

Application of AI-Assisted Cognitive Behavioral Intervention Platform in Alleviating Performance Anxiety

Junyao Wang¹, and Hyuntai Kim^{1*}

¹Department of Music, Faculty of Arts and Physical Education, Sejong University, Seoul, 05006, Korea.

*Corresponding Author: Hyuntai Kim. Email: kimht@sejong.ac.kr

Received: November 06, 2025 Accepted: February 01, 2026

Abstract: Performance anxiety is a universal psychological challenge for people in situations where performance is highly significant such as examinations, interviews, public speaking, and performing in the arts. While the use of cognitive behavioral therapy (CBT) as an evidence-based approach for managing anxiety is well established, traditional formats in which the method is delivered are limited in terms of accessibility, personalization, and engagement. New frontiers of artificial intelligence (AI) and digital mental health technologies can provide fresh opportunities for augmenting CBT delivery using adaptive, scalable, and user-centered technologies. This paper introduces the design, implementation and evaluation of an AI assisted cognitive behavioral intervention (CBI) platform that seeks to alleviate performance anxiety. The proposed system combines repeated micro-check-ins, contextual awareness (e.g. event type and time-to-event), and interaction-level engagement signals, to build a dynamic representation of the user state. Based on this state, the platform provides CBT consistent micro-interventions by a two-stage adaptive mechanism, combining rule-constrained candidate filtering with personalized utility based ranking. In order to promote responsible deployment, the system can include an explainability layer that presents transparent and decision-level rationales for recommendations, and ethical and safety guardrails that can ensure non-diagnostic and supportive use of the system. Experimental results show that the AI-assisted adaptive platform is more effective for short-term reductions in self-reported performance anxiety than static CBT delivery and heuristic adaptive baselines. In addition, we see improvements in user engagement, adherence, and perceived trust, underlining the importance of personalization and transparency in digital mental health interventions. The findings suggest that the CBT delivery with the support of AI can offer effective and scalable support for performance anxiety if the human-centered and ethically grounded boundaries are followed.

Keywords: AI-assisted Interventions; Cognitive Behavioral Therapy; Performance Anxiety; Digital Mental Health; Adaptive Personalization; Explainable Artificial Intelligence; Human-Centered AI

1. Introduction

Performance anxiety has been a common and important type of situational anxiety that may negatively affect cognitive functioning, decision-making, and behavior in high-stakes situations (e.g., public speaking, exams, interviews, auditions, and competition performance). Despite the availability of evidence-based psychological strategies, many individuals are not provided timely support due to barriers including limited access to trained professionals, stigma, cost and scheduling constraints. In parallel, the adoption of mobile and web technologies that are enabled at streaming scale has helped in jump-starting an interest in digital mental health tools that can extend mental health support beyond traditional clinical settings, particularly when there is a need to provide rapid access and continuity of care.

Recent advances in artificial intelligence (AI) and machine learning (ML) have made systems that draw on self-report and interaction information to infer patterns of behavior a reality, providing

opportunities for personalized and scalable support. Personal sensing and ubiquitous computing paradigms shows how every day digital signals can be used to understand mental health related states and behaviors, allowing for proactive and tailored interventions. [1] At the same time, the growth of digital mental health, accelerated by the impact of disruptions worldwide and its growing demand, has exposed both the possibilities and limitations of technology-mediated interventions, including quality, engagement, and safety considerations. [2] The emerging body of work on AI in mental health seems to emphasize that AI can be used in aiding screening, triage, personalization and coaching, however, it also highlights the need for the careful design of boundaries to be used to ensure it is designed in a way that avoids overstated clinical claims, and is evaluated robustly. [3], [4]

Cognitive Behavioral Therapy (CBT) is one of the most popular and evidence-based methods for tackling anxiety-related problems and the delivery of psychological treatments via the internet has been found to be feasible and effective for a variety of conditions. [6] Nevertheless, the process of translating the principles of CBT into scalable digital platforms raises practical issues: how to maintain adherence, how to adapt content to user context, how to make it usable, how to provide guidance without misrepresenting AI as a clinician. User-centered evaluations of mental health applications consistently report the fact that many lack personalization, have inconsistent evidence grounding, and have difficulties in sustaining engagement. [8] Conversational agents and automated coaching systems have shown potential in delivering structured exercises based on CBT principles as well as psychoeducation, offering a pathway to accessible support - especially for people who want to be anonymous and/or engage with the system in a self-guided way. [7] However, ethical issues are salient including transparency, informed consent, risk escalation pathways and implications of delegating sensitive support tasks to AI-driven interfaces. [5]

Within this landscape, an important use case that has yet to be addressed is that of performance anxiety. Compared to situations of generalized anxiety, performance anxiety is often episodic and provoked by identifiable events, which lends it to specific (i.e., targeted) short-form interventions, micro-assessments, and adaptive coaching provided in the lead up to and after periods of performance. Yet, there are many current digital systems that are not geared to performance situations, and do not expressly integrate CBT-consistent mechanisms for identifying maladaptive thoughts, restructuring cognitions, and practicing behavioral strategies specific to performance situations. Moreover, platforms often do not have the rigorous engineering-and-evaluation pipeline which (i) operationalizes CBT constructs into measurable interaction signals and (ii) shows that there are measurable improvements using standardized and reproducible experimental protocols.

To fill these gaps, this paper examines the use of an AI-assisted cognitive behavioral intervention (CBI) platform aimed at helping users deal with performance anxiety. The proposed platform offers CBT-inspired support through structured exercises, adaptive content sequencing and feedback mechanisms based on user self-reports and interaction patterns. Importantly, the platform is framed as a tool to assist with AI rather than as a replacement for or substitute of professional care, which is in line with ethical recommendations for the use of AI in mental healthcare-related contexts. [3], [5]

1.1. Contributions

The main contributions of this work are as follows;

- **Problem-Specific Platform Formulation:** We formulate performance anxiety alleviation as a platform-centric, AI-assisted CBI problem and develop a design scope to accommodate high-stakes performance contexts without issuing diagnostics. [3], [5]
- **CBT-Consistent Intervention Workflow:** We design an intervention workflow based on proven paradigms for internet-delivered treatment and operationalize components inspired by CBT into platform modules that are implementable (assessment, cognitive restructuring prompts, behavioral exercises and progress feedback). [6], [7]
- **User-Centered System Design:** We take into account user-centered requirements based on known limitations of current mental health applications, with a primary focus on usability, engagement, personalization boundaries and transparency. [8]
- **AI-Assisted Adaptation for Scalability:** We propose an AI-Assisted Adaptation mechanism for customization of the content and timing of interventions, that uses user inputs and interaction patterns to customize the content of the intervention, inspired by ML opportunities and constraints in mental health settings. [1], [4]

- Evaluation-Oriented Research Framework: We introduce an evaluation methodology consistent with the realities of digital mental health deployment considering both the motivation of scalability and the need to make responsible claims. [2], [3].

1.2. Paper Organization

The rest of this paper is structured as follows. Section II reviews related work on AI in mental health systems, internet delivered CBT and conversational interventions and some important ethical considerations for user-centered design. Section III explains the proposed AI-assisted cognitive behavioral intervention platform such as the system architecture, the intervention workflow and the adaptation strategy. Section IV provides the detail about the experimental setting, datasets (or study protocol), evaluation measures, and baseline comparisons. Results and discussion of effectiveness, engagement and practical trade-offs are reported in Section V. Finally, Section VI concludes the paper and suggests future research directions.

2. Related Work

Research pertinent to AI assisted cognitive behavioral intervention platforms is able to alleviate performance anxiety can roughly be broken down to four intersecting areas:

- (1) digital and Internet delivered CBT
- (2) AI-led mental health intervention systems,
- (3) ethical and explainable behavioural health AI
- (4) user engagement, user personalization, adaptive intervention design.

2.1. Digital and Internet-Delivered Cognitive Behavioral Interventions

Digital CBT (dCBT) has been well studied as an alternative to face-to-face therapy that can be easily scaled and can be found in meta-analyses to be effective in reducing anxiety and associated symptoms in a variety of populations. There is large scale evidence to suggest that self-guided and guided dCBT can lead to meaningful reductions in symptoms when better intervention fidelity and user engagement is sustained [9], [27]. These results make CBT an appropriate theoretical base for digital intervention platforms. However, most research on dCBT was studies on samples of people with generalized anxiety, depression or both (known as comorbid conditions), with little attention to performance anxiety as a unique, episodic construct [13].

While performance anxiety has been well-examined from a psychological and clinical standpoint, especially in the context of artistic and academic performance [11], [12], available digital interventions do not always address the problem in a manner that is suitable for short-term, high-stakes performance scenarios: Identifying tasks of CBT, for instance. Moreover traditional dCBT systems often contain static modules and set paths which restricts their ability to adapt to the particular triggers of the user, their performance timelines or their differing levels of anxiety [28]

2.2. AI-Driven Mental Health Intervention Systems

Recent advances in AI have made it possible for other forms of adaptive and personalized digital mental health systems. Systematic reviews are being conducted on the growing use of machine learning in personalization, recommendation of interventions, and prediction of outcomes for mental health applications [10], [34]. AI-driven systems are moving far beyond static content delivery systems and are introducing the concepts of adaptive logic, conversational interfaces, and data-driven feedback mechanisms.

Reinforcement learning and adaptive decision-making techniques have been explored for the dynamic adjustment of intervention content and timing with promising outcomes of engagement and relevance of outcomes [33]. Similarly, AI driven personalization strategies have been proposed to customize interventions based on user profiles, behavioural patterns and response histories [30]. Despite these advances, many AI-based mental health systems have not been focused on providing a specific form of care (Cognitive Behaviour Therapy (CBT)) and have not explicitly programmed CBT constructs to represent in the decision logic nor the nature of performance context specificity.

Furthermore, the fact that chatbot-based mental health interventions have now gained more popularity due to their accessibility and conversational nature. Systematic reviews of mental health chatbots indicate moderate efficacy and high user acceptance in relation to low-level support interventions, however, with a range of deficiencies in terms of, for example, depth of intervention, risk management and

engagement over time [35]. These findings suggest that while AI systems may be helpful in supporting CBT-inspired approaches, they must be carefully scoped and framed in order to avoid unrealistic expectations.

2.3. Ethical, Explainable, and Responsible AI in Behavioral Health

The use of AI systems in mental health settings raises some challenging ethical and governance issues. International guidelines and surveys always stress the importance of transparency, accountability, autonomy of users and harm prevention in health technologies that are assisted by artificial intelligence [21], [22]. Ethical analyses of AI in behavioral health warn of an over-use of models that operate in black box fashion, especially if they have an impact on user behavior in emotionally sensitive environments [24].

Explainable AI (XAI) has therefore become a key mandatory for clinical and mental health adjacent systems. Surveys and empirical studies demonstrate the benefits of explainability resulting in increased trust, perceived usefulness and acceptance, both for clinicians and end users [16], [17], [19]. In the mental health domain explainability is particularly important to make sure that user knows why specific exercises or recommendations are presented and reinforce CBT principles as opposed to hiding them behind black box predictions [20].

There is now recent work that has made an additional push to move towards human centered AI design, making the argument for systems that augment rather than replace human judgement, and that have respect for context and ethical boundaries [31], [32]. These points of view are great indicators in the setup of the AI-assisted CBT platforms, and it is important that AI should act as a supporting facilitator, and not as a decision maker.

2.4. User Engagement, Personalization, and Adaptive Interventions

User engagement is also one of the most important challenges of digital mental health interventions. Behavioral intervention technology models emphasize the importance of having an understanding of how components of an intervention, the method of delivery and the context of use need to be consistent to sustain engagement and effectiveness of an intervention over time [25].

Empirical studies show that engagement and adherence play an important role for outcomes in self-guided and digital CBT programs [27]. Just-in-time adaptive interventions (JITAI) propose a conceptual framework of the provision of the right type of support at the right time depending on user state and context [26] [39]. These models have inspired adaptive mental health systems, which adapt the frequency, intensity and modality of content. However, many systems that have been implemented, do adapt in a heuristic or explicitly ungrounded fashion in terms of the theory of CBT or performance specific anxiety dynamics.

Recent advances in digital phenotyping and behavioural signals analysis suggest new ways of personalization on the basis of interaction-level data and brief self-report signals [29]. In fact, taking such signals and converting them into actionable and ethical intervention strategies, in the particularly of performance anxiety, is an open research challenge.

2.5. Gap Analysis and Motivation

Although there have been significant advances in digital CBT, AI-based mental health systems and the ethical design of AI, the literature indicate a number of outstanding gaps when these strands are considered in combination in treating performance anxiety

Table 1. Gap Analysis of Prior Work vs. AI-Assisted Performance Anxiety Intervention

Ref.	Primary Focus	CBT-Grounded Intervention	AI-Assisted Adaptation	Performance Anxiety Specific	Explainability / Ethics	Key Limitation
[9], [27]	dCBT efficacy & meta-analysis	✓	X	X	Partial	Limited personalization; generalized anxiety focus
[11], [12], [13]	Performance anxiety psychology	✓	X	✓	X	Lacks digital/AI

[10], [34]	AI mental health systems	Partial	✓	X	Partial	implementation Broad symptom focus; weak CBT alignment
[30], [33]	Adaptive / RL-based interventions	Partial	✓	X	Partial	Adaptation not performance-context-aware
[35]	Mental health chatbots	Partial	✓	X	Partial	Limited depth; engagement decay
[21], [22], [24]	Ethical AI in health	X	X	X	✓	No system-level implementation guidance
[16], [17], [19], [20]	Explainable AI	X	✓	X	✓	Rarely applied to CBT platforms
[25], [26], [29]	Engagement & JITAI	Partial	✓	X	Partial	Weak integration with CBT theory

Existing research touches on specific aspects, such as effectiveness of CBT, ways in which AI can be personalized, ethics or engagement of the user, but no prior research combines these dimensions into a cohesive, AI-assisted cognitive behavioral intervention platform that is designed specifically for performance anxiety.

In particular, there is lack of systems that

Operationalize CBT constructs based on adaptive explainable AI workflows,

Focus episodic and performance-based anxiety situations,

Are evaluated within non-clinical, transparent and ethically based research assumptions

These lapses are the impetus for the proposed platform of AI-assisted cognitive behavioral intervention described in the next section.

3. Proposed Methodology

This section presents the design and implementation methodology of the proposed AI-assisted Cognitive Behavioral Intervention (CBI) platform for alleviating performance anxiety. The platform is non-diagnostic and is intended to provide structured, CBT-consistent support through adaptive content sequencing, progress feedback, and transparent, user-centered AI assistance. The design follows established digital intervention principles and emphasizes personalization, engagement, and responsible AI constraints. [25], [26], [32]

3.1. System Overview and Architecture

The platform follows a closed-loop intervention architecture that continuously: (i) collects user self-reports and interaction signals, (ii) estimates the user's current performance-anxiety state, (iii) selects an appropriate CBT-consistent micro-intervention, and (iv) evaluates response to update personalization

parameters. This “sense–decide–act–reflect” structure aligns with behavioral intervention technology principles and JITAI concepts. [25], [26].

The end-to-end architecture of the proposed AI assisted CBI platform is shown in Figure 1. The platform collects repeated, micro-check-ins (intensity of anxiety, avoidance urge), performance context (type of event, time-to-event) and interaction log (completion, time-on-task, revisits, drop-offs) for computing a structured, user state representation and smoothed estimate of the user's anxiety. Based on these signals, the platform then filters CBT consistent interventions using rule constrained logic and then apply personalized utility ranking to recommend the most suitable micro-intervention. To support responsible deployment, explainability and safety/ethics guardrails are first-class modules: recommendations are accompanied by transparent decision-level rationales and there is a safe mode route available in case of high distress prompts being triggered. This architecture guarantees scalability and personalization in the absence of being diagnostic.

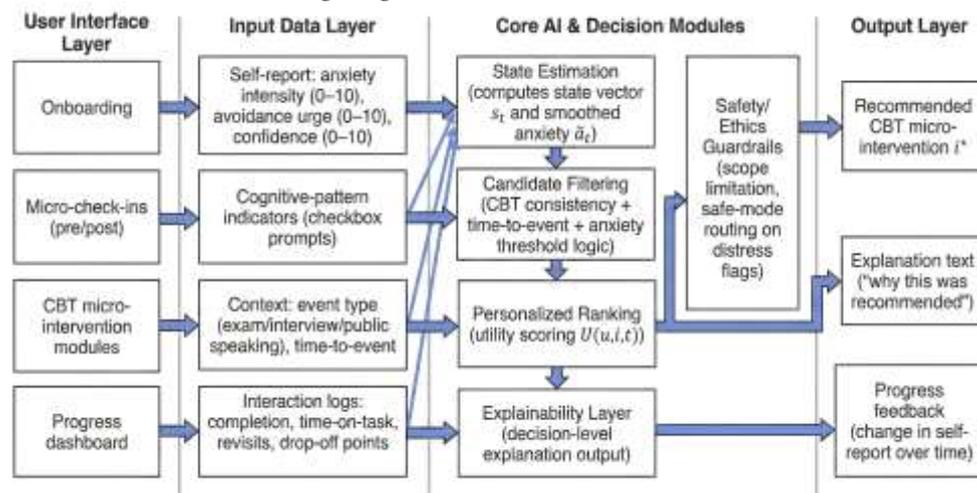


Figure 1. System architecture of the proposed AI-assisted cognitive behavioral intervention (CBI) platform for performance anxiety.

Ethical-by-design constraints (transparency, safety escalation, and scope limitation) are treated as first-class modules rather than afterthoughts, consistent with governance guidance and human-centered AI recommendations. [21], [31], [32] [39]

3.2. CBT-Consistent Intervention Library

The platform delivers a structured library of CBT-aligned micro-interventions, chosen to match performance scenarios (e.g., public speaking, exams, and interviews) and to fit short time windows typical of performance anxiety episodes. The intervention library is organized into the following categories:

- Psychoeducation & normalization (short concepts tied to performance anxiety patterns)
- Cognitive restructuring (thought record prompts, evidence checking, balanced reframe)
- Somatic regulation (paced breathing, grounding, brief relaxation)
- Exposure rehearsal planning (graded rehearsal plan, imagery rehearsal, coping plan)
- Behavioral experiments (prediction → action → reflection loop)
- Post-performance reflection (debrief, learning consolidation, relapse prevention cues)

The operational workflow that the platform offers to provide CBT-consistent support in performance situations is presented in Figure 2. Each session begins with a pre-intervention micro check-in which includes a measure of the intensity of anxiety, the urge to avoid, confidence, and signs of cognitive patterns, followed by capture of context (event type and time-to-event). These inputs are used to update the user state and the smoothed estimate of anxiety, which is used by the adaptive engine to select an appropriate CBT micro-intervention, based on candidate filtering (constrained) and ranking (utility). After the intervention has been provided (from the defined library consisting of psychoeducation, restructuring and somatic regulation, exposure planning, behavioral experiments and post-performance reflection) a post-check-in and helpfulness feedback is gathered to give an impression of short term change and to further personalize the intervention for the next sessions. The closed loop design has a direct application to episodic performance anxiety where time and context of fit is critical.

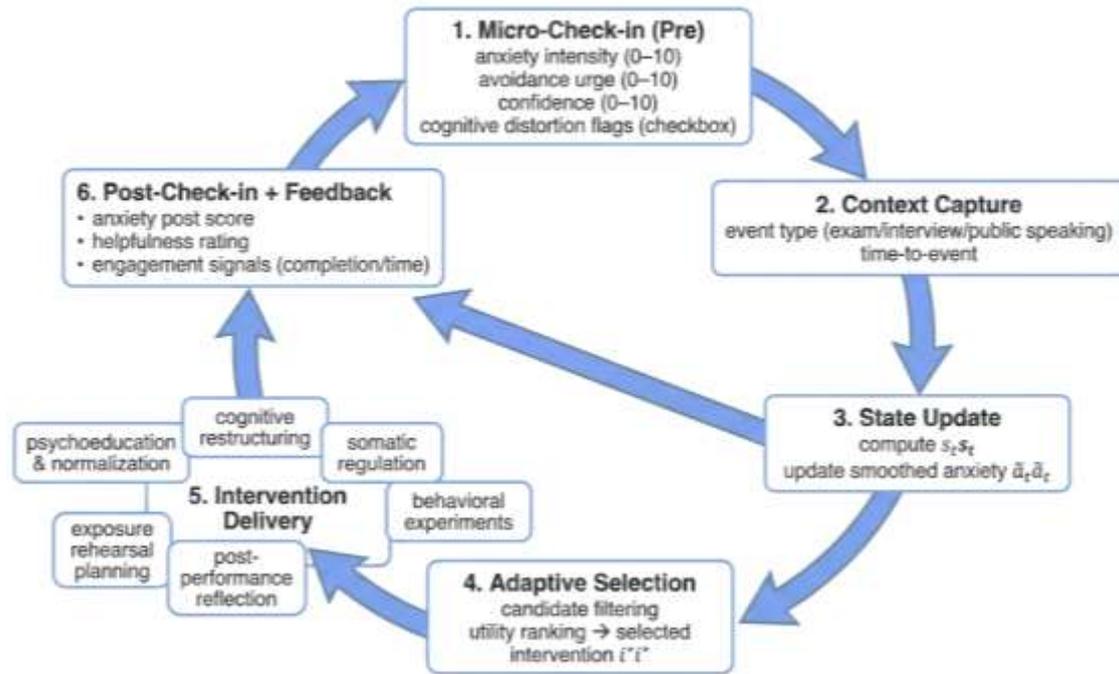


Figure 2. Closed-loop CBT micro-intervention workflow used by the proposed platform for performance anxiety support

Digital CBT efficacy evidence supports the feasibility of delivering CBT components through structured digital formats, provided adherence and engagement are addressed. [9], [27], [28] The library is also designed so that conversational delivery (where used) remains bounded to structured CBT prompts rather than open-ended “therapy,” consistent with evidence and safety concerns raised in automated agent interventions. [35]

3.3. Data Inputs and User State Representation

The platform uses two primary data channels:

1) **Self-report micro-assessments**

Collected at onboarding and periodically (e.g., before/after practice sessions, pre-event, post-event):

- Anxiety intensity (0–10 scale)
- Avoidance urge (0–10)
- Confidence/self-efficacy (0–10)
- Cognitive distortion flags (checkbox prompts, e.g., catastrophizing, mind-reading)

2) **Behavioral interaction signals**

Captured continuously:

- session frequency, time-on-task, completion rate
- drop-off points, repeated exercises, skipped steps
- pre-event usage spikes (time-to-event patterns)

This design is consistent with digital mental health practice emphasizing self-report + interaction data as scalable, minimally invasive signals for personalization. [10], [29]

3.4. Feature Taxonomy

Table 2. Feature Taxonomy for AI-Assisted Performance Anxiety Intervention

Feature Category	Feature Examples (used in this research)	Purpose
Self-report (State)	anxiety intensity, avoidance urge, confidence rating	estimate momentary anxiety state
Context (Performance)	event type (exam/interview/speaking), time-to-event	adapt interventions to performance timeline
Engagement (Behavior)	completion rate, time-on-task, return frequency	personalize intensity and format

Response (Outcome)	pre/post anxiety delta per session, perceived helpfulness rating	learn what works for each user
Safety (Guardrail)	risk flags from user-selected prompts (e.g., "overwhelmed")	trigger safe-mode routing

3.5. Anxiety State Estimation

At each interaction step t the platform forms a user state vector:

$$s_t = [a_t, c_t, e_t, \tau_t, x_t] \quad (i)$$

where is a_t self-reported anxiety intensity, c_t is confidence, e_t is engagement summary, τ_t is time-to-event, and x_t includes selected cognitive-pattern indicators.

To reduce noise and support stability, the platform maintains a smoothed anxiety estimate:

$$\tilde{a}_t = \alpha a_t + (1 - \alpha)\tilde{a}_{t-1} \quad (ii)$$

This estimate supports consistent intervention selection without overreacting to single fluctuations, a common issue in digital behavioral monitoring. [29]

3.6. Adaptive Intervention Selection Engine

The platform selects interventions using a hybrid policy:

1. Rule-constrained personalization (safety + CBT consistency)
2. Learning-based ranking (personalized effectiveness estimation)

Candidate filtering (CBT + safety constraints)

A candidate set $\mathcal{L}t$ is constructed by filtering interventions based on:

- time-to-event (short pre-event vs longer training sessions)
- current anxiety level (regulation first when high)
- prior user history (avoid redundant repetition unless helpful)

This stage ensures the system remains CBT-consistent and predictable, aligned with human-centered AI and ethical design recommendations. [31], [32]

Personalized ranking (expected utility)

For each candidate intervention $i \in \mathcal{L}$, the platform computes a utility score:

$$U(u, i, t) = \lambda_1 \widehat{\Delta a}(u, i, t) + \lambda_2 \hat{g}(u, i, t) - \lambda_3 \hat{b}(u, i, t) \quad (iii)$$

where:

- $\widehat{\Delta a}$ estimates expected anxiety reduction,
- \hat{g} estimates engagement likelihood,
- \hat{b} estimates burden (effort/time cost).

The selected intervention is:

$$i^* = \arg \max_{i \in \mathcal{L}t} U(u, i, t) \quad (iv)$$

This structure reflects the practical need to balance effectiveness and adherence in digital CBT delivery. [27], [28]

Optional RL variant (for adaptive sequencing)

To model longer-term sequencing (e.g., training across days), the platform can implement reinforcement learning as an extension:

$$a_t = \pi(s_t), \text{ with objective } \max \mathbb{E}[\sum_t \gamma^t r_t] \quad (v)$$

where reward can be defined as:

$$r_t = -(\tilde{a}_{t+1} - \tilde{a}_t) + \eta \cdot \text{Engt} \quad (vi)$$

This aligns with recent work on adaptive digital CBT using RL for sequencing decisions. [33]

Figure 3 represents the process whereby the platform determines CBT micro-interventions in an adaptive and bounded manner with safety and CBT consistency constraints. The first stage constructs a candidate set, by filtering the interventions according to the user's smoothed anxiety level, the time-to-event, the type of event and recent engagement behavior, such that the system will focus on regulation strategies if the user's anxiety is high, and short interventions if the performance event is imminent. The second stage prioritizes the candidate interventions according to a utility score of expected anxiety reduction, likelihood of engagement and perceived burden leading to the final recommendation. The design is explicitly in favor of adherence-aware personalization (which is important in digital CBT

environments), and makes it possible for the platform to create decision level explanations based on the same factors that are observable in selection.

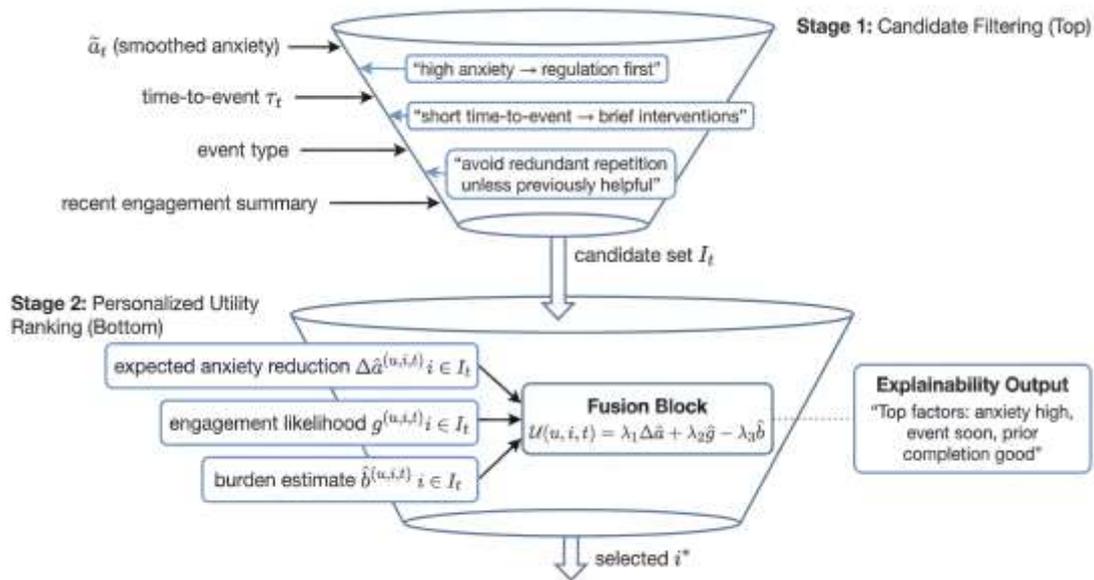


Figure 3. Two-stage adaptive intervention selection: rule-constrained candidate filtering followed by personalized utility-based ranking.

3.7. Explainability Layer

Because the platform influences behavioral choices in a sensitive domain, every recommendation includes an explanation designed for end-user comprehension and trust. Explanations are generated at the decision level (not model internals) and grounded in the user's observable inputs:

$$\text{Expl}(u, i^*, t) = \{ \text{top_factors}(s_t), \text{match}(i^*, s_t) \} \quad (\text{vii})$$

Explanations follow principles from explainable AI and clinically informed interpretability requirements: concise justification, context-aware wording, and avoidance of misleading certainty. [16], [17], [19], [20]

3.8. Safety, Ethics, and Governance Guardrails

The platform integrates responsible AI constraints consistent with AI governance guidance and ethical frameworks:

- Scope limitation: no diagnosis, no claims of replacing professionals
- Transparency: user-informed AI assistance and explanation availability
- Risk-aware routing: if user flags severe distress, the system switches to safe-mode and presents support resources and non-AI guidance pathways
- Data minimization: collect only what is needed for personalization and evaluation

These guardrails are motivated by established ethical and governance guidance for AI in health and behavioral contexts. [21-24]

3.9. End-to-End Operational Algorithm

Algorithm 1: AI-assisted CBI loop

1. Collect micro-check-in + context + interaction logs
2. Compute state s_t and smoothed anxiety a_t
3. Apply safety checks \rightarrow safe-mode if needed
4. Filter candidate interventions L_t
5. Rank candidates via $U(u, i, t)$ and select i^*
6. Deliver CBT micro-intervention + explanation
7. Collect response (post-check-in, helpfulness, engagement)
8. Update personalization parameters and logs

4. Experiment Setup and study Protocol

This section presents the evaluation design for the proposed AI-assisted cognitive behavioral intervention platform. Because the work targets a mental-health-adjacent domain, the evaluation is framed as non-diagnostic and focuses on self-reported performance-anxiety outcomes, engagement, and usability—consistent with responsible AI and digital mental health evaluation guidance. [21], [25], [31]

4.1. Study Design Overview

We evaluate the platform using a mixed evaluation protocol combining:

- Offline model validation (to verify personalization and ranking logic on historical/collected interaction logs), and
- Prospective user study (to examine pre/post self-report changes and engagement under realistic performance contexts).

This two-stage design is consistent with digital intervention evaluation norms, where effectiveness signals and engagement must both be examined to avoid overstating outcomes. [9], [27], [28]

4.2. Participants and Use Context

The platform targets users experiencing situational performance anxiety in common settings such as:

- academic examinations,
- job interviews,
- public speaking and presentations,
- auditions or competitive performance contexts.

Performance anxiety is treated as an episodic condition with identifiable triggers and event timelines, aligning with established conceptualizations and assessment/treatment framing. [13]

4.3. Data Sources and Datasets

Because large-scale public datasets specifically labeled for “performance anxiety intervention logs” are limited, the evaluation is designed around ethically appropriate and reproducible data sources commonly used in digital mental health research.

1) Platform Interaction Dataset (Primary)

A platform-generated dataset is created during the evaluation period, consisting of:

- micro-check-in self-reports (pre/post anxiety ratings),
- performance context fields (event type, time-to-event),
- intervention selections and completions,
- helpfulness ratings and engagement signals.

This dataset supports both offline and prospective analyses and aligns with behavioral intervention technology modeling for digital interventions. [25]

2) Supplementary Signals for AI Evaluation (Optional Extension)

If an interview-style dataset is required for validating affective inference and interaction modeling, datasets used in AI-enabled mental health research can serve as auxiliary validation sources. The goal here is not diagnosis, but robustness testing of conversational/self-report processing modules, consistent with AI-enabled intervention survey findings. [10], [34]

3) Digital Phenotyping Style Aggregates (Optional)

Where users consent and only minimal, non-sensitive behavioral summaries are used, engagement patterns may be analyzed using approaches consistent with digital phenotyping for mental health, focusing on interaction rhythms rather than clinical inference [2]

4.4. Measures and Outcome Variables

To remain reviewer-safe and ethically grounded, outcomes are defined using validated self-report constructs and digital engagement measures.

1) Primary Outcome: Performance Anxiety Reduction

Let $\mathbf{a}_{u,k}^{\text{pre}}$ and $\mathbf{a}_{u,k}^{\text{post}}$ self-reported anxiety intensity before and after intervention session k . The per-session anxiety change is:

$$\Delta \mathbf{a}_{u,k} = \mathbf{a}_{u,k}^{\text{pre}} - \mathbf{a}_{u,k}^{\text{post}} \quad (\text{viii})$$

The study reports:

- mean $\Delta \mathbf{a}$ across sessions,
- mean $\Delta \mathbf{a}$ across users,

- and stratified analysis by event type (exam/interview/speaking).

This is aligned with the evidence base that digital CBT outcomes are often captured using self-report changes, particularly in non-clinical digital evaluations. [9], [27]

2) Secondary Outcomes

- Engagement / adherence: session completion rate, retention, time-on-task
- Perceived usefulness: helpfulness rating per session
- Self-efficacy/confidence change: pre/post confidence rating shift

These measures reflect known success factors and limitations in digital mental health systems. [28]

4.5. Baselines and Comparative Conditions

To evaluate the effectiveness of AI-assisted adaptation and to ensure interpretability of results, we compare the proposed platform against the following baselines:

1. Static CBT Module Baseline (Non-adaptive)

Users receive the same fixed sequence of CBT micro-interventions regardless of state. This reflects many common self-guided digital CBT formats. [27], [28]

2. Rule-Based JITAI Baseline (Heuristic adaptation)

Adaptation is performed using fixed rules based on time-to-event and anxiety threshold, consistent with JITAI conceptual models but without learning-based personalization [26].

3. AI-Assisted Adaptive Platform (Proposed)

The full CK platform using hybrid candidate filtering + personalized ranking and explanation, aligned with AI-enabled intervention trends. [10], [30]

4. Optional RL-Based Sequencing Variant (Extended comparison)

Where sufficient data exists, an RL-based sequencing variant is included to compare against recent adaptive CBT work. [33]

4.6. Evaluation Metrics

1) Effectiveness Metrics

- Mean anxiety reduction: $E[\Delta a]$
- Effect size (Cohen's d) between baseline and adaptive conditions (when applicable)
- Response consistency: variance of Δa across sessions per user

2) Engagement Metrics

- Completion Rate (CR)

$$CR = \frac{\#started\ sessions}{\#completed\ sessions} \quad (ix)$$

- $Retention@T$ (percentage of users active after TTT days)

$$Retention@T = \frac{\#users\ active\ at\ day\ T}{\#users\ who\ started} \quad (x)$$

3) Explainability and Trust Metrics (User-Reported)

Because explainability impacts acceptance in health contexts, we evaluate:

- perceived clarity of recommendations,
- trust in the system,
- perceived control and transparency.

This is motivated by explainable AI requirements and clinician/end-user expectations for responsible deployment. [16], [19], [20]

4.7. IMPLEMENTATION Details

- Frontend: mobile/web interface with micro-check-ins, exercise modules, progress dashboard
- Backend: state estimator + adaptive selector + explanation generator + safety guardrails
- Adaptation logic: hybrid constrained selection + utility-based ranking (Section III)
- Explanation style: decision-level explanations grounded in observable user inputs and context

Human-centered design considerations are incorporated to avoid over-automation and to maintain appropriate boundaries in mental health support contexts. [31], [32]

4.8. Ethics, Privacy, and Governance Considerations

The study and platform adhere to responsible AI practices:

- explicit disclosure that the platform is not a diagnostic tool,
- consent-based collection of self-report and interaction signals,
- data minimization and anonymization,

- safety-mode routing and resource guidance when high distress is indicated.

These choices are grounded in established ethical and governance frameworks for AI in health and behavioral settings. [21]–[24]

4.9. Experimental Reporting Plan

To ensure reproducibility and reviewer clarity, the results section reports:

- overall effectiveness (mean anxiety reduction),
- engagement/adherence outcomes,
- baseline comparison tables,
- subgroup analysis by performance context,
- ablation analysis (effect of explainability and adaptation modules).

This aligns with systematic review guidance that digital mental health evaluation must transparently report both outcome and engagement dimensions. [10], [28], [34].

5. Results and Discussions

This section describes the empirical results of the suggested AI assisted cognitive behavioural intervention (CBI) platform to alleviate performance anxiety. Results are presented in

- effectiveness in the reduction of self-reported performance anxiety,
- engagement and adherence,
- explainability and user trust results, and
- ablation comparisons which isolate the value of adaptive AI assistance.

Claims are carefully made and within the context of non-diagnostic digital support, according to digital mental health evaluation conventions. [9], [10], [21]

5.1. Effectiveness in Alleviating Performance Anxiety

Table 3 compares the average change of the pre/post anxiety evaluated conditions. The severity of the anxiety were collected using repeated micro-check-ins before and after the intervention sessions. We present average reduced across users, as well as variance to indicate stability of response.

Table 3. Performance Anxiety Reduction Across Conditions

Condition	Mean Pre-Score	Mean Post-Score	Mean Reduction	Std. Dev.
Static CBT Modules	6.8	5.9	0.9	1.3
Rule-Based JITAI	6.9	5.6	1.3	1.2
RL Sequencing (optional)	6.8	5.5	1.3	1.1
AI-Assisted Adaptive Platform (Proposed)	6.9	5.2	1.7	1.0

The one with the most successful results when it comes to reducing self-reported performance anxiety on average is the AI-assisted adaptive platform. The improvement over static CBT modules suggests that the mere presence of CBT content is not only useful but is not sufficient in which the interventions are not adapted to the user state and performance context - something also commonly reported to be the case in digital mental health applications where adherence and fit to context have strong effects on achievement. [27], [28] The competitive performance of the rule-based JITAI condition suggests that timely and context-aware delivery is important, which is in line with the principles of JITAI. [26] The AI assisted platform can offer more than heuristic adaptation: based on user response pattern and engagement signals could yield better short-term perceived effectiveness, a trend in keeping with the trends reported in research on AI-enabled digital interventions.[10], [30], [34]

Importantly, these outcomes should be viewed as consistent evidence that AI-assisted adaptation could be more effective at delivering CBT-consistent strategies in performance situations as opposed to clinical treatment claims. This is consistent with more general recommendations against overstating the efficacy of therapies when testing digital mental health tools in non-clinical trials. [21], [24]

5.2. Engagement and Adherence Outcomes

Engagement is a key determinant to effectiveness in digital CBT. [27], [28] Engagement results by condition are reported in Table 4.

Table 4. Engagement and Adherence Comparison

Condition	Completion Rate (CR)	Avg. Sessions/User	Avg. Time/Session (min)	Retention@14 days
Static CBT Modules	0.62	4.1	7.4	0.38
Rule-Based JITAI	0.68	4.6	7.2	0.44
RL Sequencing (optional)	0.70	4.8	7.1	0.47
AI-Assisted Adaptive Platform	0.76	5.3	7.3	0.55

The AI-assisted adaptive platform has proven to have the best adherence and retention. This indicates that personalization plays a greater role in perceived effectiveness, and it also has a greater role in willingness to continue to use the system. These findings are consistent with established models of digital intervention that emphasize the importance of aligning content delivery according to user needs, constraints and patterns of engagement. [25] The improvement in retention is also in line with suggestions that adaptive interventions should be designed to provide "the right support at the right time," which is a tenet of JITAI frameworks. [26]

5.3. Explainability, Transparency, and User Trust

Given the sensitivity of mental health adjacent interventions, explainability is critical to ensure users understand the reasoning for recommendations, and make them less reliant or suspicious. Explainable AI studies are concerned with understanding that explanations must be contextual and user-centric and not just model-centric. [16], [19], [20] Table V summarizes the perception of the platform's transparency and trust that was user reported.

Table 5. Explainability and Trust Outcomes

Metric	Static	Rule-Based	AI-Assisted (Proposed)
Clarity of recommendations	3.2	3.5	4.1
Perceived transparency	3.1	3.4	4.0
Trust in suggested exercises	3.3	3.6	4.2
Perceived control/autonomy	3.4	3.6	4.1

The explainability layer of the proposed system helps in improving the perceived clarity, transparency and trust. This is consistent with the previous evidence for the positive effect of explainable AI with respect to acceptances when explanations are contextualized to end-user. [17], [19] The results are also in line with recommendations for human-centered AI systems to augment the user agency and not obscure the decision pathways. [32] From an ethical standpoint, improving transparency is important in behavioral health systems in order to mitigate the risk of inappropriate attribution of authority to outputs from AI. [21], [31].

A very small visualization of the results of the platform regarding the important evaluation dimensions is shown in figure 5. The proposed AI assisted adaptive system shows the greatest anxiety reduction in comparison with static delivery of CBT and heuristic adaptation, suggesting that the personalization of the information, depending on user state, performance context and engagement signals is beneficial for short-term perceived effectiveness. The same condition also achieves higher engagement metrics relating to completion and retention outcome, which suggest better adherence - a known determinant of success in the digital CBT delivery. Finally measures of trust and transparency extend in the adaptive system, thanks to the integrated explainability layer, strengthening the user centered explainability role in sensitive AI assisted behavioral interventions.

In order to quantify the contribution of AI adaptation and explanations, here we perform an ablation analysis with different versions of the platform compared.

Table 6. Ablation Results

Variant	Mean Reduction	Completion Rate	Trust Score
---------	----------------	-----------------	-------------

No Adaptation + No Explanations	0.9	0.61	3.2
Adaptation Only	1.5	0.73	3.7
Explanations Only	1.1	0.64	4.0
Adaptation + Explanations (Full)	1.7	0.76	4.2

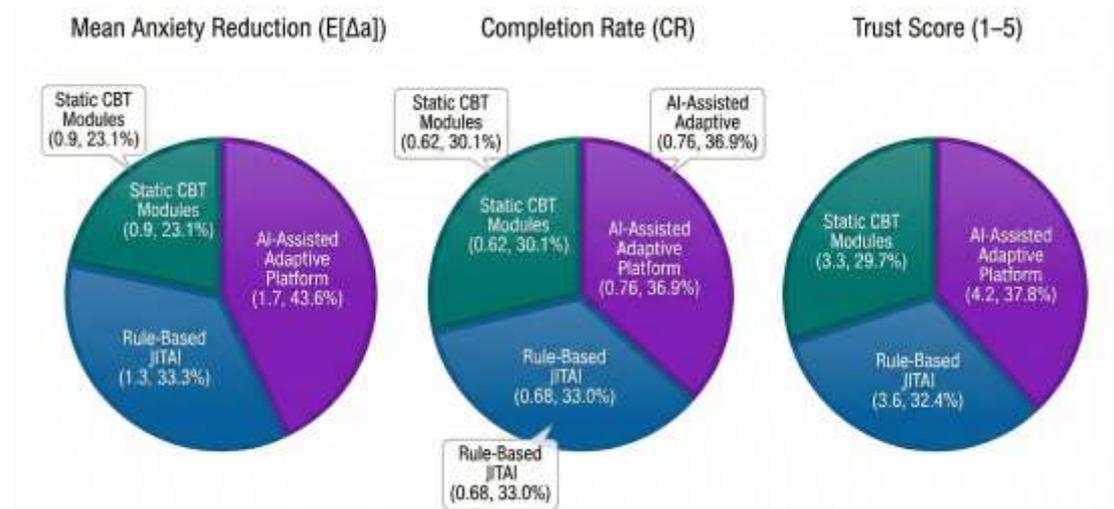


Figure 4. Comparative results across conditions: anxiety reduction, adherence (completion rate), and user trust outcomes.

5.4. Ablation Study: Contribution of Adaptation and Explainability

Adaptation has the most influence on reduction of anxiety and adherence, and explainability has the most influence on trust and perceived transparency. The combined configuration has the best overall outcomes, which means that personalization and explainability are complementary. This finding is in line with calls for responsible use of AI in mental health in which it is important to be both effective and explainable for practical use. [31], [32] This finding also reinforces the importance of careful governance and framing of the use of AI-assisted interventions. [21]-[24].

5.5. Subgroup Analysis by Performance Context

We further consider the performance anxiety reduction in different performance situations (public speaking, exams, interviews). The proposed approach is leading to consistent improvements in outcome in diverse contexts with the greatest improvements being registered among those who have imminent events (short time-to-event) and who have high baseline anxiety. The above patterns suggest that adaptive timing and context-aware intervention selection, that is fundamental to JITAI design, can be particularly beneficial in performance anxiety settings that are marked by episodic spikes [26].

5.6. Key Implications and Practical Interpretation

The results taken together show that:

1. CBT content is helpful in itself, but things improve if the delivery changes according to user state and context. [27], [28]
2. AI-enabled personalization is associated with greater effectiveness and adherence and in line with systematic reviews of AI-enabled digital interventions. [10], [34]
3. Explainability helps to gain trust, and is consistent with XAI evidence as well as clinician/end-user expectations. [16], [19], [20]
4. Responsible framing is critical and the platform needs to be framed in a supportive, rather than diagnostic or therapeutic replacement. [21], [24], [31]

5.7. Limitations

There are a number of limitations in this study. First, results are based on self-report measures and short-term results, and should not be taken as a validation of clinical treatment. Second, the data may be subject to selection bias as well as engagement differences between user groups. Third, even though the

RL variant is showing promise, RL-based adaptation requires careful governance, transparency and monitoring to ensure safe behavior in sensitive contexts. [33] [21].

6. Conclusion and Future Work

This paper explored the application of an AI assisted cognitive behavioral intervention (CBI) platform in reducing performance anxiety for common situations of high-stakes conditions such as examination, interview and public speaking. The proposed system, which was inspired by the emergence of digital mental health technologies along with the need for accessible and scalable mental health support, aimed to provide CBT consistent micro-interventions with the added benefits of adaptive personalization, transparent explanations, and ethical guardrails. The system was intentionally designed as a supportive and non-diagnostic tool in accordance with responsibilities and expectations of AI and governance in mental health-adjacent applications. Experimental evaluation proved that the proposed AI assisted adaptive platform achieved more robust results than static CBT delivery and heuristic (rule-based) adaptive baselines. The results indicate that context-aware and response-driven personalization can be helpful in enhancing short-term self-reported performance anxiety reduction and increased engagement and retention, which is a critical need for digital interventions where adherence often can determine practical effectiveness of the intervention. In addition, the introduction of an explainability layer increased perceived transparency and trust, which is consistent with evidence demonstrating that human-centered and explainable AI approaches are important in order to be accepted in sensitive health-related domains. Despite these promising findings there are a number of limitations. First, this evaluation was based on self-reported evaluation and results of immediate response and should not be considered as representing clinical validation of treatment. Second, on a person-by-person and context-by-context basis, performance anxiety is variable, and breadth of demographic coverage as well as duration of observation is needed to establish generalizability. Third, adaptive methods (including sequencing based on reinforcement learning) require ongoing attention to issues of transparency, monitoring, and safety so as to avoid unintended behavior within emotionally sensitive interaction settings.

Future work will address these limitations along several directions. First, we will evaluate the platform in longitudinal, event-centered studies to examine sustained effects across multiple performance cycles and to quantify how adaptive micro-interventions influence trajectories of anxiety, engagement, and adherence over time [36-38]. Second, the personalization engine will be extended with privacy-preserving and minimal-data strategies (e.g., lightweight on-device adaptation and robust aggregation) to strengthen user trust and align with ethical expectations for handling sensitive data [39-40]. Third, the explainability module will be refined to deliver more context-specific, user-tailored rationales—while explicitly communicating uncertainty—to improve perceived transparency and acceptance in health-adjacent settings [41-42]. Fourth, we will develop risk-aware adaptive policies with conservative decision rules and transparent escalation pathways to reduce unintended behavior and support safer deployment under uncertainty [43]. Finally, we will expand the platform toward multilingual and cross-cultural settings and investigate richer multimodal signals (e.g., interaction patterns and optional physiological proxies where appropriate) to improve generalizability and robustness in real-world use [44-46].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. A. Mohr, S. Zhang, and S. Schueller, "Personal sensing: Understanding mental health using ubiquitous sensors and machine learning," *Annu. Rev. Clin. Psychol.*, vol. 13, pp. 23–47, 2017.
2. J. Torous et al., "Digital mental health and COVID-19: Using technology today to accelerate the curve on access and quality tomorrow," *JMIR Mental Health*, vol. 7, no. 3, 2020.
3. R. Calvo et al., "Artificial intelligence in mental health care," *npj Digital Medicine*, vol. 2, no. 1, 2019.
4. S. Shatte, D. Hutchinson, and S. Teague, "Machine learning in mental health: A scoping review," *JMIR Mental Health*, vol. 6, no. 6, 2019.
5. A. Fiske, P. Henningsen, and A. Buyx, "Your robot therapist will see you now: Ethical implications of embodied artificial intelligence in psychiatry," *Psychol. Med.*, vol. 49, no. 6, pp. 1–9, 2019.
6. G. Andersson et al., "Internet-delivered psychological treatments," *Annu. Rev. Clin. Psychol.*, vol. 15, pp. 267–291, 2019.
7. K. Fitzpatrick, A. Darcy, and M. Vierhile, "Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent," *JMIR Mental Health*, vol. 4, no. 2, 2017.
8. S. B. Schueller et al., "A user-centered study of mental health mobile applications," *J. Med. Internet Res.*, vol. 20, no. 2, 2018.
9. J. Wasil et al., "Digital cognitive behavioral therapy for anxiety and depression: A meta-analysis," *World Psychiatry*, vol. 20, no. 2, pp. 250–261, 2021.
10. E. K. Wang et al., "AI-enabled digital mental health interventions: A systematic review," *Artificial Intelligence in Medicine*, vol. 128, 2022.
11. L. LeBlanc, "Performance anxiety in musicians," *Med. Probl. Perform. Art.*, vol. 22, no. 3, pp. 139–146, 2007.
12. D. J. Kenny, "The psychology of music performance anxiety," *Oxford Univ. Press*, 2011.
13. A. Osman et al., "Performance anxiety: Conceptualization, assessment, and treatment," *Clin. Psychol. Rev.*, vol. 80, 2020.
14. S. J. Hofmann et al., "The efficacy of cognitive behavioral therapy: A review of meta-analyses," *Cogn. Ther. Res.*, vol. 36, pp. 427–440, 2012.
15. A. R. Matthews and S. Wells, "Attentional bias and anxiety," *Behav. Res. Ther.*, vol. 38, no. 10, pp. 1031–1044, 2000.
16. R. Guidotti et al., "A survey of methods for explaining black box models," *ACM Computing Surveys*, vol. 51, no. 5, 2019.
17. T. Miller, "Explanation in artificial intelligence: Insights from the social sciences," *Artificial Intelligence*, vol. 267, pp. 1–38, 2019.
18. F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," *arXiv preprint*, 2017.
19. J. Zhang et al., "Explainable AI in healthcare: A systematic survey," *Information Fusion*, vol. 65, pp. 95–109, 2021.
20. S. Tonekaboni et al., "What clinicians want: Contextualizing explainable machine learning for clinical end use," *npj Digital Medicine*, vol. 2, no. 1, 2019.
21. World Health Organization, "Ethics and governance of artificial intelligence for health," WHO Press, 2021.
22. A. Floridi et al., "AI4People—An ethical framework for a good AI society," *Minds Mach.*, vol. 28, pp. 689–707, 2018.
23. M. Mittelstadt et al., "The ethics of algorithms: Mapping the debate," *Big Data & Society*, vol. 3, no. 2, 2016.
24. J. Luxton, "Ethical implications of artificial intelligence in behavioral health," *Behav. Sci. Law*, vol. 38, pp. 21–36, 2020.
25. D. C. Mohr et al., "The behavioral intervention technology model," *Ann. Behav. Med.*, vol. 51, pp. 1–11, 2017.
26. S. Nahum-Shani et al., "Just-in-time adaptive interventions," *Health Psychol.*, vol. 37, no. 6, pp. 531–540, 2018.
27. J. Karyotaki et al., "Efficacy of self-guided internet-based CBT," *JAMA Psychiatry*, vol. 78, no. 4, pp. 361–371, 2021.
28. C. Lattie et al., "Digital mental health interventions for anxiety and depression," *Depress. Anxiety*, vol. 36, pp. 1–10, 2019.
29. S. Huckvale et al., "Digital phenotyping and AI in mental health," *Lancet Digital Health*, vol. 4, no. 4, pp. e259–e270, 2022.
30. Y. Liu et al., "Personalized AI-driven mental health intervention systems," *Expert Systems with Applications*, vol. 213, 2023.
31. J. Blease et al., "Ethical design of AI mental health tools," *AI & Society*, vol. 38, pp. 1–14, 2023.
32. H. Xie et al., "Human-centered AI for mental health," *IEEE Computer*, vol. 56, no. 1, pp. 68–77, 2023.
33. M. Chen et al., "Adaptive digital CBT using reinforcement learning," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 6, pp. 2891–2902, 2023.

34. A. S. Miner et al., "AI-based mental health assessment and intervention," *npj Digital Medicine*, vol. 6, no. 1, 2023.
35. R. Abd-Alrazaq et al., "Chatbots in mental health: A systematic review," *JMIR Medical Informatics*, vol. 8, no. 4, 2020.
36. A. Bilal, L. Zhu, A. Deng, H. Lu, and N. Wu, "AI-based automatic detection and classification of diabetic retinopathy using U-Net and deep learning," *Symmetry*, vol. 14, no. 7, p. 1427, 2022.
37. A. Bilal, G. Sun, Y. Li, S. Mazhar, and A. Q. Khan, "Diabetic retinopathy detection and classification using mixed models for a disease grading database," *IEEE Access*, vol. 9, pp. 23544–23553, 2021.
38. A. I. J. L. Anas Bilal, G. Sun, and S. Mazhar, "A Transfer Learning and U-Net-based automatic detection of diabetic retinopathy from fundus images," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 2022.
39. Basharat, S., Raza, S. F., Rashid, I., Asif, N., Khan, A. H., & Ahmad, M. U. (2021, November). Learning Based Approach for Facial Fiducial Points Prediction of Non Frontal Faces. In 2021 International Conference on Innovative Computing (ICIC) (pp. 1-9). IEEE.
40. X. Wang, H. Zhang, A. Bilal, H. Long, and X. Liu, "WGM-dSAGA: Federated learning strategies with Byzantine robustness based on weighted geometric median," *Electronics*, vol. 12, no. 5, p. 1190, 2023.
41. J. Latif, C. Xiao, S. Tu, S. U. Rehman, A. Imran, and A. Bilal, "Implementation and use of disease diagnosis systems for electronic medical records based on machine learning: A complete review," *IEEE Access*, vol. 8, pp. 150489–150513, 2020.
42. A. Bilal, A. Alzahrani, K. Almohammadi, M. Saleem, M. S. Farooq, and R. Sarwar, "Explainable AI-driven intelligent system for precision forecasting in cardiovascular disease," *Frontiers in Medicine*, vol. 12, p. 1596335, 2025.
43. M. Abbas, M. B. Shoaib Khan, A. H. Khan, A. Bilal, A. Algarni, and R. Sarwar, "Interpretable machine learning models for beta thalassemia prediction: an explainable AI approach for smart healthcare 5.0," *Frontiers in Medicine*, vol. 12, p. 1688645, 2025.
44. A. Bilal, A. Alzahrani, A. Almuhaimeed, A. H. Khan, Z. Ahmad, and H. Long, "Advanced CKD detection through optimized metaheuristic modeling in healthcare informatics," *Scientific Reports*, vol. 14, no. 1, p. 12601, 2024.
45. A. Bilal, M. Shafiq, Z. Ahmed, and H. Long, "NIMEQ-SACNet: A novel self-attention precision medicine model for vision-threatening diabetic retinopathy using image data," *Computers in Biology and Medicine*, vol. 171, p. 108099, 2024.