2.2. Eligibility criteria

The eligibility criteria when searching for journal articles were carefully applied to ensure they linked with the research questions. Just those papers published in the last 11 years (2012–2023 inclusively) were included. This covers the period in which digital interventions incorporating personalisation were more common in the literature. Articles that were reviews or theoretical work were not included in the articles collected, nor were purely predictive or diagnostic studies as these did not include mental health interventions. Furthermore, papers needed to have a substantial component of primary research connected to personalisation/recommendations and not be predominantly secondary literature review or recommendations for future research. Initial searches produced large volumes of articles, and many were not applicable to the study. These included articles regarding personalisation of prescription drugs or personalised face to face mental

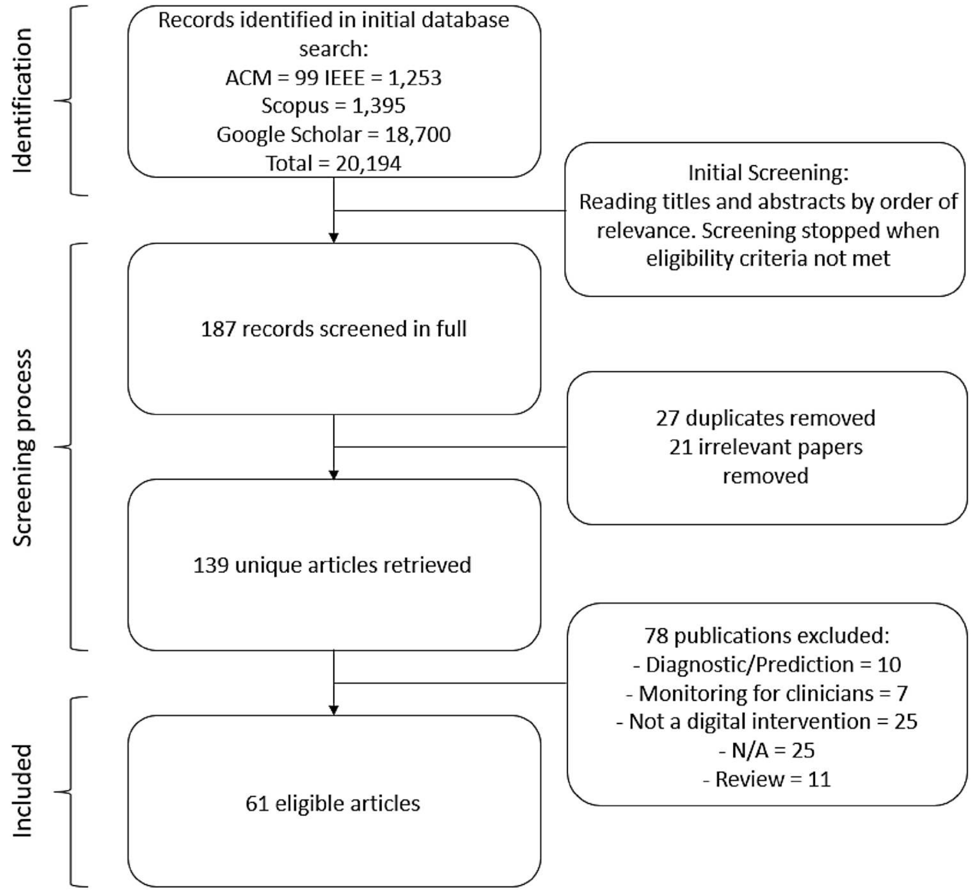

Figure 1. PRISMA chart.

health interventions. Only studies that were written in English were included in this review but there was no restriction regarding study setting or country of publication.

2.3. Search

The search strategy included the main elements and synonyms of personalisation as well as different mental health problems. The search strategy used for the IEEE database, for example, were ('All Metadata': personalis* OR 'All Metadata': personaliz* OR 'All Metadata': tailor* OR 'All Metadata': recommender OR 'All Metadata': recommendation) AND ('All Metadata': 'mental health' OR 'All Metadata': 'digital mental health' OR 'All Metadata': anxiety OR 'All Metadata': depression OR 'All Metadata': loneliness OR 'All Metadata': stress OR 'All Metadata': psychosis OR 'All Metadata': wellbeing OR 'All Metadata': schizophrenia OR 'All Metadata': ptsd OR 'All Metadata': panic OR 'All Metadata': bipolar OR 'All Metadata': ocd OR 'All Metadata': adhd OR 'All Metadata': dissocial). The * within this allows different derivatives of the terms shown with the brackets demonstrating the different options that must be included.

2.4. Study selection

As shown in Figure 1, the process of selecting articles for review consisted of three steps. Throughout the process the two authors worked on two databases each, continuously discussing the process. The databases were ordered by relevance to improve time management due to the large volume of articles, once it was shown that articles were not relevant to the eligibility criteria the screening process ended. Both authors then went through half of the articles each, reading them in full, to identify further ineligible articles alongside assigning eligible articles to the criteria described above. The N/A for publications excluded in Figure 1 describes a small group of articles that did not match with the other categories chosen. It included CONSORT documents, articles found to not have mental health as a primary outcome, those that did not include the full articles and those that did not include personalisation/recommendation of the apps.

2.5. Coding and validation

139 articles were chosen for full review and these were then read in full to decide on whether they matched the full criteria. Those that were deemed eligible were then coded into their study type, study rigour, level of personalisation and type of personalisation. The type of personalisation was based on the taxonomy of op den Akker et al. (2014) and Monteiro-Guerra et al. (2020) but adapted by us during the coding. This is shown in Table 2. The level of personalisation was weak, intermediate and strong, with the authors agreeing criteria for allocation to these categories as shown in Table 1.

Each author worked through half of the journals using the eligibility criteria, then to ensure intercoder reliability was high 10 papers were randomly selected and *r* reliability was calculated. The Kappa's Cohen for agreement was 0.75 which showed an acceptably high level of agreement between the authors. Any disagreements were discussed, and a decision was made about which classification aligned the most with the articles. Once this was completed it was found that 61 articles met the eligibility criteria.

3. Results

All the articles were encoded and classified in full and Table 3 shows the results. 65.6% of the articles were considered more useful, incorporating design and build with evaluation or being purely evaluative. These articles were used to determine if personalisation/recommendation improved user engagement within the applications. The papers in this category often provided limited information on the design and build of the recommendation techniques and it was difficult to identify the level of personalisation used in practice. For the evaluative parts, these were both qualitative and quantitative and were encoded according to the study rigour, describing the quality of the evaluation designs – this included sample sizes, reproducibility, detail of the methods and the existence of controls. The control groups were either treatment as usual for those who were receiving counselling alongside the use of an application, or they were waitlisted and instructed to not use the application until the end of the study.

Most of the articles adopted user targeting as the personalisation method, which meant personalisation/recommendation was often based on questionnaires or onboarding option sets administered at the sign-up point, e.g:

> we will use the information provided by the answers to the pre-test questionnaire, as these answers give an evidence of the strengths and needs of a caregiver [user] before starting to use the intervention plan proposed by the app.
>
> (Ferré-Bergadà et al. 2021, 3, our brackets)

Alternatively, assessments and interventions were targeted to people based on demographic or clinical categories, e.g.

Table 1. Personalisation level (inspired by Oinas-Kukkonen 2018).

Weak	Broad-brush or very basic user targeting, or where there are few content/intervention options to determine personalisation
Medium	Finer grained user targeting, adaptation or self-learning
Strong	Multiple personalisation methods used; Personalisation to the individual level; Responsive to changes in individuals' state / task / location
Not Applicable	No real personalisation or not stated

Table 2. Personalisation types (op den Akker, Jones, and Hermens 2014) with personalisation concepts from Monteiro-Guerra et al. (2020). Original definitions in normal type, our own application and interpretation in bold type.

Type	Description
Feedback	Presenting individuals with information about themselves, obtained during assessment or elsewhere. Might be from other users, admins or automated
Inter-human interaction	Support for any form of interaction with other real human beings – though if mainly in the form of feedback use that instead
Adaptation	Adaptation 'attempts to direct messages to individuals' status on key theoretical determinants (knowledge, outcome expectations, normative beliefs, efficacy and/or skills) of the behaviour of interest' – behaviour or temporal-based changes in recommended activities. App alters as time goes on based on user activities
User Targeting	User targeting attempts to increase attention or motivation to process messages by conveying, explicitly or implicitly, that the communication is designed specifically for 'you' – personalisation based on user profiling through e.g. initial survey or menu choices. Information from user used to determine the content presented
Goal Setting	Goal setting is a technique used to present the user with short-term, as well as long-term goals that can instil a feeling of progress over the course of an intervention or the day. Goal setting is a tailoring concept that can only be used in combination with feedback
Context Awareness	A system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task – use adaptation in most cases
Self-learning	A self-learning application is able to update its internal model of the user by recording and learning from the various interactions the user has with the application. (ML-based personalisation). use adaptation instead
Multiple Areas	*Use only where more than two apply and it is difficult to choose one*

Table 3. Case classification results.

Classification		Number of studies
Study Type	Protocol only	8
	Framework or theory	3
	Design / Build Only	10
	Design / Build with Evaluation	17
	Evaluation Only	23
Study Rigour	Poor – Small Sample/Self Report	16
	Average – Qualitative or No Control	12
	Good – Controlled	16
	N/A	17
Type of Personalisation	Adaptation	7
	Context Awareness	2
	Feedback	7
	Goal Setting	4
	Inter-Human Interaction	4
	Self-Learning	1
	User Targeting	29
	Multiple Areas	3
	N/A	4
Level of Personalisation	Weak	30
	Intermediate	15
	Strong	9
	N/A	7

> To achieve a greater level of personalization, these assessments can also be tailored to individuals by demographic (e.g. age, gender) or clinical information (e.g. endorsement of self-harm, depression score).
>
> (Iorfino et al. 2019, 3)

Personalisation in such cases was therefore not continuous throughout the use of the app and not individualised in a strict sense. These were therefore usually classified as weak personalisation, unless there were additional mechanisms used.

Stronger personalisation, which we have tended to link with adaptive or context-sensitive support, is something which was noted in user-centred design orientated studies as potentially desirable:

> participants often welcomed opportunities to give feedback, hoping to improve how well a tool would support them in the future.
>
> (Kornfield et al. 2022, 10)

Though even these noted an accompanying desire on the part of users to also be surprised, to be exposed to content not seen before.

Some studies with adaptive systems, such as Luan et al. medication management app for depression (2020), did indicate greater user satisfaction with time:

> the longer users use the system, the more satisfied they are with the recommended content. The main reason is that with the interaction with users, the system is more and more aware of users, and the recommended content is more and more accurate.
>
> (Luan et al. 2020, 8)

Of the eligible papers employing more adaptive machine learning as a basis of personalisation and/or recommendation, many were only reporting on their design stages and lacked a controlled evaluation with end users. Among those classified as having strong personalisation, five were in the framework/design stage, while four had design-build with evaluation or evaluation only. For intermediate personalisation, nine were in the framework/design stage, and seven had design-build with evaluation or evaluation only. Lewis et al. (2022) were among the four papers classified as having strong personalisation with design-build and evaluation. However, this study evaluated data collected from an app and simulated a personalisation technique, thereby lacking a controlled evaluation involving end users.

Examples of these more adaptive machine learning techniques included: Regression analysis to match student stress levels to categories of art for relaxation (Chandrasiri et al. 2021), a Naïve Bayes content recommender for therapeutic activities (Rohani et al. 2020), Model Predictive Control for PTSD intervention (Noble 2013), natural language processing of patient input to extract symptoms and match to treatment (Mukhiya et al. 2020b) and reinforcement learning (Q-Learning) to recommend interventions in an app for stress (Clarke, Jaimes and Labrador 2017).

Overall, there was conflicting evidence about whether personalisation was a viable route to improve engagement and increased the wellbeing of the users. As shown in Table 4, the Hedge's g statistics vary from paper to paper, due to the nature of the designs and target groups, showing either no effect or a strong one. Many articles did not supply the data to calculate these effects and so it is not possible to discern a strong trend. For example, Weber et al. (2019) suggest that their study demonstrates that personalisation of a mental health app to help manage and prevent work related stress was effective in comparison to the waitlist control group. The personalised approach in this study was simple, with personalised feedback on questionnaire scores and in-depth sleep data. There was a small effect size in two groups, indicating a small reduction in stress at the final timepoint.

Given the study designs, it was not possible to isolate the effect specifically due to personalisation features. But we can compare the effects to quite strong effect sizes of Hedge's $g = 0.6$ (0.45–0.75) for a review of digital interventions more generally (i.e. *non-personalised* apps) (Fu et al. 2020) and generally positive, though inconclusive engagement-to-outcome effects of $g = 0.40$, again across standardised apps (0.097–0.705) (Gan et al. 2021).

Table 4. Hedge's *g* for eligible articles.

Reference	Type of personalisation	Group	Hedge's *g* statistics	Comments
Weber, Lorenz and Hemmings (2019)	User Targeting	Intervention vs Waitlist Control – wellbeing scale	0.0136	Improved wellbeing – no effect
		Intervention vs Waitlist Control – general stress scale	0.1364	Decreased stress – very small effect
		Intervention vs Waitlist Control – cognitive stress scale	0.2065	Decreased stress – small effect
Lattie et al. (2020)	Inter-Human Interaction	Single study: baseline to 8 week – anxiety literacy questionnaire	0.2174	Anxiety improved – very small effect
Tsirmpas et al. (2022)	User Targeting	Single study: baseline to 8 weeks – PHQ-9	0.5814	Depression improved – medium effect
		Single study: baseline to 16 weeks – PHQ-9	1.2994	Depression improved – very large effect
		Single study: 8 week to 16 weeks – PHQ-9	0.7841	Depression improved – medium/ large effect
		Single study: baseline to 8 weeks – GAD-7	0.8261	Anxiety improved – large effect
		Single study: baseline to 16 weeks – GAD-7	1.0870	Anxiety improved – very strong effect
		Single study: 8 week to 16 weeks – GAD-7	0.4615	Anxiety improved – medium effect
Hwang et al. (2022)	Goal Setting	Intervention Vs Waitlist Control – PSS baseline	−0.3023	stress better (control) – small effect
		Intervention Vs Waitlist Control – PSS follow-up	0.9226	Stress decreased (intervention) – large effect

Studies looking at actual usage patterns were revealing. In terms of app usage and engagement, uptake of recommended self-help items may be as low as 4% of available activities and those recommendations selected may be shorter-term and easier to implement interventions (Rohani et al. 2020). Elsewhere, apps with recommendations, while increasing intention-to-use among participants, did not actually translate to more engagement (Currey and Torous 2023).

Qualitative evaluation work (prospective and retrospective) also revealed mixed attitudes toward personalisation and recommendations. For instance, in a design study just 42% of participants expressed a preference for personalised, intelligent wellbeing recommendations over more group-generalised recommendations (Sellak and Grobler 2020). Another study found participants would 'sometimes' welcome system recommendations (Kornfield et al. 2022).

4. Discussion

While caution about long-term or standalone effectiveness is warranted, mental health apps have the potential to comprise part of modern-day treatment (Ahmed et al. 2021). Apps can be used at different stages of treatment/prevention and can be used for different aftercare techniques and help with education, symptom assessment or even treatment adherence (Kerst, Zielasek, and Gaebel 2020). Due to the increase in demand for mental health services, mental health apps are likely to remain more available and accessible to the public then professional help.

Personalisation and recommendation approaches are becoming more and more familiar in everyday digital life, with many different retail and social media applications personalising the interface as they learn more about our preferences and behaviour. But personalisation can lead to harms as well as benefits (Stray et al. 2022) and there needs to be both oversight and transparency around how it operates.

This review identified forty-seven articles concerned with DMH support and personalisation. These papers provide a good sample of both the types and level of personalisation in proposed and current apps together with how likely app users are likely to find these useful.

Our analysis of the content and quality of these articles has the following implications for our research questions:

1. What different approaches to personalisation and recommendation have been used in digital MH research and research-based products?

Our classification saw the majority of articles rated as cases of weak personalisation, with coarse profiling or targeting being the most common method. This approach may commit participants to a common group at the outset, without the possibility of later refinement or movement between groups. This form of clustering may give functional simplicity for the provider (and improve usability for many users) but may miss longitudinally dynamic aspects of personal states and preferences.

Products described in many of the articles did not clearly state the underlying method of personalisation used. We can conclude that the term is sometimes used for emphasis or hyperbole only, when in fact the experience received by users is in fact relatively uniform. In other cases, authors may be reluctant, or not at liberty to describe technical intellectual property. This would be, in our view, a retrogressive tendency, and a failure to acknowledge that, in the area of mental health, algorithmic transparency should be essential and not merely a nice-to-have.

Of examples using an underlying machine learning model, these were largely found to be static (trained offline on a dataset before being deployed). Such models run the risk of validity decay, as predictions are most relevant to the data the model is trained on, potentially less so to new data provided by new users (Fröhlich et al. 2018).

2. Does personalisation and recommendation in MH apps lead to better user engagement and outcomes?

This question can be broken down into two parts: that personalisation can lead to better engagement, and that engagement with DMH products can be a determinant of better mental health outcomes. These are discussed below. Essentially, we find that the evidence for these premises is currently weak and conflicting, despite a strong discourse in the DMH community that personalisation is an unproblematic advantage.

For *engagement*, we find that personalisation may have little or no impact on indicators, with studies admitting, for instance, that:

> Deploying the app recommendation algorithm demonstrated feasibility but did not in itself change engagement. (Currey and Torous 2023, 6)

> it was found that these personalized behavioural goals did not have a significant impact on engagement levels compared to non-personalized behaviour goals. (James et al. 2022, 1574)

While participants often state preferences for personalised and recommended content, in practice they may behave differently. In their study of 218 app users assigned randomly to guided or autonomous conditions, Pieritz et al. (2021) found that although *all* participants declared they would prefer guided content, in practice those in the autonomous condition completed significantly more activities and also rated them higher. The authors recommended optimising for both efficacy and engagement by applying recommendations sparingly (Pieritz et al. 2021) Others have noted that a balance between choices and more explicit direction may be appropriate where people are less motivated or confident to choose themselves (Kornfield et al. 2022). Strongly recommendation-driven interfaces may negatively impact users' perception of agency and control. As one study participant has expressed it:

> I'd prefer to be in control, and not the computer … it's like the machine making choices … rather than you making your own decisions. (Blom 2002)

On the more positive side, Kornfield et al. (2022) stated in their study of potential users of a text messaging intervention that 'participants also suggested that they would sometimes welcome system recommendations as a way of maintaining momentum or moving out of their comfort zones'. Certainly, some trials based on digital personalised treatments have shown high completion rates (e.g. 85% for an 8 module web-based treatment for paranoia) though notably when the treatment was completed in the face to face setting of a clinical appointment (Ward et al. 2022).

As for demonstrating positive *impact* on mental health indicators, the majority of the evaluative studies that we reviewed used a personalisation technique and a waitlist control but did not include a group that were assigned a standardised form of the intervention. So, while studies did sometimes show strong effect sizes compared to controls (Table 4), these were not necessarily larger than those found in systematic studies of general (i.e. largely non-personalised) interventions. So there remains doubt as to whether personalised treatments are more impactful.

3. What gaps and opportunities exist for better understanding and promoting effective use of personalisation and recommendation?

From reviewing these papers we find that there has been no standard conception of what personalisation is in mental health apps, with many works defining personalisation as generic feedback, periodic questions or broad targeting, approaches that we have labelled as 'weak' according to Oinas-Kukkonen's conceptualisation (Oinas-Kukkonen 2018). Hornstein's recent work (2023) makes progress here by breaking down the concept into different dimensions and mechanisms which help in characterising and distinguishing different personalisation levels. This provides the opportunity for much finer grained and evidence-based feature comparisons within single app design and evaluations. To date there is little if any published research on this.

An additional gap lies in the need for a more thorough investigation into the actual interactions within the app, as opposed to solely relying on either the stated intentions of users or the system designers. Better HCI-based observational and data-driven work might expand on the intriguing work quoted above, which indicates that choice restriction – even if only perceived – might harm exploration and engagement.

A small number of projects have started to incorporate machine learning and statistical algorithms for user profiling and content recommendation. That said, in the literature we encountered there is a large skew towards proof-of-concept prototypes. Many of the studies we reviewed discussed design and build stages only, with the effect of the final intervention products on the intended user groups unknown. And operationalising machine learning is itself a research and development issue that can be better addressed in future work.

As noted above, machine learning approaches to personalisation that are available in our literature sample were often static and there are significant opportunities to trial online and iterated, offline learning that can be used to dynamically tune individual experiences. Coupled with some form of evaluation this provides a powerful tool to understand the effect of different forms of personalisation (or the varying of emphases on past as opposed to present context and preferences). At a more basic level, it is also worth comparing state-of-the-art adaptive personalisation to more social collaborative filtering (CF – where greater emphasis is placed on the overall preferences of the app user community) which is simpler to implement and may give comparative results. There is a little evidence from offline simulation studies such as Lewis et al (2022) that adaptive CF algorithms with contextual variables can provide more accurate predictions of recommendation suitability. Such *in silico* work provides a next best alternative to naturalistic and online learning evidence, which can be more expensive to collect.

Finally, the small sample sizes and lack of power of the evaluative studies reviewed here meant that it could be questionable whether the results shown were due to chance or a poor level of internal/external validity. Larger sample sizes should be included within future studies and better designs which evaluate different kinds of personalisation and compare these with more standardised approaches in multi-armed designs.

5. Conclusion

Personalisation of the DMH products in our study was often quite weak with people tending to be allocated to groups or types which are treated similarly in terms of interaction, content and messaging. This is some way from a user experience that really learns from and adapts to an individual's unique preferences.

In terms of impacts, despite a strong discourse as to the importance of personalisation to engagement and ensuing benefits in digital mental health support, the evidence for this remains to some extent contradictory. If anything, both behavioural studies and stated preferences suggest a blended approach, with either optional or limited personalisation of content which preserves user agency and control.

The engagement benefits of personalisation shown in face-to-face interventions, particularly to serve under-reached groups, still have the potential to be transferred to the digital space. Stronger and more responsive personalisation might be a way to achieve this, but more precise evaluation designs need to be employed to confirm if this is indeed the case.

DMH products are gaining popularity for providing immediate and anonymous support for individuals struggling with their mental health. Personalisation could be a valid route for creators of DMH to enable users to receive appropriate support based on their individual preferences. To help build the knowledge base, our review highlights the importance of researchers and private companies collaborating to openly and transparently share research on personalisation in DMH. This collaborative effort is crucial for evaluating the effectiveness of personalisation and understanding its potential limitations.

5.1. Success/strengths

This review followed the PRISMA guidelines for scoping reviews and is one of the first such reviews aimed at characterising personalisation in mental health apps, helping to demonstrate the gaps in research for future researchers.

We have taken strong yet diverse inclusion criteria to ensure that the right type of and range of articles were being selected.

5.2. Limitations

This review did exclude studies that did not include or state personalisation as a significant component, which could have limited the amount of evidence gathered. The eligibility criteria may also have eliminated papers relevant to this scoping review, with the article search being restricted to 2012–2023 and may have eliminated some interesting earlier articles.

In addition, as Hornstein et al. have pointed out (2023), published research tends to exclude many developments in the private /commercial sector, many of which are considerably more advanced in terms of implementation than in the academic domain. We therefore encourage more sharing of impact and engagement data from such platforms, perhaps through academic partnerships.

5.3. Future research

The purpose of this scoping review was to identify the different approaches used within DMH research and whether these approaches then lead to better user engagement and mental wellbeing.

As mental health and wellbeing apps are becoming more popular there are new apps being released regularly but without evidence being supplied on their effectiveness. There needs to be an ongoing effort to design evaluative studies for a more detailed assessment of the benefits of the different possible approaches to personalisation.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Ahmed, A., N. Ali, A. Giannicchi, A. A. Abd-alrazaq, M. A. S. Ahmed, S. Aziz, and M. Househ. 2021. “Mobile Applications for Mental Health Self-Care: A Scoping Review.” *Computer Methods and Programs in Biomedicine Update* 1: 100041. https://doi.org/10.1016/j.cmpbup.2021.100041.

Alqahtani, F., and R. Orji. 2020. “Insights from User Reviews to Improve Mental Health Apps.” *Health Informatics Journal* 26 (3): 2042–2066. https://doi.org/10.1177/1460458219896492.

Antezana, G., N. Bidargaddi, V. Blake, G. Schrader, B. Kaambwa, S. Quinn, S. Orlowski, M. Winsall, and M. Battersby. 2015. “Development of an Online Well-being Intervention for Young People: An Evaluation Protocol.” *JMIR Research Protocols* 4: 2. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84971015837&doi=10.2196%2fresprot.4098&partnerID=40&md5=3e056ab5f7632280bf5469d7946db3b3.

Athanas, A., J. McCorrison, J. Campistron, N. Bender, J. Price, S. Smalley, and N. J. Schork. 2021. “Characterizing Emotional State Transitions During Prolonged Use of a Mindfulness and Meditation App: Observational Study.” *JMIR Mental Health* 8: 3. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85103224704&doi=10.2196%2f19832&partnerID=40&md5=0f1a43a49b793742288505782bbb8d27.

Ben-Yehuda, O., E. Dreazen, D. Koren, and M. Peleg. 2022. “Participatory Design of a Mobile App to Safeguard Mental Resilience in the Context of Drug Use in Young Adults: Multi-Method Study.” *JMIR Formative Research* 6: 2. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85126129268&doi=10.2196%2f34477&partnerID=40&md5=895a87105bd9e6e330291524a99c3d90.

Bertacco, M., O. Soyeux, R. Durand, P. Boudrias, L. Wiseman, P. Rompré, P. Rainville, E. Emami, and N. Gosselin. 2022. “Effect of Personalized Musical Intervention on Burden of Care in Dental Implant Surgery: A Pilot Randomized Controlled Trial.” *Journal of Dentistry* 120. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85127033605&doi=10.1016%2fj.jdent.2022.104091&partnerID=40&md5=5552b43318950c46b72eeae3fa8dddfa.

Bhattacharya, A., C. Liang, E. Y. Zeng, K. Shukla, M. E. R. Wong, S. A. Munson, and J. A. Kientz. 2019. “Engaging Teenagers in Asynchronous Online Groups to Design for Stress Management.” In *Proceedings of the 18th ACM International Conference on Interaction Design and ChildrenIDC ‘19*, 26–37. New York, NY: Association for Computing Machinery. Available from: https://doi.org/10.1145/3311927.3323140 [Accessed 7 August 2023].

Bidargaddi, N., P. Musiat, M. Winsall, G. Vogl, V. Blake, S. Quinn, S. Orlowski, G. Antezana, and G. Schrader. 2017. “Efficacy of a Web-Based Guided Recommendation Service for a Curated List of Readily Available Mental Health and Well-Being Mobile Apps for Young People: Randomized Controlled Trial.” *Journal of Medical Internet Research* 19: 5. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85019937415&doi=10.2196%2fjmir.6775&partnerID=40&md5=844baaa7ef144670976f950f453856c9.

Blom, J. 2002. “A theory of personalized recommendations.” In *CHI ‘02 Extended Abstracts on Human Factors in Computing SystemsCHI EA ‘02*, 540–541. New York, NY: Association for Computing Machinery. . Available from: http://doi.org/10.1145/506443.506471 [Accessed 23 July 2022].

Bonn, S. E., M. Löf, C.-G. Östenson, and Y. Trolle Lagerros. 2019. “App-technology to Improve Lifestyle Behaviors Among Working Adults – The Health Integrator Study, a Randomized Controlled Trial.” *BMC Public Health* 19 (1): 1. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85062618856&doi=10.1186%2fs12889-019-6595-6&partnerID=40&md5=ab42e45b0d01af228f3c478cdb9ce162.

Borghouts, J., et al. 2021. “Barriers and Facilitators to User Engagement with Digital Mental Health Interventions: A Systematic Review (Preprint).” *Journal of Medical Internet Research.* 23: 3.

Burley, C., et al. 2020. “Increasing Engagement in eHealth Interventions Using Personalization and Implementation Intentions.” In *2020 Systems and Information Engineering Design Symposium (SIEDS)*, 1–5.

Cabrita, M., H. op den Akker, M. Tabak, H. J. Hermens, and M. M. R. Vollenbroek-Hutten. 2018. “Persuasive Technology to Support Active and Healthy Ageing: An Exploration of Past, Present, and Future.” *Journal of Biomedical Informatics [online* 84: 17–30.

Chandrasiri, K. G. P. R., A. A. Chandrasena, L. H. C. R. De Silva, H. W. V. O. Jayasinghe, G. T. Dassanayake, and O. Seneweera. 2021. “Mellow: Stress Management System For University Students In Sri Lanka.” In *2021 6th International Conference on Information Technology Research (ICITR)*, 1–6.

Chaturvedi, A., et al. 2023. “Content Recommendation Systems in Web-Based Mental Health Care: Real-world Application and Formative Evaluation.” *JMIR Formative Research* 7. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85149102069&doi=10.2196%2f38831&partnerID=40&md5=a28be1ba60a485526f5ce97aa93776fa.

Clarke, S., L. G. Jaimes, and M. A. Labrador. 2017. “mStress: A Mobile Recommender System for Just-in-time

Interventions for Stress." In *2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, 1–5.

Currey, D., and J. Torous. 2022. "Digital Phenotyping Data to Predict Symptom Improvement and App Personalization: Protocol for a Prospective Study." *JMIR Research Protocols* 11 (11): e37954. https://doi.org/10.2196/37954.

Currey, D., and J. Torous. 2023. "Digital Phenotyping Data to Predict Symptom Improvement and Mental Health App Personalization in College Students: Prospective Validation of a Predictive Model." *Journal of Medical Internet Research* 25 (1): e39258. https://doi.org/10.2196/39258.

Ferré-Bergadà, M., A. Valls, L. Raigal-Aran, J. Lorca-Cabrera, N. Albacar-Riobóo, T. Lluch-Canut, and C. Ferré-Grau. 2021. "A Method to Determine a Personalized Set of Online Exercises for Improving the Positive Mental Health of a Caregiver of a Chronically Ill Patient." *BMC Medical Informatics and Decision Making* 21 (1): 1. https://doi.org/10.1186/s12911-021-01445-6.

Figueroa, C. A., N. Deliu, B. Chakraborty, A. Modiri, J. Xu, J. Aggarwal, J. Jay Williams, C. Lyles, and A. Aguilera. 2022. "Daily Motivational Text Messages to Promote Physical Activity in University Students: Results from a Microrandomized Trial." *Annals of Behavioral Medicine* 56 (2): 212–218. https://doi.org/10.1093/abm/kaab028.

Fröhlich, H., et al. 2018. "From Hype to Reality: Data Science Enabling Personalized Medicine." *BMC Medicine* 16 (1): 150. https://doi.org/10.1186/s12916-018-1122-7.

Fu, Z., H. Burger, R. Arjadi, and C. L. H. Bockting. 2020. "Effectiveness of Digital Psychological Interventions for Mental Health Problems in Low-income and Middle-income Countries: a Systematic Review and Meta-analysis." *The Lancet Psychiatry* 7 (10): 851–864. https://doi.org/10.1016/S2215-0366(20)30256-X.

Gan, D. Z. Q., L. McGillivray, J. Han, H. Christensen, and M. Torok. 2021. "Effect of Engagement with Digital Interventions on Mental Health Outcomes: A Systematic Review and Meta-Analysis." *Frontiers in Digital Health* 3. . Available from: https://www.frontiersin.org/articles/10.3389fdgth.2021.764079 [Accessed 10 March 2023].

Germain, A., et al. 2021. "Enhancing Behavioral Sleep Care with Digital Technology: Study Protocol for a Hybrid Type 3 Implementation-Effectiveness Randomized Trial." *Trials* 22 (1): 1. https://doi.org/10.1186/s13063-020-04974-z.

Ghandeharioun, A., D. McDuff, M. Czerwinski, and K. Rowan. 2019. "EMMA: An Emotion-Aware Wellbeing Chatbot." In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, 1–7.

Ghaznavi, I., D. Gillies, D. Nicholls, and A. Edalat. 2020. "Photorealistic Avatars to Enhance the Efficacy of Selfattachment psychotherapy." In *2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, 60–67.

Gire, N., et al. 2021. "Care Co-ordinator in My Pocket': a Feasibility Study of Mobile Assessment and Therapy for Psychosis (TechCare)." *BMJ Open* 11 (11): 11. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85120032764&doi=10.1136%2fbmjopen-2020-046755&partnerID=40&md5=e4068ece6af687e08a31a8efca2d5816.

Görtz-Dorten, A., M. Frank, A. Fessel, L. Hofmann, and M. Döpfner. 2022. "Effects of a Smartphone App-Augmented Treatment for Children with Oppositional Defiant Disorder/Conduct Disorder and Peer-Related Aggressive Behavior – A Pilot Study." *Trials* 23 (1): 1. https://doi.org/10.1186/s13063-022-06325-6.

Heffernan, M., et al. 2019. "Maintain Your Brain: Protocol of a 3-Year Randomized Controlled Trial of a Personalized Multi-Modal Digital Health Intervention to Prevent Cognitive Decline among Community Dwelling 55 to 77 Year Olds." *Journal of Alzheimer's Disease* 70 (s1): S221–S237. https://doi.org/10.3233/JAD-180572.

Hornstein, S., K. Zantvoort, U. Lueken, B. Funk, and K. Hilbert. 2023. "Personalization Strategies in Digital Mental Health Interventions: A Systematic Review and Conceptual Framework for Depressive Symptoms." *Frontiers in Digital Health* 5. Available from: https://www.frontiersin.org/articles/10.3389fdgth.2023.1170002 [Accessed 18 August 2023].

Hung, G. C.-L., P.-C. Yang, C.-Y. Wang, and J.-H. Chiang. 2015. "A Smartphone-Based Personalized Activity Recommender System for Patients with Depression." In *MOBIHEALTH 2015 – 5th EAI International Conference on Wireless Mobile Communication and Healthcare - Transforming Healthcare through Innovations in Mobile and Wireless Technologies*. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85034964323&doi=10.4108%2feai.14-10-2015.2261655&partnerID=40&md5=fb44798ac169a7d6c5aae535025c3deb.

Hwang, H., S. M. Kim, B. Netterstrøm, and D. H. Han. 2022. "The Efficacy of a Smartphone-Based App on Stress Reduction: Randomized Controlled Trial." *Journal of Medical Internet Research* 24: 2. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85124609806&doi=10.2196%2f28703&partnerID=40&md5=fac5e08b8304b1d36135cc5f0af5f36f.

Insel, T. R. 2017. "Digital Phenotyping: Technology for a New Science of Behavior." *JAMA* 318 (13): 1215–1216. https://doi.org/10.1001/jama.2017.11295.

Iorfino, F., S. P. Cross, T. Davenport, J. S. Carpenter, E. Scott, S. Shiran, and I. B. Hickie. 2019. "A Digital Platform Designed for Youth Mental Health Services to Deliver Personalized and Measurement-Based Care." *Frontiers in Psychiatry* 10), Available from: https://www.frontiersin.org/articles/10.3389fpsyt.2019.00595 [Accessed 1 November 2023].

Jakob, R., S. Harperlink, A. Rudolf, E. Fleisch, S. Haug, J. Mair, A. Salamanca-Sanabria, and T. Kowatsch. 2022. "Factors Influencing Adherence to mHealth Apps for Prevention or Management of Noncommunicable Diseases: Systematic Review." *Journal of Medical Internet Research [online* 24 (5): e35371. https://doi.org/10.2196/35371.

James, L. J., J. M. A. van Heugten, P. M. E. Van Gorp, R. C. Y. Nuijten, B. Montagne, M. A. Hagenaars, and L. E. Frank. 2022. "Evaluation of Personalized Treatment Goals on Engagement of SMI Patients with an mHealth app." In *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 1568–1573.

Jarrahi, M. H., N. Gafinowitz, and G. Shin. 2018. "Activity Trackers, Prior Motivation, and Perceived Informational and Motivational Affordances." *Personal and Ubiquitous Computing* 22 (2), https://doi.org/10.1007/s00779-017-1099-9.

Kennard, B. D., et al. 2018. "As Safe As Possible (ASAP): A Brief App-Supported Inpatient Intervention to Prevent Postdischarge Suicidal Behavior in Hospitalized, Suicidal Adolescents." *American Journal of Psychiatry* 175 (9): 864–872. https://doi.org/10.1176/appi.ajp.2018.17101151.

Kerst, A., J. Zielasek, and W. Gaebel. 2020. "Smartphone Applications for Depression: a Systematic Literature Review and a Survey of Health care Professionals' Attitudes Towards Their Use in Clinical Practice." *European Archives of Psychiatry and Clinical Neuroscience* 270 (2): 139–152. https://doi.org/10.1007/s00406-018-0974-3.

Khwaja, M., M. Ferrer, J. O. Iglesias, A. Aldo Faisal, and A. Matic. 2019. "Aligning Daily Activities with Personality: Towards a Recommender System for Improving Wellbeing." In *RecSys 2019 - 13th ACM Conference on Recommender Systems*, 368–372. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85073332019&doi=10.1145%2f3298689.3347020&partnerID=40&md5=2d5fc38f06f2b8d032e243013c374cd3.

Khwaja, M., S. Pieritz, A. A. Faisal, and A. Matic. 2021. "Personality and Engagement with Digital Mental Health Interventions." In *UMAP '21: 29th ACM Conference on User Modeling, Adaptation and Personalization*, 235–239. Utrecht: ACM. . Available from: https://dl.acm.org/doi/10.11453450613.3456823 [Accessed 2 August 2022].

Kop, R., M. Hoogendoorn, and M. C. A. Klein. 2014. "A Personalized Support Agent for Depressed Patients: Forecasting Patient Behavior Using a Mood and Coping Model." In *2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 3*, 302–309.

Kornfield, R., J. Meyerhoff, H. Studd, A. Bhattacharjee, J. J. Williams, M. Reddy, and D. C. Mohr. 2022. "Meeting Users Where They Are: User-centered Design of an Automated Text Messaging Tool to Support the Mental Health of Young Adults." In *Proceedings of the 2022 CHI Conference on Human Factors in Computing SystemsCHI '22*, 1–16. New York, NY: Association for Computing Machinery. Available from: https://doi.org/10.1145/3491102.3502046 [Accessed 31 August 2022].

Lanata, A., G. Valenza, M. Nardelli, C. Gentili, and E. P. Scilingo. 2015. "Complexity Index from a Personalized Wearable Monitoring System for Assessing Remission in Mental Health." *IEEE Journal of Biomedical and Health Informatics* 19 (1): 132–139. https://doi.org/10.1109/JBHI.2014.2360711.

Lattie, E., K. A. Cohen, N. Winquist, and D. C. Mohr. 2020. "Examining an App-based Mental Health Self-care Program, Intellicare for College Students: Single-arm Pilot Study." *JMIR Mental Health* 7: 10. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85097853366&doi=10.2196%2f21075&partnerID=40&md5=534aa31e94c9ea0680027072618e94df.

Leary, M., K. Pursey, A. Verdejo-García, J. Skinner, M. C. Whatnall, P. Hay, C. Collins, A. L. Baker, and T. Burrows. 2022. "Designing an Online Intervention for Adults with Addictive Eating: a Qualitative Integrated Knowledge Translation Approach." *BMJ Open* 12 (6): 6. https://doi.org/10.1136/bmjopen-2021-060196.

Leightley, D., R. J. Rona, J. Shearer, C. Williamson, C. Gunasinghe, A. Simms, N. T. Fear, L. Goodwin, and D. Murphy. 2020. "Evaluating the Efficacy of a Mobile App (Drinks:Ration) and Personalized Text and Push Messaging to Reduce Alcohol Consumption in a Veteran Population: Protocol for a Randomized Controlled Trial." *JMIR Research Protocols* 9: 10. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85095776628&doi=10.2196%2f19720&partnerID=40&md5=2fd4680961b34fbe75796c3ab01dbd3a.

Lewis, R., C. Ferguson, C. Wilks, N. Jones, and R. W. Picard. 2022. "Can a Recommender System Support Treatment Personalisation in Digital Mental Health Therapy? A Quantitative Feasibility Assessment Using Data from a Behavioural Activation Therapy App." In *Conference on Human Factors in Computing Systems – Proceedings*. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85129738507&doi=10.1145%2f3491101.3519840&partnerID=40&md5=af9df38944c20c4e84d2a7fc35a0d79a.

Lindow, J. C., J. L. Hughes, C. South, A. Minhajuddin, L. Gutierrez, E. Bannister, M. H. Trivedi, and M. J. Byerly. 2020. "The Youth Aware of Mental Health Intervention: Impact on Help Seeking, Mental Health Knowledge, and Stigma in U.S. Adolescents." *Journal of Adolescent Health* 67 (1): 101–107. https://doi.org/10.1016/j.jadohealth.2020.01.006.

Luan, L., W. Xiao, K. Hwang, M. S. Hossain, G. Muhammad, and A. Ghoneim. 2020. "MEMO Box: Health Assistant for Depression with Medicine Carrier and Exercise Adjustment Driven by Edge Computing." *IEEE Access* 8: 195568–195577. https://doi.org/10.1109/ACCESS.2020.3031725.

Lungu, A., R. E. Wickham, S.-Y. Chen, J. J. Jun, Y. Leykin, and C. E.-J. Chen. 2022. "Component Analysis of a Synchronous and Asynchronous Blended Care CBT Intervention for Symptoms of Depression and Anxiety: Pragmatic Retrospective Study." *Internet Interventions* 28), Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85127874836&doi=10.1016%2fj.invent.2022.100536&partnerID=40&md5=58f3e53430e61735e5816aca4203611f.

Matthews, P., P. Topham, and P. Caleb-Solly. 2018. "Interaction and Engagement with an Anxiety Management App: Analysis Using Large-Scale Behavioral Data." *JMIR Mental Health* 5 (4): e9235. https://doi.org/10.2196/mental.9235.

Mayer, G., et al. 2022. "User Experience and Acceptance of Patients and Healthy Adults Testing a Personalized Self-management App for Depression: A Non-randomized Mixed-methods Feasibility Study." *Digital Health* 8. https://doi.org/10.1177/20552076221091353.

Miura, C., S. Chen, S. Saiki, M. Nakamura, and K. Yasuda. 2022. "Assisting Personalized Healthcare of Elderly People: Developing a Rule-Based Virtual Caregiver System Using Mobile Chatbot." *Sensors* 22: 10. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85130203562&doi=10.3390%2fs22103829&partnerID=40&md5=779e6dd32cb24b3739549baded9fa3f5.

Moe-Byrne, T., J. Shepherd, D. Merecz-Kot, M. Sinokki, P. Naumanen, L. H. Roijen, and C. V. D. Feltz-Cornelis. 2022. "Effectiveness of Tailored Digital Health

Interventions for Mental Health at the Workplace: A Systematic Review of Randomised Controlled Trials." *PLOS Digital Health [online]* 1 (10): e0000123.

Monteiro-Guerra, F., O. Rivera-Romero, L. Fernandez-Luque, and B. Caulfield. 2020. "Personalization in Real-Time Physical Activity Coaching Using Mobile Applications: A Scoping Review." *IEEE Journal of Biomedical and Health Informatics* 24 (6): 1738–1751. https://doi.org/10.1109/JBHI.2019.2947243.

Moral-Munoz, J. A., A. Salazar, M. Dueñas, H. De Sola, and I. Failde. 2022. "Smartphone-based Exercise Intervention for Chronic Pain: PainReApp Randomized Clinical Trial Protocol." *Journal of Advanced Nursing* 78 (2): 569–576. https://doi.org/10.1111/jan.15095.

Mukhiya, S. K., U. Ahmed, F. Rabbi, K. I. Pun, and Y. Lamo. 2020a. "Adaptation of IDPT System Based on Patient-Authored Text Data using NLP." In *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, 226–232.

Mukhiya, S. K., J. D. Wake, Y. Inal, and Y. Lamo. 2020b. "Adaptive Systems for Internet-Delivered Psychological Treatments." *IEEE Access* 8: 112220–112236. https://doi.org/10.1109/ACCESS.2020.3002793.

Nebot, À, S. Domènech, N. Albino-Pires, F. Mugica, A. Benali, X. Porta, O. Nebot, and P. M. Santos. 2022. "LONG-REMI: An AI-Based Technological Application to Promote Healthy Mental Longevity Grounded in Reminiscence Therapy." *International Journal of Environmental Research and Public Health* 19: 10. https://doi.org/10.3390/ijerph19105997.

Neumeier, W. H., N. Guerra, M. Thirumalai, B. Geer, D. Ervin, and J. H. Rimmer. 2017. "POWERS forID: Personalized Online Weight and Exercise Response System for Individuals with Intellectual Disability: Study protocol for a randomized controlled trial." *Trials* 18 (1): 1. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85032866742&doi=10.1186%2fs13063-017-2239-2&partnerID=40&md5=c9ea1a372f806173e3f6356713dc26a0.

Newbold, A., et al. 2020. "Promotion of Mental Health in Young Adults Via Mobile Phone App: Study Protocol of the ECoWeB (Emotional Competence for Well-being in Young Adults) Cohort Multiple Randomised Trials." *BMC Psychiatry* 20 (1): 1. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85091571230&doi=10.1186%2fs12888-020-02857-w&partnerID=40&md5=9160ab1f2203498c016d92ed0ea3f236.

Noble, S. L. 2013. "Control-Theoretic Treatment Scheduling for Posttraumatic Stress Disorder." In *2013 American Control Conference*, 722–727.

O'Cillin, R. B. 2022. "The Engagement and Treatment Effectiveness of Personalised Interventions in Youth Mental Health: A Mixed Method Systematic Review in a European Context." *Journal of European Psychology Students* 13 (1): 18–77. https://doi.org/10.5334/jeps.518.

Oinas-Kukkonen, H. 2018. "Personalization Myopia: A Viewpoint to True Personalization of Information Systems." In *Proceedings of the 22nd International Academic Mindtrek ConferenceMindtrek '18*, 88–91. New York, NY: Association for Computing Machinery. Available from: http://doi.org/10.1145/3275116.3275121 [Accessed 23 July 2022].

op den Akker, H., V. M. Jones, and H. J. Hermens. 2014. "Tailoring Real-Time Physical Activity Coaching Systems: A Literature Survey and Model." *User Modeling and User-Adapted Interaction* 24 (5): 351–392. https://doi.org/10.1007/s11257-014-9146-y.

Pieritz, S., M. Khwaja, A. A. Faisal, and A. Matic. 2021. "Personalised Recommendations in Mental Health Apps: The Impact of Autonomy and Data Sharing." In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, CHI '21*, 1–12. Yokohama: ACM. Available from: https://dl.acm.org/doi/10.11453411764.3445523 [Accessed 8 April 2022].

Rabbi, M., M. S. H. Aung, G. Gay, M. C. Reid, and T. Choudhury. 2018. "Feasibility and Acceptability of Mobile Phone-Based Auto-Personalized Physical Activity Recommendations for Chronic Pain Self-Management: Pilot Study on Adults." *Journal of Medical Internet Research* 20: 10. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85055631376&doi=10.2196%2f10147&partnerID=40&md5=e73e3dbbc7f77c66f3a29284e828958f.

Riese, H., L. von Klipstein, R. A. Schoevers, D. C. van der Veen, and M. N. Servaas. 2021. "Personalized ESM Monitoring and Feedback to Support Psychological Treatment for Depression: a Pragmatic Randomized Controlled Trial (Therap-i)." *BMC Psychiatry* 21: 1. https://doi.org/10.1186/s12888-021-03123-3.

Rohani, D. A., A. Quemada Lopategui, N. Tuxen, M. Faurholt-Jepsen, L. V. Kessing, and J. E. Bardram. 2020. "MUBS: A Personalized Recommender System for Behavioral Activation in Mental Health." In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI '20*, 1–13. Honolulu, HI: ACM. . Available from: https://dl.acm.org/doi/10.11453313831.3376879 [Accessed 2 August 2022].

Rohani, D. A., A. Springer, V. Hollis, J. E. Bardram, and S. Whittaker. 2021. "Recommending Activities for Mental Health and Well-Being: Insights from Two User Studies." *IEEE Transactions on Emerging Topics in Computing [online]IEEE Transactions on Emerging Topics in Computing* 9 (3): 1183–1193. https://doi.org/10.1109/TETC.2020.2972007.

Rosario, D. F. T. D., A. E. D. Mariano, and M. J. C. Samonte. 2019. "SmileTeq: An Assistive and Recommendation Mobile Application for People with Anxiety, Depression or Stress." In *2019 International Conference on Information and Communication Technology Convergence (ICTC)*, 1304–1309.

Schlosser, D. A., T. R. Campellone, B. Truong, K. Etter, S. Vergani, K. Komaiko, and S. Vinogradov. 2018. "Efficacy of PRIME, a Mobile App Intervention Designed to Improve Motivation in Young People with Schizophrenia." *Schizophrenia Bulletin* 44 (5): 1010–1020. https://doi.org/10.1093/schbul/sby078.

Sellak, H., and M. Grobler. 2020. "mHealth4U: Designing for Health and Wellbeing Self-Management." In *2020 35th IEEE/ACM International Conference on Automated Software Engineering Workshops (ASEW)*, 41–46.

Shalaby, R., W. Vuong, E. Eboreime, S. Surood, A. J. Greenshaw, and V. I. O. Agyapong. 2022. "Patients' Expectations and Experiences With a Mental Health–Focused Supportive Text Messaging Program: Mixed

Methods Evaluation." *JMIR Formative Research* 6 (1): e33438. https://doi.org/10.2196/33438.

Stray, J., et al. 2022. Building Human Values into Recommender Systems: An Interdisciplinary Synthesis [online]. Available from: http://arxiv.org/abs/2207.10192 [Accessed 30 July 2022].

Tang, D., R. Macniven, N. Bender, C. Jones, and B. Gopinath. 2022. "Development, Implementation and Evaluation of the Online Movement, Interaction and Nutrition for Greater Lifestyles in the Elderly (MINGLE) program: The Protocol for a Pilot Trial." *PLoS One* 17 (5). Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85129985998&doi=10.1371%2fjournal.pone.0267581&partnerID=40&md5=19ffac0894e91aeaf6fd83dab6dc4150.

Tricco, A., et al. 2018. "PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation." *Annals of Internal Medicine*. https://www.acpjournals.org/doi/10.7326M18-0850 [Accessed 17 May 2023]

Tsirmpas, C., D. Andrikopoulos, P. Fatouros, G. Eleftheriou, J. A. Anguera, K. Kontoangelos, and C. Papageorgiou. 2022. "Feasibility, Engagement, and Preliminary Clinical Outcomes of a Digital Biodata-Driven Intervention for Anxiety and Depression." *Frontiers in Digital Health* 4. https://doi.org/10.3389/fdgth.2022.868970.

Vairavasundaram, S., V. Varadarajan, D. Srinivasan, V. Balaganesh, S. B. Damerla, B. Swaminathan, and L. Ravi. 2022. "Dynamic Physical Activity Recommendation Delivered through a Mobile Fitness App: A Deep Learning Approach." *Axioms* 11: 7. https://doi.org/10.3390/axioms11070346.

Valentine, L., S. D'Alfonso, and R. Lederman. 2023. "Recommender Systems for Mental Health Apps: Advantages and Ethical Challenges." *AI & SOCIETY* 38 (4): 1627–1638. http://doi.org/10.1007/s00146-021-01322-w.

Varnfield, M., K. Rajesh, C. Redd, S. Gibson, L. Gwillim, and S. Polkinghorne. 2019. "Health-e Minds: a Participatory Personalised and Gamified mHealth Platform to Support Healthy Living Behaviours for People with Mental Illness." In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 6943–6947. https://doi.org/10.1109/EMBC.2019.8857286.

Vonk, T., E. A. Bakker, E. S. Zegers, M. T. E. Hopman, and T. M. H. Eijsvogels. 2021. "Effect of a Personalised mHealth Home-based Training Application on Physical Activity Levels During and After cEntre-based Cardiac Rehabilitation: Rationale and Design of the Cardiac RehApp Randomised Control Trial." *BMJ Open Sport and Exercise Medicine* 7 (3): 3. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85115034152&doi=10.1136%2fbmjsem-2021-001159&partnerID=40&md5=dda1cd87c3e34ad0108901f356221889.

Wang, R., et al. 2016. "CrossCheck: Toward Passive Sensing and Detection of Mental Health Changes in People with Schizophrenia." In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*, 886–897. New York, NY: Association for Computing Machinery. Available from: https://doi.org/10.1145/2971648.2971740 [Accessed 31 August 2022].

Ward, T., et al. 2022. "SlowMo Therapy, a New Digital Blended Therapy for Fear of Harm from Others: An Account of Therapy Personalisation within a Targeted Intervention." *Psychology and Psychotherapy: Theory, Research and Practice* 95 (2): 423–446. https://doi.org/10.1111/papt.12377.

Weber, S., C. Lorenz, and N. Hemmings. 2019. "Improving Stress and Positive Mental Health at Work via an App-Based Intervention: A Large-Scale Multi-Center Randomized Control Trial." *Frontiers in Psychology* 10. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85077277195&doi=10.3389%2ffpsyg.2019.02745&partnerID=40&md5=e8a4e14f87865cb294bc7658856eac04.

Weisel, K. K. 2019. "Link to External Site, This Link Will Open in a New Window, Fuhrmann, L.M., Matthias, B., Harald, B., Pim, C. and Ebert, D.D. (2019) Standalone Smartphone Apps for Mental Health—a Systematic Review and Meta-analysis." *NPJ Digital Medicine* 2 (1), Available from: https://www.proquest.com/docview/2525627502/abstract/69DCD7B44FB44DBAPQ/1 [Accessed 10 March 2023].

Xin, Y., Y. Chen, L. Jin, Y. Cai, and L. Feng. 2017. "TeenRead: An Adolescents Reading Recommendation System Towards Online Bibliotherapy." In *2017 IEEE International Congress on Big Data (BigData Congress)*, , Honolulu, HI, USA, 431–434. https://doi.org/10.1109/BigDataCongress.2017.63.

Appendix. Table showing the list of coded eligible articles

Classification		Studies
Study Type	Protocol only	(Currey and Torous 2022; Moral-Munoz et al. 2022; Tang et al. 2022; Riese et al. 2021; Leightley et al. 2020; Newbold et al. 2020; Neumeier et al. 2017; Antezana et al. 2015)
	Framework or theory	(Germain et al. 2021; Vonk et al. 2021; Khwaja et al. 2019)
	Design/Build Only	(Leary et al. 2022; Vairavasundaram et al. 2022; Chandrasiri et al. 2021; Luan et al. 2020; Mukhiya et al. 2020a; Varnfield et al. 2019; Clarke, Jaimes and Labrador 2017; Hung et al. 2015; Lanata et al. 2015; Noble 2013)
	Design/Build with Evaluation	(Chaturvedi et al. 2023; Ben-Yehuda et al. 2022; Görtz-Dorten et al. 2022; Kornfield et al. 2022; Lewis et al. 2022; Miura et al. 2022; Nebot et al. 2022; Athanas et al. 2021; Burley et al. 2020; Ghaznavi et al. 2020; Rohani et al. 2020; Sellak and Grobler 2020; Rosario, Mariano and Samonte 2019; Rabbi et al. 2018; Xin et al. 2017; Wang et al. 2016; Kop, Hoogendoorn and Klein 2014)
	Evaluation Only	(Currey and Torous 2023; Bertacco et al. 2022; Figueroa et al. 2022; Hwang et al. 2022; James et al. 2022; Lungu et al. 2022; Mayer et al. 2022; Shalaby et al. 2022; Tsirmpas et al. 2022; Ward et al. 2022; Gire et al. 2021; Khwaja et al. 2021; Rohani et al. 2021; Lattie et al. 2020; Bhattacharya et al. 2019; Bonn et al. 2019; Ghandeharioun et al. 2019; Heffernan et al. 2019; Weber, Lorenz, and Hemmings 2019; Kennard et al. 2018; Rabbi et al. 2018; Schlosser et al. 2018)
Study Rigor	Poor – Small Sample/Self Report	(Bertacco et al. 2022; Görtz-Dorten et al. 2022; James et al. 2022; Mayer et al. 2022; Tsirmpas et al. 2022; Gire et al. 2021; Ghaznavi et al. 2020; Lattie et al. 2020; Bhattacharya et al. 2019; Rosario, Mariano and Samonte 2019; Rabbi et al. 2018, 2018; Schlosser et al. 2018; Xin et al. 2017)
	Average – Qualitative or No Control	(Currey and Torous 2023; Ben-Yehuda et al. 2022; Figueroa et al. 2022; Hwang et al. 2022; Miura et al. 2022; Nebot et al. 2022; Ward et al. 2022; Khwaja et al. 2021; Burley et al. 2020; Rohani et al. 2020; Sellak and Grobler 2020; Ghandeharioun et al. 2019)
	Good – Controlled	(Chaturvedi et al. 2023; Currey and Torous 2022; Lewis et al. 2022; Lungu et al. 2022; Moral-Munoz et al. 2022; Shalaby et al. 2022; Athanas et al. 2021; Riese et al. 2021; Rohani et al. 2021; Newbold et al. 2020; Bonn et al. 2019; Heffernan et al. 2019; Khwaja et al. 2019; Weber, Lorenz and Hemmings 2019; Kennard et al. 2018; Bidargaddi et al. 2017; Neumeier et al. 2017; Antezana et al. 2015)
	N/A	(Kornfield et al. 2022; Leary et al. 2022; Tang et al. 2022; Vairavasundaram et al. 2022; Chandrasiri et al. 2021; Germain et al. 2021; Vonk et al. 2021; Leightley et al. 2020; Luan et al. 2020; Mukhiya et al. 2020a; Varnfield et al. 2019; Clarke, Jaimes and Labrador 2017; Hung et al. 2015; Lanata et al. 2015; Noble 2013)
Type of Personalisation	Adaptation	(Kornfield et al. 2022; Rohani et al. 2021; Luan et al. 2020; Ghandeharioun et al. 2019; Clarke, Jaimes and Labrador 2017; Xin et al. 2017; Noble 2013)
	Context Awareness	(Görtz-Dorten et al. 2022; Wang et al. 2016)
	Feedback	(Figueroa et al. 2022; Leary et al. 2022; Mayer et al. 2022; Ward et al. 2022; Riese et al. 2021; Ghaznavi et al. 2020; Heffernan et al. 2019)
	Goal Setting	(Hwang et al. 2022; Tang et al. 2022; Vonk et al. 2021; Varnfield et al. 2019)
	Inter-Human Interaction	(Lungu et al. 2022; Miura et al. 2022; Lattie et al. 2020; Bonn et al. 2019)
	Self-Learning	(Rohani et al. 2020)
	User Targeting	(Chaturvedi et al. 2023; Currey and Torous, 2022, 2023; Bertacco et al. 2022; James et al. 2022; Lewis et al. 2022; Moral-Munoz et al. 2022; Nebot et al. 2022; Tsirmpas et al. 2022; Vairavasundaram et al. 2022; Athanas et al. 2021; Chandrasiri et al. 2021; Germain et al. 2021; Gire et al. 2021; Burley et al. 2020; Mukhiya et al. 2020a; Newbold et al. 2020; Bhattacharya et al. 2019; Rosario, Mariano and Samonte 2019; Weber, Lorenz and Hemmings 2019; Kennard et al. 2018; Rabbi et al. 2018; Schlosser et al. 2018; Bidargaddi et al. 2017; Neumeier et al. 2017; Hung et al. 2015; Lanata et al. 2015; Kop, Hoogendoorn and Klein 2014)
	Multiple Areas	(Leightley et al. 2020; Khwaja et al. 2019; Antezana et al. 2015)
	N/A	(Ben-Yehuda et al. 2022; Shalaby et al. 2022; Khwaja et al. 2021; Sellak and Grobler 2020)
Level of Personalisation	Weak	(Chaturvedi et al. 2023; Bertacco et al. 2022; Görtz-Dorten et al. 2022; Hwang et al. 2022; James et al. 2022; Kornfield et al. 2022; Leary et al. 2022; Lungu et al. 2022; Mayer et al. 2022; Miura et al. 2022; Moral-Munoz et al. 2022; Nebot et al. 2022; Chandrasiri et al. 2021; Gire et al. 2021; Vonk et al. 2021; Burley et al. 2020; Ghaznavi et al. 2020; Lattie et al. 2020; Newbold et al. 2020; Bhattacharya et al. 2019; Bonn et al. 2019; Heffernan et al. 2019; Rosario, Mariano and Samonte 2019; Varnfield et al. 2019; Weber, Lorenz, and Hemmings 2019; Kennard et al. 2018; Bidargaddi et al. 2017; Wang et al. 2016; Antezana et al. 2015)
	Intermediate	(Currey and Torous 2022; Tsirmpas et al. 2022; Vairavasundaram et al. 2022; Ward et al. 2022; Germain et al. 2021; Rohani et al. 2020, 2021; Leightley et al. 2020; Mukhiya et al. 2020a; Ghandeharioun et al. 2019; Schlosser et al. 2018; Clarke, Jaimes and Labrador 2017; Neumeier et al. 2017; Kop, Hoogendoorn and Klein 2014; Noble 2013)
	Strong	(Currey and Torous 2023; Lewis et al. 2022; Riese et al. 2021; Luan et al. 2020; Khwaja et al. 2019; Rabbi et al. 2018; Xin et al. 2017; Hung et al. 2015; Lanata et al. 2015)
	N/A	(Ben-Yehuda et al. 2022; Figueroa et al. 2022; Shalaby et al. 2022; Tang et al. 2022; Athanas et al. 2021; Khwaja et al. 2021; Sellak and Grobler 2020)
Level of Interest	Interesting method	(Currey and Torous 2022; Görtz-Dorten et al. 2022; Leary et al. 2022; Lewis et al. 2022; Miura et al. 2022; Tang et al. 2022; Tsirmpas et al. 2022; Vairavasundaram et al. 2022; Chandrasiri et al. 2021; Germain et al. 2021; Gire et al. 2021; Riese et al. 2021; Ghaznavi et al. 2020; Leightley et al. 2020; Luan et al. 2020; Mukhiya et al. 2020a; Khwaja et al. 2019; Varnfield et al. 2019; Clarke, Jaimes and Labrador 2017; Neumeier et al. 2017; Xin et al. 2017; Wang et al. 2016; Antezana et al. 2015; Hung et al. 2015; Kop, Hoogendoorn and Klein 2014; Noble 2013)

(*Continued*)

Continued.

Classification		Studies
	Interesting evaluation	(Chaturvedi et al. 2023; Hwang et al. 2022; Kornfield et al. 2022; Lungu et al. 2022; Mayer et al. 2022; Shalaby et al. 2022; Athanas et al. 2021; Lattie et al. 2020; Bonn et al. 2019; Heffernan et al. 2019; Kennard et al. 2018)
	Interesting outcome	(Bertacco et al. 2022; Figueroa et al. 2022; James et al. 2022; Khwaja et al. 2021; Bhattacharya et al. 2019; Ghandeharioun et al. 2019)
	Multiple areas	(Currey and Torous 2023; Ben-Yehuda et al. 2022; Moral-Munoz et al. 2022; Nebot et al. 2022; Ward et al. 2022; Rohani et al. 2020, 2021; Vonk et al. 2021; Burley et al. 2020; Sellak and Grobler 2020; Rosario, Mariano and Samonte 2019; Weber, Lorenz and Hemmings 2019; Rabbi et al. 2018; Schlosser et al. 2018; Bidargaddi et al. 2017; Lanata et al. 2015)
	N/A	(Newbold et al. 2020)