Design Thinking the Human-AI Experience of Neurotechnology for Knowledge Workers*

Troy R. Weekes[1[0000−0002−0371−1997]] and Thomas. C. Eskridge[1,2[0000−0002−2117−529]]

L3 Harris Institute for Assured Information,
Department of Human-Centered Design,
Florida Institute of Technology, Melbourne FL 32901, USA
tweekes@fit.edu, teskridge@fit.edu

Abstract. Neurotechnology promises cognitive enhancement as a way for humanity to extend its information-processing capability without invasive brain surgeries and pharmacological side effects. Notable advancements in this field have achieved high-bandwidth wireless communication interfaces between human brains and computers. Human-centered design proposes that human-technology experiences should focus on human needs. This paper explains how design thinking has been applied as a methodology to design the user experience of an attention-based neurotechnology solution that leverages artificial intelligence (AI) to enhance the flow performance and cognitive well-being of knowledge workers (KWs). Using the d.school design thinking process, we started with a mindset that favored empathy, creative confidence, and ambiguity to discover and define the problems confronting KWs. After diverging with deep empathy and converging on user personas and problem definition, the design thinking process branched into an iterative prototyping cycle that transformed our initial ideas into a human-centered AI-powered neurotechnology. We utilized the functional prototypes for testing assumptions and performing a comprehensive design evaluation. Our final solution incorporated a gamified user interface with visual elements, affordances, and a coherent human-AI experience. Expert software evaluators conducted a series of cognitive walkthroughs and heuristic evaluations by simulating the user personas and performing an end-to-end user scenario with the prototype. The design thinking process generated a neurotechnology service with a human-AI experience that enables KWs to achieve healthy flow performance while enhancing cognitive well-being.

Keywords: Knowledge Worker · Neurotechnology · Design Thinking.

* Supported by L3 Harris Institute for Assured Information.

1 Introduction

A growing body of literature on neurotechnology recognizes the importance of bio-sensing and biofeedback [44, 40, 17]. In this paper, the term 'neurotechnology' refers to the methods and instruments that enable a direct connection with the nervous system [34]. Recent studies by computational neuroscientists have shown how neurotechnologies use bio-sensing electrodes to record signals from the brain and transform them into biofeedback displays of useful control commands and stimuli [31, 24]. This exploratory research focused on human knowledge workers (KWs) and attention-based neurotechnology-as-a-service shown in Figure 1.

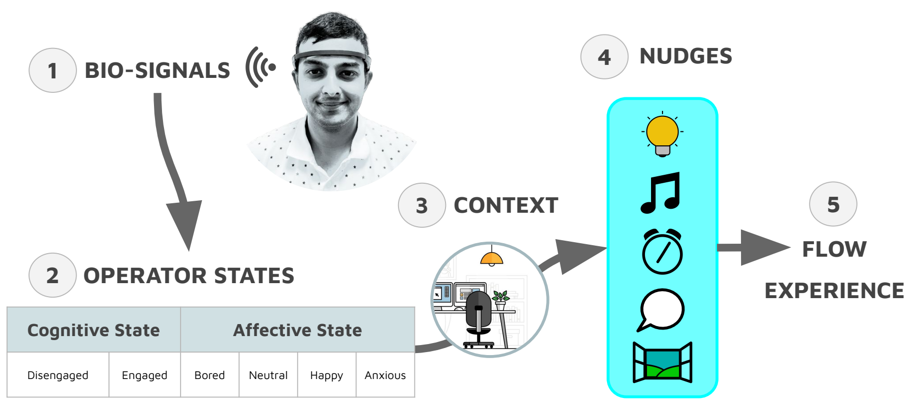

Fig. 1. FCA bio-senses, contextualizes, and nudges KWs into flow

In this work, we design a human-aware and context-aware neurotechnology artificial intelligence (AI) known as the **Flow Choice Architecture (FCA)** that "nudges" [51, 46] KWs to increase the healthy time that they spend in the flow state [10, 11, 13]. Our design thinking story outlines the development of FCA's human-AI experience to strengthen the cognitive abilities of KWs by deepening their levels of cognitive work rather than automating their jobs.

KWs are essential to maintain our standard of living and quality of life. Their well-being is paramount to economic development and human advancement. Since the pandemic caused by the coronavirus disease, KWs have experienced an accelerated shift towards remote working [3] in virtual and hybrid work environments [49] that are augmented by AI [54]. To remain competitive, KWs need to create more value in less time while improving their performance and maintaining their well-being.

In Section 2, "Background," we review the literature on cognitive enhancement, neurofeedback, and healthy flow performance. This section analyzes how neurofeedback can be adopted to help KWs effortlessly focus their attention on the task stimulus during knowledge work.

In Section 3, “Design Thinking,” we describe the phases of the design thinking process used in this paper. We empathize by conducting remote, semi-structured interviews with 12 KWs to generate qualitative results, quotes, and insights about their personal experiences, expectations, and preferences related to knowledge work. We validate the interview results by conducting an online survey with 468 participants. We articulate the results from a need-finding synthesis, which were consistent problem statements evidenced by supporting user personas. These data-driven design assets guided the future steps of the design thinking process. We explore the qualitative dataset to discover commonly used vocabulary and consolidate an intuitive information architecture. We adopt rapid prototyping techniques to build a functional FCA for testing and evaluation.

In Section 4, “Evaluation,” we conduct cognitive walkthroughs and heuristic evaluations of the prototype with six evaluators who are subject matter experts in knowledge work and software engineering. The evaluation results generated improvements for future development.

In Section 5, “Results,” we present the findings from the cognitive walkthroughs and heuristic evaluations. The evaluation results highlight how FCA succeeded on the tasks and identify areas for improvement.

In Section 6, “Discussion,” we explain the implications of the findings and discuss recommendations for improving the learnability and usability of FCA.

In Section 7, “Conclusion,”, we conclude with the research outcomes. We discuss the next steps to advance the FCA neurotechnology prototype to become a beneficial tool for the cognitive enhancement of KWs.

2 Theoretical Background

FCA is a bio-sensing and contextual bio-feedback nudging system that enhances KW flow performance and cognitive well-being. FCA contributes to bridging the gap of growing global demand for more creative and productive human output in knowledge-based industries by helping KWs perform their cognitive work with fewer distractions and attentional load. The references synthesized in this section identify relevant findings from scholarly sources on cognitive enhancement, neurofeedback, and healthy flow performance.

2.1 Cognitive Enhancement

Cognitive enhancement aims to reach one’s personal best without necessarily outperforming others [9]. Bostrom and Sandberg [7] define cognitive enhancement as “ the amplification or extension of core capacities of the mind through improvement or augmentation of internal or external information processing systems.” Contemporary cognitive enhancement methods involve an array of nootropics, brain implants, brain training games, neurofeedback, and transcranial electric stimulation devices for modifying brain function [19, 21].

Despite the many positive effects of cognitive enhancement, there are likely negative aspects to be considered. Given that cognitive enhancements are likely to be used for extended periods across the lifespan, the long-term effectiveness and safety are crucial concerns to be determined.

In the knowledge economy, the value of human capital far outweighs more traditional, tangible forms, such as plants and equipment [35]. For this primary reason, we regarded the cognitive performance and well-being of KWs as quintessential elements of organizational success. We value the KW's happiness before, during, and after knowledge work in terms of its immediate and long-term impacts on the KW. By doing so, we designed FCA to guide KWs towards the flow state and enhance cognitive well-being through well-timed nudges and psychological flexibility routines that cultivate mindfulness and commitment.

2.2 Healthy Flow Performance

Csikszentmihalyi [10] defined flow as a state of concentration so focused that it amounts to absolute absorption in an activity. Concentration is a trainable cognitive state that may aid in the activation and maintenance of flow during goal-directed behaviors [48]. We hypothesize that concentration during knowledge work tends to activate flow. To this end, we designed FCA to train KWs to regulate their concentration and benefit from the positive effects experienced during and after flow.

On the contrary, fatigue is the debilitating cognitive state associated with feeling exhausted, sleepy, and tired, which diminishes the ability to function efficiently on a task [18]. Although work may be completed under conditions of high cognitive fatigue, the quality of performance and the quality of work outcomes tend to decrease [29]. Basic research found that an increase in cognitive fatigue correlated with increased reaction times, misses and false alarms, and time-on-task in an attention-dependent task [6]. Matthews and Desmond [32] observed the detrimental impact of cognitive fatigue on performance during highly demanding cognitive tasks. This observation makes the management of task demand an essential aspect for FCA to perform successfully and effectively.

Despite the rich literature on the topic of flow, these studies have been primarily qualitative inductive analyses [47]. Ambiguities exist in its definitions and inconsistencies are evident in how flow is operationalized [1]. In this research, we operationalized flow with the nine components defined by Csikszentmihalyi [10], which include challenge-skill balance, action-awareness merging, clear goals, unambiguous feedback, concentration on the task at hand, sense of control, loss of self-consciousness, time-transformation, and an autotelic experience.

Flow occurs when individuals, acting solo or in teams, operate with optimal concentration, which yields a heightened sense of satisfaction, intrinsic motivation, and peak performance [10, 12, 37]. In this work, we claim that healthy flow performance is not the excessive attainment of the flow state, which may lead to exhaustion and burnout, but sufficient flow to accomplish one's work while maintaining cognitive and emotional well-being.

2.3 Neurofeedback

Within the neurofeedback domain, training protocols provide audio-visual signals based on site-specific electroencephalography (EEG) frequency bands or combinations thereof [31]. EEG data make up a reliable bio-signal stream that may reify cognitive performance into measurable neural activity. FCA tests this hypothesis with its neurofeedback AI technology for flow augmentation by building upon other neurofeedback interventions grounded in the training of respondent and operant behaviors [17]. The main distinction between FCA and other neurofeedback tools is the use of intuitive and comprehensible nudges that reduce cognitive workload rather than signals that require monitoring and decoding [40].

There are varied results from experiments that correlate EEG to human performance. In a study by Katahira et al. [24], participants performed arithmetic tasks of varying difficulty levels to induce three conditions: flow, boredom, and overload. The researchers analyzed the variance of EEG data between the three conditions. Results from the study demonstrated that theta power in the brain's frontal areas was higher in the flow and overload conditions compared to the boredom condition. According to Katahira et al. [24], high theta power reflects the subjective states of maximum cognitive control and absorption in the task. The flow condition exhibited decreased alpha activity compared to the overload condition, which suggested a relatively low cognitive load on working memory during flow. This study concluded that the flow state was indicated by high frontal theta power and moderate alpha power [24].

Researchers have explored artificial neural networks and deep learning techniques to classify operator states using EEG signals. Wilson et al. [56] performed two-class cognitive workload classification based on artificial neural networks, and achieved 86% classification accuracy. Tripathi et al. [53] used a deep neural network and a convolutional neural network (CNN) to classify valence and arousal measures using EEG signals from the DEAP dataset [27]. Their neural networks provided 58% and 56% classification accuracy for valence and arousal, respectively, and their CNN model provided 67% and 58% for valence and arousal, respectively. Zheng et al. [57] used a deep neural network architecture to process EEG and eye movement features. The fusion of multimodal bio-signals with deep neural networks significantly enhanced the model's performance compared with a single modality, and the best mean accuracy of 85% was achieved for four emotional states [57]. Eskridge & Weekes [20] used the SEED-IV EEG dataset [57] to run dimensionality reduction on the power spectral density features from five EEG frequency bands using linear discriminant analysis followed by an artificial neural network to gain average overall classification accuracy of 99%.

This paper discusses the use of EEG bio-signals in the form of EEG power indices as reliable indicators for effort, concentration, relaxation, absorption, fatigue, arousal, and valence during knowledge work tasks. We computed seven EEG power indices for use in the analysis from evidence-based correlates in the literature of computational neuroscience [25, 5, 22, 30, 23, 22, 43, 5, 50, 36, 16, 6, 15, 8, 4, 14, 41, 45, 52, 2, 45, 26].

3 Design Thinking the Human-AI Experience

In this work, knowledge workers (KWs) are human workers who perform complex deep work [38] that requires considerable amounts of concentration and creativity. We adopted the d.school design thinking process model [42] in Figure 2 to actively engage the participation of KWs in the design and evaluation of FCA. The Institutional Review Board (IRB) at Florida Tech approved our research with human subjects.

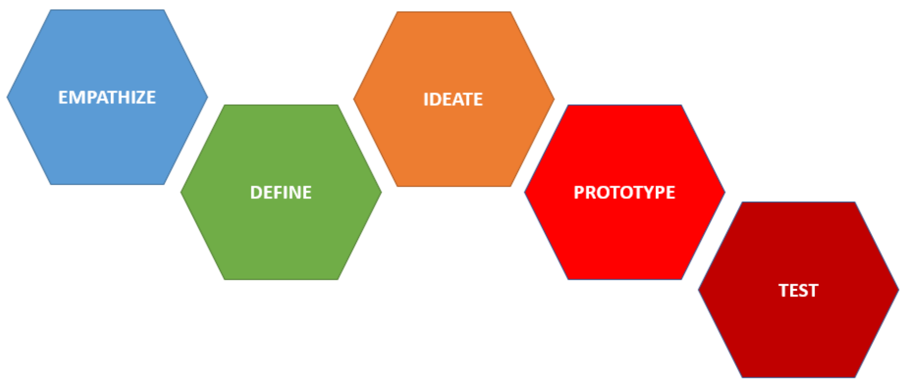

Fig. 2. The Five Stages of the d.school Design Thinking Process Model

3.1 Stage 1 - Empathize

In the "Empathize" stage, we formulated an understanding of potential FCA users, i.e., KWs seeking flow experiences. We conducted remote, semi-structured, one-on-one interviews with KWs and subsequently applied the interview results to develop a cross-sectional survey with KWs to generate quantitative data and validate insights from the interviews.

We conducted the interviews with a sample of 12 KWs (6 females, age M = 29.5 years [SD = 7.91]) who represented KWs from diverse domains in engineering, creative writing, project management, supply chain management, research, and philosophy. The interview study was generative in nature and centered on building a deep empathy with KWs as the starting point for the design thinking method of the innovation process.

The problem under investigation was: how might we describe the flow performance of KWs according to their tasks, workspaces, tools, and heuristics? We hypothesized that KWs explore and exploit their factors of production to maximize performance even though stressors may negatively impact their well-being. To test our claim, we pursued the following research questions. In which domains do KWs work? What types of tasks do KWs perform? Where do KWs perform their work? Which heuristics and tools do KWs use? What is flow at work for KWs? What is effective work time for KWs? How does KW productivity impact well-being?

The interview questions were derived from the research questions and grouped into the following ten sections: activity tracking, favorite activities, knowledge work, productivity, work tracking scenario, task definition & execution, flow at work, distraction, traction, and changing work states. After the interviews, we applied quantitative content analysis and qualitative thematic analysis to synthesize the results. The process generated a conceptual framework for classifying KWs based on distinct personal traits and work preferences. User personas relevant to productivity multiplier tools help humanize FCA.

The KWs worked in three main work domains, i.e., sciences, business, and arts. The types of knowledge work tasks that the KWs performed included writing, research, coding, design, documentation, reading, finding, visualization, collaboration, and managing subordinates. The types of tasks varied across work domains. The foremost task features that KWs considered included urgency, duration, challenge, importance, sensitivity, and priority.

KWs worked in the office, lab, cubicle, and living room. Participants reported the shift to working from home due to the COVID-19 pandemic. KWs described their efforts to create a distraction-free environment by turning off or stowing phones, wearing headphones, listening to music, and closing or locking doors. They take breaks to relax, relieve stress, refocus, eat, drink, and deflate.

KWs reported that flow was a zone, work mode, or head-space that occurs naturally and is goal-oriented, structure-driven, and distraction-free. The KWs supported the need for clear task goals and complete task absorption to achieve flow. Most of them experience a loss of self-consciousness and a faster passage of time during flow. KWs reported a positive feeling of satisfaction after flow.

KWs identified detrimental impacts of productivity on their well-being, e.g., procrastination, sleep issues, and developing hyper-focus and tunnel vision. Several KWs admitted to missing lunches, not drinking water, ignoring eating, and eating too quickly. Others complained about poor posture, being stationary at the desk, and lack of exercise. On the other hand, KWs identified some positive impacts of productivity on their well-being. KWs reported feeling in a better mood, confident, happy, and more energetic. Some KWs used the positive energy as an opportunity to perform activities outside of work.

A key finding was that KWs considered flow at work in terms of being in a zone and head-space when they are focused on making progress, completing tasks, and achieving results without interruptions and distractions. This finding has significant implications for the design of FCA to increase attention on the task at hand and mitigate external distractions. One of the most important findings to emerge from this study was that KWs balance the positive and negative impacts of productivity with their well-being to seek growth and happiness, which suggests a role for FCA in promoting healthy flow and work-life balance.

We conducted the cross-sectional survey with 468 KWs from MTurk to generate quantitative data, validate earlier insights, and understand what makes flow enjoyable for individual KSs. We used the survey results to identify user preferences and qualify excerpts for the user personas. The survey was randomized and cross-sectional in design to build inclusion across KWs from the three main work domains, i.e., sciences, business, and arts.

The survey yielded significant effects of specific situations on the KW's enjoyment of a task. Success on the initial attempt had an extremely positive effect on the KW's enjoyment of the task. The same positive effect occurred during situations of consistent and unexpected success over time. Contrary to expectations, KWs tended to embrace failure since consistent failure over time also has a significantly positive effect on the KW's enjoyment. The use of incentives for success also had a significantly positive effect on the KW's enjoyment.

In terms of how frequently KWs experienced the state of flow during knowledge work, consistent failure over time had a significantly negative effect on the KW's enjoyment. During flow, KWs felt that failure on the initial attempt had a significantly negative effect on the KW's enjoyment. Similarly, random failure over time had a significantly negative effect on the KW's enjoyment.

3.2 Stage 2 - Define

In the "Define" stage, we synthesized the research from the interviews and survey to discover where KWs were experiencing work-related problems that interfered with their performance and well-being. Our need-finding synthesis generated the user persona in Figure 3 and a set of user Point-Of-View (POV) problem statements, which guided the remaining stages of the design thinking process.

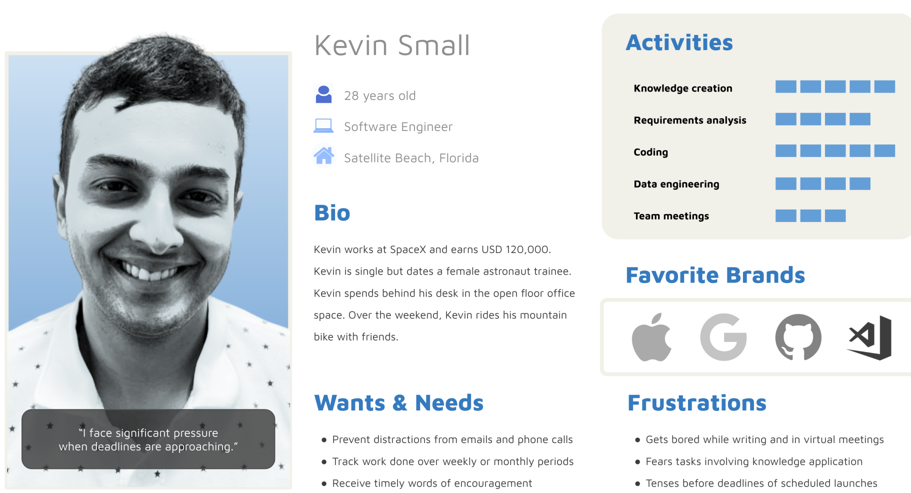

Fig. 3. The Software Engineer User Persona - Kevin Small

The following actionable Point-Of-View (POV) statements formulated contextualized problems confronting KWs and identified their needs and insights.

1. Before knowledge work, KWs need a way to prepare for work because they tend to procrastinate and lose focus without precise tasks or goals.
2. During knowledge work, KWs need a way to stay engaged with work because they tend to become distracted and stressed over lost productive time.
3. After knowledge work, KWs need a way to account for and reflect on work accomplishments because there are key work patterns to learn.
4. During boring knowledge work, KWs need a way to stimulate and challenge themselves because they tend to lose motivation and underperform.
5. During overwhelming knowledge work, KWs need a way to relax and calm down because they tend to become anxious and underperform.
6. During enjoyable knowledge work, KWs need encouragement and reinforcement because they tend to perform better and achieve more healthy flow.

3.3 Stage 3 - Ideate

In the "Ideate" stage, we generated creative ideas from the macro-scale to the micro-scale. We used AI to mine the qualitative datasets from the interviews and surveys to discover useful vocabulary, labels, and interactions that were recognizable and appealing to KWs. We abstracted the user scenarios into a generic task list where the interactions and user interface (UI) modeled a "minimalist" version of a task management system. We ideated the FCA through the lens of safe, explainable, and responsible AI.

The design philosophy of FCA's user experience exploited wearable technology that is non-invasive, lightweight, and easy to use. Once we obtained a comfortable hardware setup, the FCA operator needed a simple and effective neurofeedback UI. Our approach to FCA's UI design leveraged research about the operator's biases, behaviors, and preferences. Three "**flow principles**" were incorporated as fundamental tenets of FCA's design philosophy.

1. **Dynamic Flow - In flow, time stands still.** The dynamic visualization of deep flow was represented as minimal motion, whereas shallow flow was moderate motion, and distraction was significant motion.
2. **Cumulative Flow - More flow yields better work.** FCA rewarded the operator with flow points based on the flow state of each epoch. The cumulative visualization of flow applied a heuristic that humans employ, i.e., more is better.
3. **Deep Flow - Never interrupt deep flow.** FCA used a recommender system that delivered nudges based on specific learned criteria or when the system "explored" and tried something novel to learn new knowledge. However, a rule was that FCA would never interrupt "deep flow."

We leveraged the qualitative datasets from the interviews and surveys to construct the information architecture with abstracted keywords, i.e., profile, workspace, device, project, task, and work session, which formed the basis of the functional user requirements for the prototypes.

3.4 Stage 4 - Prototype

In the “Prototype” stage, we incrementally built out FCA by diverging and converging on multiple ideas. We developed four prototypes with different capabilities to answer questions and clarify risky assumptions.

Prototype 1 tested facial expressions as operator state indicators. One of the riskiest components of our research plan was discovering how to classify human bio-signals to provide reliable operator state indicators. This question became a core focus of Prototype 1, which focused on extracting steady streams of facial landmarks and action units to train models that predict the operator’s state. Prototype 1 provided insights into which states the predictive model computed from the camera feed overlay of facial landmarks.

Fig. 4. Prototype 1 classifying Facial Expression Bio-signals

Figure 4 shows the near real-time classification of facial expression bio-signals from an operator performing an experiment watching emotionally-charged videos. This initial prototype demonstrated that it was feasible to classify operator states given the facial expression bio-signal time series vector. After proving that it was possible to classify operator states with a measure of reliability using facial bio-signals, we advanced to the most challenging aspect of our research plan. We needed to discover how to simultaneously classify multiple bio-signals to provide reliable indicators of operator state.

Prototype 2 tested the integration of multimodal bio-signals from wearables devices. We selected two different wearable devices based on their capabilities. Muse headbands are affordable, commercially available off-the-shelf EEG devices developed by InteraXon Inc. Muse headbands aimed to enhance meditation practice by combining instruction and tracking with EEG sensor biofeedback during mindfulness exercises [28]. The Empatica E4 is a wrist-worn photoplethysmography (PPG) bio-sensor device that Empatica Inc developed. The E4 calculated heart rate, inter-beat-interval (IBI), and skin temperature [33].

Prototype 2 demonstrated that it was feasible to classify operator states from multimodal bio-signals. However, Prototype 2 was too slow to reach a consensus due to the various bio-signals with different timescales.

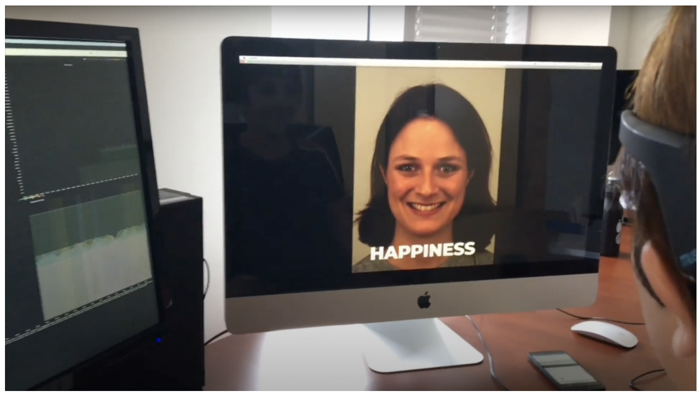

Fig. 5. Prototype 2 showing an Operator performing a Mirroring Experiment

Prototype 3 in Figure 6 tested EEG bio-signals as operator state indicators. Prototype 3 proved that it was feasible to compute operator states from EEG bio-signals. This finding supported our decision to focus on neurofeedback.

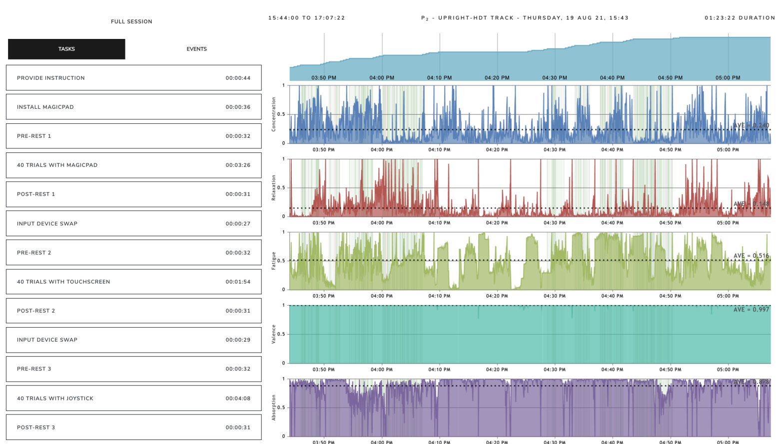

Fig. 6. Prototype 3 showing computed EEG Indices & Task Interval Markers

3.5 Stage 5 - Test

In the “Test” stage, we evaluated the effectiveness of the prototypes. We conducted cognitive walkthroughs and heuristic evaluations with KWs to inform the design of the human-AI experience. We developed Prototype 4 in Figure 7 to evaluate FCA as a gamified neurotechnology in cognitive walkthroughs and heuristic evaluations.

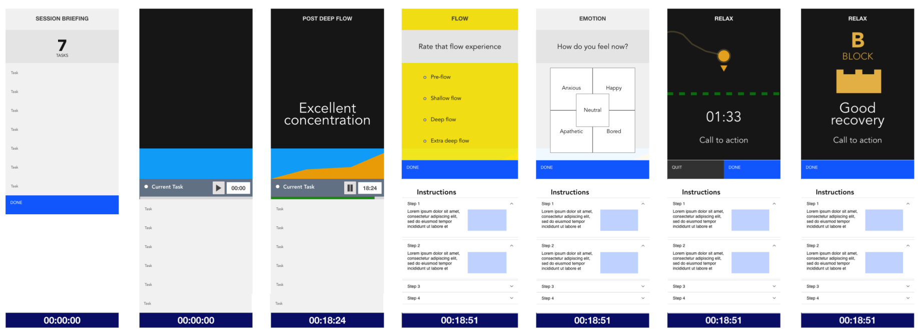

Fig. 7. Prototype 4 showing features of the FCA UI

The cognitive walkthrough was a detailed, step-by-step evaluation of FCA on a set of tasks. The purpose of the walkthrough was to empathize with KWs to uncover design errors in the FCA UI that would interfere with their learning by exploration and cause confusion during interactions. Examples of such errors are poorly worded labels, misguiding layout flows, and inadequate feedback about the consequences of an action.

The heuristic evaluations applied Jakob Nielsen’s usability heuristics [39]. The heuristic evaluations identified usability issues in the FCA UI for remediation. Responses from the evaluators were comments on the violations of the usability guidelines supplemented by severity ratings.

4 Evaluation

4.1 Methodology

The remote, one-on-one cognitive walkthroughs were conducted by 3 KW evaluators (2 females, age M = 26.3 years [SD = 4.61]). The evaluators simulated the personas and evaluated the FCA prototype from the perspective of the potential users. KWs started FCA and configured it to plan and complete tasks in a work session. After interpreting FCA’s UI and responding to nudges, the KWs completed and reviewed the work session, then shut down FCA. The KWs evaluated FCA by describing how the UI fulfilled each task.

The reseacher administered the cognitive walkthroughs. In the preparatory phase, the evaluator became familiar with the assigned user persona. The user personas from the "Define" phase allowed the evaluators to judge what needs, knowledge, preferences, and limitations the users may have relative to the tasks. During the walkthrough briefing, the researcher discussed the tasks to be analyzed. The evaluator interacted with FCA on the following tasks by describing how the UI performed on each task.

1. Startup FCA
2. Calibrate FCA
3. Plan a new work session
4. Start a work session
5. Complete a task
6. Interpret the signals during a task
7. Respond to the nudges during a task
8. Complete a work session
9. Review a completed work session
10. Shut down FCA

In the analysis phase of the cognitive walkthrough, the evaluators examined each action in the solution path and attempted to tell a credible story that explained why the expected users would choose that action. Credible stories were based on assumptions about the background knowledge of users and the problem-solving process that enables the user to guess the correct action.

If there was a major problem with the UI, the researcher noted the problem and proceeded to the next task as though the correct action had been performed. The state of the UI at the beginning of each action was always assumed to be the correct state and never the state after an incorrect action was performed.

The remote, one-on-one heuristic evaluations were conducted by 3 evaluators (3 males, age M = 39.0 years [SD = 17.32]). The evaluators simulated the Kevin Small persona over ten tasks. The KWs started FCA and configured it to plan and complete a work session. After interpreting FCA's UI and responding to nudges, they completed and reviewed the work session and then shut down FCA. The KWs evaluated FCA by describing how the UI fulfilled each task.

The preparatory phase of the remote heuristic evaluations involved a series of questions about FCA's compliance with Jakob Nielsen's ten usability heuristics. During the evaluation briefing, the researcher discussed the ten tasks above to be analyzed.

During the heuristic evaluations, the evaluators reviewed, interacted with, and evaluated FCA on the given tasks by describing how the UI performed on each task and then performing the correct action sequence to complete each task. In the analysis phase, the evaluators examined each task with each usability heuristic. If there were violations of the design guidelines, the evaluator made a descriptive comment and associated it with the task and the heuristic.

The researcher reviewed the comments with the evaluators where clarity was necessary. If there was a problem with the UI, the researcher noted the problem and proceeded to the next task, as if the correct action had been performed. The UI state at the beginning of each action was always assumed to be the correct state. The evaluators' comments were tabulated and severity ratings were assigned to each comment. In addition to the heuristic, each comment was categorized by the type of design error.

5 Results

The cognitive walkthroughs generated the following key findings and recommendations. We recommended button groups to replace the slider bars in the user profile, which negatively affected the configuration of FCA. We proposed to clarify what the UI does by changing the label "Manage Tasks" to "Plan Work-session." There should be a feedback screen that presents a meaningful summary of the work accomplished in the work session.

Figure 8 shows the tabulated results from the heuristic evaluations. There were 82 design issues covering the ten usability heuristics over the ten tasks. The totals indicated which tasks and heuristics contained a majority of the issues.

	STEPS										
HEURISTICS	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**TOTALS**
Visibility of system status	1	3	4	1	3	2	2		4		20
Match between system and the real world	1	2	3		2		1	1	3		13
User control and freedom		3	1	1	1	1	1		1		9
Consistency and standards		1	4	1	1	1			3		11
Error prevention	2	2	1	2			1			1	9
Recognition rather than recall		1							2	1	4
Flexibility and efficiency of use		1				1			1		3
Aesthetic and minimalist design	1										1
Help users recognize, diagnose, and recover from errors	1			1	1						3
Help and documentation	2	1	2	1	1	1	1				9
TOTALS	8	14	15	7	9	6	6	1	14	2	82

Fig. 8. Summary of Design Issues from the Heuristic Evaluations

6 Discussion

The relevance of flow among KWs strongly supported our findings. KWs considered flow at work in terms of being in a zone and head-space when they are focused on making progress, completing tasks, and achieving results without interruptions and distractions. The need for flow has significant implications for understanding how to design FCA in a way that increases their focused attention and mitigates distractions.

The cognitive walkthrough rationalized the design problems so that the FCA prototype would promote the discoverability and learnability of its users while providing adequate feedback on their tasks early in the implementation. Overall the cognitive walkthroughs demonstrated that FCA fits the KW's mental model. The concepts of user accounts, profiles, workspaces, devices, and tasks were all very familiar to the evaluators. The streamlined user flows to complete the UI interaction tasks guided the KWs from initial use onward.

The top-3 tasks that contained the majority of the design issues were tasks 3, 2, and 9, i.e., "Plan a new work session," "configure FCA," and "Review a completed work session." The top-3 heuristics that exhibited the highest frequencies were "Visibility of system status," "Match between system and the real world," and "Consistency and standards." The severity of the design issues factored into the prioritization of the fixes.

In addition to uncovering design issues that degrade the learnability and usability of FCA, this design thinking process reinforced the need for FCA to help KWs balance positive and negative impacts of productivity with their well-being. This finding underpins the primary goal of FCA to promote healthy flow performance and work-life balance.

7 Conclusion

This paper discussed the design thinking of a human-centered AI system that seeks to enhance the flow performance and well-being of individual KWs. We effectively applied the d.school design thinking process model to iteratively integrate lessons learned across the entire AI design and development life cycle.

The design thinking process reinforced that system design should start with the correct user to find the right problem. The application of design thinking to the human-AI experience of FCA involved a high level of sensemaking to decide which questions about the KW required clarity [55]. We leveraged samples of KWs and Amazon MTurk's pool of KWs to generate sufficient qualitative and quantitative data to ensure that FCA was developed to fit their needs.

Each prototyping cycle solved specific problems. Prototype 1 visualized outputs from the predictive model as time series and heatmaps. Prototype 2 tested the feasibility of classifying multimodal bio-signals and confirmed the decision to pursue a neurofeedback-based solution. Prototype 3 extended the neurofeedback approach and determined the efficacy of the computed EEG indices. Prototype 4 evaluated FCA in cognitive walkthroughs and heuristic evaluations.

In the context of FCA, an organization may mandate that individual KWs use FCA. The organization may wish to fire specific KWs if they do not achieve a high flow state for more than five hours a day. FCA proposes to protect against this type of misuse and abuse by treating bio-signal data as personal health information, defined as protected information under the Health Insurance Portability and Accountability Act (HIPAA). FCA digitalizes and scales the role of a personal workplace coach who helps KWs work healthier, happier, and more productively of their own volition.

The most challenging aspect of this research plan was developing a multimodal Prototype 2. This prototype required the simultaneous classification of multiple bio-signals to provide reliable indicators of the KW state. The prototype was too slow to reach a consensus due to various bio-signals with different timescales and inexplicable classifications. Another significant challenge was the downtime to train, test, and tune the AI models that punctuated the rapid exploratory iterations of the prototypes.

Future research includes experiments to streamline protocols, collect self-reports, and compute EEG indices. Other UI-focused advancements will center on a wider variety of flow-inducing conditions. We will continue to collect data from KWs performing measurable knowledge work tasks in randomized controlled trials and longitudinal playtests. These data will help to improve the human-AI experience of FCA for the benefit of KWs.

References

1. Abuhamdeh, S.: Investigating the “flow” experience: Key conceptual and operational issues. Frontiers in Psychology (2020)
2. Alves, N.T., Fukusima, S.S., Aznar-Casanova, J.A.: Models of brain asymmetry in emotional processing. Psychology & Neuroscience **1**, 63–66 (2008)
3. Amankwah-Amoah, J., Khan, Z., Wood, G., Knight, G.: Covid-19 and digitalization: The great acceleration. Journal of Business Research **136**, 602–611 (2021)
4. Barwick, F., Arnett, P., Slobounov, S.: Eeg correlates of fatigue during administration of a neuropsychological test battery. Clinical Neurophysiology **123**(2), 278–284 (2012)
5. Berka, C., Levendowski, D.J., Lumicao, M.N., Yau, A., Davis, G., Zivkovic, V.T., Olmstead, R.E., Tremoulet, P.D., Craven, P.L.: Eeg correlates of task engagement and mental workload in vigilance, learning, and memory tasks. Aviation, space, and environmental medicine **78**(5), B231–B244 (2007)
6. Boksem, M.A., Meijman, T.F., Lorist, M.M.: Effects of mental fatigue on attention: an erp study. Cognitive brain research **25**(1), 107–116 (2005)
7. Bostrom, N., Sandberg, A.: Cognitive enhancement: methods, ethics, regulatory challenges. Science and engineering ethics **15**(3), 311–341 (2009)
8. Cheng, S.Y., Hsu, H.T.: Mental fatigue measurement using EEG. IntechOpen (2011)
9. Colzato, L.S., Hommel, B., Beste, C.: The downsides of cognitive enhancement. The Neuroscientist **27**(4), 322–330 (2021)

10. Csikszentmihalyi, M.: Flow: The psychology of optimal performance.(1990) (1990)
11. Csikszentmihalyi, M.: Happiness and creativity. The Futurist **31**(5), S8 (1997)
12. Csikszentmihalyi, M.: Play and intrinsic rewards. In: Flow and the foundations of positive psychology, pp. 135–153. Springer (2014)
13. Csikszentmihalyi, M.: Toward a psychology of optimal experience. In: Flow and The Foundations of Positive Psychology, pp. 209–226. Springer (2014)
14. Dasari, D., Shou, G., Ding, L.: Ica-derived eeg correlates to mental fatigue, effort, and workload in a realistically simulated air traffic control task. Frontiers in neuroscience **11**, 297 (2017)
15. De Gennaro, L., Marzano, C., Veniero, D., Moroni, F., Fratello, F., Curcio, G., Ferrara, M., Ferlazzo, F., Novelli, L., Pellicciari, M.C., et al.: Neurophysiological correlates of sleepiness: a combined tms and eeg study. Neuroimage **36**(4), 1277–1287 (2007)
16. DeLosAngeles, D., Williams, G., Burston, J., Fitzgibbon, S.P., Lewis, T.W., Grummett, T.S., Clark, C.R., Pope, K.J., Willoughby, J.O.: Electroencephalographic correlates of states of concentrative meditation. International Journal of Psychophysiology **110**, 27–39 (2016)
17. Demos, J.N.: Getting Started with EEG Neurofeedback. Norton & Company (2019)
18. Dittner, A.J., Wessely, S.C., Brown, R.G.: The assessment of fatigue: a practical guide for clinicians and researchers. Journal of psychosomatic research **56**(2), 157–170 (2004)
19. Dubljević, V., Saigle, V., Racine, E.: The rising tide of tdcs in the media and academic literature. Neuron **82**(4), 731–736 (2014)
20. Eskridge, T.C., Weekes, T.R.: Opportunities for case-based reasoning in personal flow and productivity management. In: International Conference on Case-Based Reasoning. pp. 349–354. Springer (2020)
21. Farah, M.J.: The unknowns of cognitive enhancement. Science **350**(6259), 379–380 (2015)
22. Freeman, F.G., Mikulka, P.J., Prinzel, L.J., Scerbo, M.W.: Evaluation of an adaptive automation system using three eeg indices with a visual tracking task. Biological psychology **50**(1), 61–76 (1999)
23. Grammer, J.K., Xu, K., Lenartowicz, A.: Effects of context on the neural correlates of attention in a college classroom. npj Science of Learning **6**(1), 1–4 (2021)
24. Katahira, K., Yamazaki, Y., Yamaoka, C., Ozaki, H., Nakagawa, S., Nagata, N.: Eeg correlates of the flow state: A combination of increased frontal theta and moderate frontocentral alpha rhythm in the mental arithmetic task. Frontiers in Psychology **9**, 300 (2018)
25. Keller, B.M.: Cognitive avionics toolset for operator state classifacation based on physiological signals. In: 2007 IEEE/AIAA 26th Digital Avionics Systems Conference. pp. 6–A. IEEE (2007)
26. Kim, M.K., Kim, M., Oh, E., Kim, S.P.: A review on the computational methods for emotional state estimation from the human eeg. Computational and mathematical methods in medicine **2013** (2013)
27. Koelstra, S., Muhl, C., Soleymani, M., Lee, J., Yazdani, A., Ebrahimi, T., et al.: Ieee transactions on affective computing. DEAP: A database for emotion analysis. Using Physiol Signals **3**, 18–31 (2012)
28. Krigolson, O.E., Williams, C.C., Norton, A., Hassall, C.D., Colino, F.L.: Choosing muse: Validation of a low-cost, portable eeg system for erp research. Frontiers in neuroscience **11**, 109 (2017)

29. Langner, R., Steinborn, M.B., Chatterjee, A., Sturm, W., Willmes, K.: Mental fatigue and temporal preparation in simple reaction-time performance. Acta psychologica **133**(1), 64–72 (2010)
30. Lim, S., Yeo, M., Yoon, G.: Comparison between concentration and immersion based on eeg analysis. Sensors **19**(7), 1669 (2019)
31. Marzbani, H., Marateb, H.R., Mansourian, M.: Neurofeedback: a comprehensive review on system design, methodology and clinical applications. Basic and clinical neuroscience **7**(2), 143 (2016)
32. Matthews, G., Desmond, P.A.: Task-induced fatigue states and simulated driving performance. The Quarterly Journal of Experimental Psychology: Section A **55**(2), 659–686 (2002)
33. Milstein, N., Gordon, I.: Validating measures of electrodermal activity and heart rate variability derived from the empatica e4 utilized in research settings that involve interactive dyadic states. Frontiers in Behavioral Neuroscience **14** (2020)
34. Müller, O., Rotter, S.: Neurotechnology: Current developments and ethical issues. Frontiers in Systems Neuroscience **11**, 93 (2017)
35. Murray, A.J., Greenes, K.A.: From the knowledge worker to the knowledge economy: six billion minds co-creating the future. Vine (2007)
36. Nacke, L.: Affective ludology: Scientific measurement of user experience in interactive entertainment. Ph.D. thesis, Blekinge Institute of Technology (2009)
37. Nakamura, J., Csikszentmihalyi, M.: The concept of flow. In: Flow and the foundations of positive psychology, pp. 239–263. Springer (2014)
38. Newport, C.: Deep work: Rules for focused success in a distracted world. Hachette UK (2016)
39. Nielsen, J.: Enhancing the explanatory power of usability heuristics. In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems. pp. 152–158 (1994)
40. Norman, D.: The design of everyday things: Revised and expanded edition. Basic books (2013)
41. Ota, T., Toyoshima, R., Yamauchi, T.: Measurements by biphasic changes of the alpha band amplitude as indicators of arousal level. International journal of psychophysiology **24**(1-2), 25–37 (1996)
42. Plattner, H., Meinel, C., Leifer, L.: Design thinking research. Springer (2012)
43. Prinzel, L.J., Freeman, F.G., Scerbo, M.W., Mikulka, P.J., Pope, A.T.: A closed-loop system for examining psychophysiological measures for adaptive task allocation. The International journal of aviation psychology **10**(4), 393–410 (2000)
44. Putze, F., Jarvis, J.P., Schultz, T.: Multimodal recognition of cognitive workload for multitasking in the car. In: 2010 20th International Conference on Pattern Recognition. pp. 3748–3751. IEEE (2010)
45. Reuderink, B., Mühl, C., Poel, M.: Valence, arousal and dominance in the eeg during game play. International journal of autonomous and adaptive communications systems **6**(1), 45–62 (2013)
46. Sunstein, C.R.: Nudging: a very short guide. In: The Handbook of Privacy Studies, pp. 173–180. Amsterdam University Press (2018)
47. Swann, C., Piggott, D., Schweickle, M., Vella, S.A.: A review of scientific progress in flow in sport and exercise: normal science, crisis, and a progressive shift. Journal of Applied Sport Psychology **30**(3), 249–271 (2018)
48. Tang, Y.Y., Posner, M.I.: Attention training and attention state training. Trends in cognitive sciences **13**(5), 222–227 (2009)

49. Team, H.R.R.: Redesigning work for the hybrid world (Apr 2021), https://www.gartner.com/en/documents/4001104/redesigning-work-for-the-hybrid-world, accessed on 2020-11-03
50. Teplan, M., Krakovská, A., Špajdel, M.: Spectral eeg features of a short psycho-physiological relaxation. Measurement Science Review **14**(4), 237–242 (2014)
51. Thaler, R.H., Sunstein, C.R.: Nudge: Improving decisions about health, wealth, and happiness. Penguin (2009)
52. Tomarken, A.J., Davidson, R.J., Henriques, J.B.: Resting frontal brain asymmetry predicts affective responses to films. Journal of personality and social psychology **59**(4), 791 (1990)
53. Tripathi, S., Acharya, S., Sharma, R.D., Mittal, S., Bhattacharya, S.: Using deep and convolutional neural networks for accurate emotion classification on DEAP dataset. In: Twenty-Ninth IAAI Conference (2017)
54. Vaishya, R., Javaid, M., Khan, I.H., Haleem, A.: Artificial intelligence (ai) applications for covid-19 pandemic. Diabetes & Metabolic Syndrome: Clinical Research & Reviews **14**(4), 337–339 (2020)
55. Verganti, R., Vendraminelli, L., Iansiti, M.: Innovation and design in the age of artificial intelligence. Journal of Product Innovation Management **37**(3), 212–227 (2020)
56. Wilson, G.F., Russell, C.A.: Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. Human factors **45**(4), 635–644 (2003)
57. Zheng, W.L., Liu, W., Lu, Y., Lu, B.L., Cichocki, A.: Emotionmeter: A multimodal framework for recognizing human emotions. IEEE transactions on cybernetics **49**(3), 1110–1122 (2018)