2026 ASEE ANNUAL CONFERENCE June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte
Paper ID #51138

Cultivating Design Empathy through the PCAR Diary Framework

Jixiang Fan, Virginia Polytechnic Institute and State University
Jiahui Song, Northeastern University
Wei-Lu Wang, Virginia Polytechnic Institute and State University
Yuhao Chen, Georgia Institute of Technology
Scott McCrickard, Virginia Polytechnic Institute and State University

Cultivating Design Empathy through the PCAR Diary Framework

Abstract

Empathy refers to the ability to understand others' emotions, needs, and circumstances, and it serves as an essential foundation for recognizing user experiences, identifying latent requirements, and making informed design decisions. Recent studies have reported a noticeable decline in empathy among undergraduate students, particularly in highly technical fields such as nursing and engineering, which makes supporting user-centered design awareness an important instructional challenge. This study implemented a diary-based learning activity grounded in the PCAR (Plan, Collect, Analyze, Reflect) framework across seven groups of undergraduate computing students, with a focus group discussion incorporated into the Reflect phase to guide perspective integration and structured reflection. A total of 63 students participated, including six groups from a U.S. university (50 students) and one group from a Chinese university (13 students). Each group engaged with a different software application over two weeks, documented their experiences through diary entries, and later took part in a reflective discussion. The EMPA-D scale was used to examine changes in design empathy across three stages of the learning process, focusing on Emotional Interest and Perspective-taking, Personal Experience, and Self-awareness. Results showed variation in the magnitude of change across groups, while overall design empathy exhibited a consistent upward trend. The paper presents EMPA-D results across three measurement points and examines both overall patterns of change and cross-application consistency within the PCAR diary learning framework.

Keywords: Human-Computer Interaction; Empathy; Design; Assessment

Introduction

In technology-driven computing programs, students typically receive systematic training in areas such as algorithms, system architecture, software engineering processes, and programming techniques. As computing systems become increasingly embedded in social and engineering contexts, however, the quality of engineering design is shaped not only by technical performance but also by user experience, usage context, and understanding of user needs [1], [2]. This makes it essential for courses to help students learn to consider problems from the user's perspective. Yet for many novices, understanding real user experiences and needs remains challenging because they often lack methodological grounding and do not have sufficient experience linking abstract concepts to everyday technology use. Prior research has shown that students may encounter various difficulties when learning design-related content, including confusion about design concepts and uncertainty about how to use available resources, which collectively limit their ability to develop a deeper understanding of user problems [3] [1].

[1] AI tools were used solely for linguistic refinement and translation. The intellectual content, analysis, and conclusions

We adopt the PCAR (Plan, Collect, Analyze, Reflect) diary study framework [5], integrating continuous diary work with structured reflective activities to support computing students in developing design empathy through authentic software use and sustained reflection. In the Reflect phase, we incorporate a focus group discussion (FGD) [6] as a reflective strategy that encourages students to deepen their understanding and broaden their perspectives through peer comparison, perspective sharing, and reinterpretation of their cases. The study was conducted across seven groups of undergraduate computing students, with a total of 63 participants, including 50 students from a university in the United States and 13 students from a university in China. Each group engaged with a different software application for two weeks, completed diary entries documenting their experiences, and participated in a reflective discussion at the end of the activity.

The study aims to examine the implementation and scalability of this structured instructional mechanism and to investigate how students' design empathy changes over the course of the diary study and reflection process. Using data collected with the EMPA-D empathy scale [7] across three measurement points, we analyze overall trends in design empathy and examine whether similar patterns of change emerge across the seven learning groups. Rather than comparing groups to identify relative differences, the analysis focuses on the consistency of empathy development across different application contexts within the same instructional structure. The contributions of this work are twofold. First, the study provides a concrete, course-ready implementation of a PCAR-based diary study with facilitator-led group reflection, along with evidence of how students' design empathy changes over the activity. Second, by examining seven application contexts, the study offers evidence that this instructional mechanism yields consistent patterns of change across tools, suggesting it can be adopted without requiring a specific platform. Together, these contributions provide the ASEE community with empirical evidence and a practical instructional model for supporting user-centered thinking in computing education.

Background

Empathy in Engineering and Computing Education

In engineering and design education, empathy has been discussed as an important capability for understanding users' situations and perspectives. Kouprie and Visser defined empathy in design as a process of entering into and stepping out of users' lived situations, emphasizing the formation of user understanding through experience and reflection in the design process [8]. In engineering education contexts, Hess et al. discussed how empathy for users can be measured in engineering design settings and characterized it as a user-oriented cognitive and affective disposition [9]. Kim et al. further examined how perspective-mixing within design teams supports shared empathic understanding [10]. Fila et al. proposed an integrated

are the authors' own.

framework of user empathy for engineering students, outlining different dimensions of empathy in engineering design learning [11]. These studies indicate that engineering education research has engaged in systematic discussion of user-oriented empathy and has sought to define its dimensions from multiple perspectives.

Debate continues regarding whether empathy can be supported through instruction. Hess et al. examined the development of empathic perspective-taking in an engineering ethics course [12]. James et al. discussed instructional practices that use design challenges to guide students' attention to user situations [13]. Mitchell and Light reported on the introduction of immersive user experience activities in first-year design education and analyzed students' engagement in these activities [14]. Shah et al. investigated the relationship between a targeted assignment in an engineering design course and students' empathy measures, addressing whether empathy can be supported through course design [15]. Overall, these studies explore the relationship between instructional practice and student empathy, while leaving open questions regarding how different instructional structures may systematically support empathy development.

Although empathy has been examined extensively in engineering education, much of this work has been situated in traditional engineering design course contexts. In these settings, students typically collaborate on tangible products or clearly defined design problems, where user situations are relatively concrete. In contrast, computing education emphasizes algorithm development, system implementation, and technical optimization. Course content tends to be more abstract, and students' interaction with real users is often indirect. Accordingly, in computing education contexts, how specific instructional structures can guide students to understand user situations, and how changes in empathy may be observed in this process, remain open questions. The present study is situated within this context.

Diary Study in Education

Diary studies are a qualitative method commonly used in HCI to examine users' experiences, behaviors, and reflections across an extended period of time [16]. In practice, participants are asked to keep records at intervals or in response to prompts defined by the researcher. These records typically describe how they interact with particular software or technologies, the scenarios and processes involved, and the thoughts or emotions that arise during use. The format of diary studies can vary: some encourage open-ended, free writing that allows participants to describe their experiences in detail. In contrast, others rely on structured prompts or scales that facilitate subsequent qualitative or quantitative analysis [17]. This flexibility allows diary studies to produce both detailed personal narratives and standardized data, enabling systematic comparison. First-hand accounts of this kind provide researchers with a clear view of how people's thoughts, feelings, and practices evolve over time in real-world contexts. When set against one-time interviews or laboratory studies, diary methods are often more effective in capturing change and subtle aspects of experience, making it possible to reveal user needs and challenges that other approaches might overlook [18].

Diary studies have increasingly been introduced into educational discussions related to design-oriented computing, particularly within human-computer interaction (HCI) education. This trend is supported by both the ongoing development of diary methods themselves and the emergence of technologies that facilitate reflection on diary content, making diary studies more adaptable to educational practice. For example, timeline features in diary systems allow individuals to revisit and gain insights into the rhythms and patterns of their daily activities [19]. Li et al. combined multimodal data processing with the contextual prediction capabilities of large language models to design DiaryHelper, an automated agent for recording contextual information [20]. Similarly, Oh et al. applied fine-tuned large language models to summarize diary content, helping users engage in more effective reflection on their records [21]. Fan et al. highlighted the potential value of diary studies for HCI teaching, particularly in complex outdoor environments [22]. Factors such as weather and lighting can significantly influence user experience and behavioral patterns; diary studies enable students to gain a more direct understanding of these effects. To further advance the use of diary studies as an HCI teaching tool, researchers proposed the PCAR framework: Plan, Collect, Analyze, and Reflect, which provides detailed elements and recommendations for each stage [23]. The authors emphasized the importance of the Reflect stage as the key point at which students revisit and consolidate their diary content, allowing them to identify underlying user needs and potential human factors. They also suggested that educators adapt reflection strategies to align with specific learning objectives, class sizes, and available resources.

Focus Group Discussion

Focus Group Discussions (FGDs) are a qualitative method widely used to explore participants' perspectives, attitudes, and experiences through group interaction [24]. Unlike individual interviews, FGDs create a social context in which participants can respond to one another, compare viewpoints, and generate new ideas. This interactive dynamic makes FGDs particularly valuable for revealing collective patterns, uncovering diverse perspectives, and identifying issues that may not emerge in one-on-one settings [25]. However, prior research has also noted that FGDs are not fully insulated from risks such as groupthink, which may shape or suppress individual perspectives. Boateng [26] argues that while FGDs can serve as a useful stand-alone method for non-sensitive topics, they are best employed in combination with other approaches, for example, as part of triangulation or mixed-method designs, to ensure richer and more reliable insights. In HCI studies, FGDs have been applied to evaluate technologies, probe user needs, and investigate the social dimensions of technology use [27].

FGDs have also been used as a way to support reflective and collaborative learning [28], [29]. Students can put their own ideas into words, listen to how peers see the same issue, and rethink their positions through conversation. Prior work shows that peer discussion of this kind can encourage stronger engagement, more critical thinking, and greater awareness of different viewpoints [30], [31]. Building on these insights, we position FGDs as a reflection strategy within the Reflect stage of the PCAR framework for diary-based learning. In our approach, students first record their experiences individually in diaries and then revisit those reflections in group discussions. This two-step process helps them retain the depth of personal reflection while also benefiting from the wider perspectives of collective exchange. In turn, it supports a fuller understanding of user experience and user needs in HCI education.

Method

This study adopts the PCAR (Plan, Collect, Analyze, Reflect) diary study framework to guide both instruction and research. The framework was proposed by Fan et al. in 2025 [23] and was developed through a systematic review of prior diary study implementations in research and education, together with the authors' teaching experience in HCI-related courses. PCAR is intended to support the effective integration of diary studies into HCI and design-oriented computing education. The framework organizes diary study activities into four interrelated stages. The Plan stage focuses on defining learning goals, selecting study targets, and organizing activities. The Collect stage guides students to document their experiences continuously in authentic usage contexts. The Analyze stage supports students in reviewing and organizing their diary entries to develop preliminary interpretations. The Reflect stage emphasizes structured reflection activities that help students integrate experiences and deepen their understanding of user needs. In this study, we designed and implemented instructional activities aligned with each stage of the PCAR framework.

Data Collection

Plan Phase

A total of 63 undergraduate students majoring in computing participated in this study. 50 students were enrolled at a university in the United States, and 13 students were enrolled at a university in China. All participants were taking computing-related courses and had basic experience with software use and technical systems. However, prior to the study, they had not received systematic training in user-centered design or empathy-oriented design approaches.

During the Plan phase, the research team selected representative interactive software applications for each learning group based on course objectives and instructional context. Six learning groups at the U.S. university used Merlin Bird ID, Monster Hunter Now, Platemate, Pokémon GO, ReciMe, and Relive, respectively, while the learning group at the Chinese university used Keep. The selection of these applications was guided by three considerations. First, all applications support sustained use in everyday contexts, enabling repeated authentic interaction throughout the two-week diary period. Second, the selected tools represent diverse usage domains and interaction patterns, including outdoor activity and location-based systems (Relive, Pokémon GO, Monster Hunter Now), health and fitness management (Keep), knowledge logging and identification support (Merlin Bird ID), recipe and cooking assistance (ReciMe), and food-related social interaction (PlateMate). This diversity ensured variation in interaction modes, system feedback mechanisms, and intended user groups. Third, all applications are consumer-facing systems designed for general public use rather than specialized professional tools, allowing undergraduate students to participate as authentic users and examine design issues from a user-oriented perspective. For students without prior exposure to user-oriented design, such applications offer not only diverse interaction experiences but also multiple entry points for observing user experience and user needs. This allowed students to analyze design issues from different perspectives, including feature use, contextual variation, and user behavior.

Prior to the implementation of the diary study, the research team conducted preliminary use and analysis of all selected applications to ensure alignments. Based on this preparatory work, a corresponding diary template was designed for each application. The templates combined quantitative items, reflection prompts, and open-ended questions to guide students in documenting their usage processes, emotional responses, encountered challenges, interpretations of design features, and observations related to potential user needs. By employing structured templates, the study aimed to support students' expressive freedom while improving data comparability and analytical consistency across learning groups.

Collect Phase

During the Collect phase, students engaged in a two-week diary study centered on their assigned software applications. The study period excluded weekends, during which students were required to complete a total of ten diary entries. Each entry was based on students' actual use of the software on that day, and participants were instructed to complete their entries as soon as possible after use to minimize recall bias and maintain the authenticity of their reflections. This continuous recording approach was intended to capture students' evolving perceptions of software functionality, interaction processes, and user experience across different usage contexts.

Throughout the diary study, students completed their entries using the predefined diary templates. These templates guided students to systematically reflect on usage context, interaction processes, emotional responses, encountered challenges, and interpretations of design choices. At the same time, open-ended prompts were included to encourage students to freely articulate their observations and reflections. Students were explicitly instructed not to retrospectively complete multiple entries at the end of the study period, but rather to treat diary writing as an integral part of their day-to-day software use.

Analyze Phase

After the two-week diary study was completed, the Analyze phase was conducted by the facilitators who later led the focus group discussions. Each learning group was supported by one or two facilitators with prior experience in HCI research and qualitative inquiry. The facilitators first organized and reviewed the diary entries submitted by students in their respective groups to gain an overview of participants' experiences.

During this phase, facilitators examined the diary entries to identify recurring experiences, salient moments, and issues relevant to software use, user experience, and perceived user needs. Rather than performing formal coding for result generation, the analysis focused on surfacing themes and contrasts that could support meaningful group reflection. Based on this review, facilitators prepared discussion materials, including representative diary excerpts that reflected a range of perspectives and experiences within each group.

These materials were then used to structure the subsequent focus group discussions. By grounding the Reflect phase in patterns and examples drawn directly from students' diary entries, the Analyze phase helped ensure that group reflection was anchored in participants' lived experiences and supported deeper discussion of user experience, system behavior, and design trade-offs.

Reflect Phase

During the Reflect phase, students engaged in facilitator-led focus group discussions to collectively reflect on their diary study experiences. Each learning group was guided by one or two facilitators with prior experience in HCI research. The discussions were grounded in the diary entries and representative excerpts identified in the previous phase and were designed to help students connect individual experiences with group-level perspectives.

To keep individual viewpoints visible and minimize the influence of peer conformity, each participant was first asked to write down their thoughts on their usage experiences, perceived user needs, and design issues before the group discussion began. During the discussion, facilitators intentionally balanced speaking turns and time to ensure that all students had opportunities to contribute, while also taking notes on key remarks. When important design issues emerged or when further clarification was needed, facilitators followed up with probing questions to encourage deeper reflection. Through this process, students were able to reflect individually through diary writing and then engage in group dialogue, comparing their ideas and learning from others' experiences. This movement from individual reflection to collective exchange represents a core component of the instructional framework and formed the basis for subsequent analyses.

When a high level of agreement emerged among participants, facilitators deliberately introduced alternative or opposing perspectives to prompt deeper discussion. For example, facilitators might raise hypothetical counterexamples, encourage students to consider different user groups or usage contexts, or question the applicability of a design choice under alternative constraints. By challenging surface-level consensus, these interventions encouraged students to reconsider their assumptions and engage more deeply with diverse user perspectives, thereby cultivating empathy-oriented thinking in design decision-making.

Data Analysis

To characterize changes in students' design empathy within the PCAR diary study framework in computing education, this study employed the Empathy in Design Scale (EMPA-D) [7] as the primary quantitative measurement instrument. Empathy is widely recognized as a core capability in user-centered design, yet measuring empathy in design and engineering education contexts remains challenging. Existing instruments include Davis's Interpersonal Reactivity Index (IRI) [32], which primarily assesses general interpersonal trait empathy, and engineering-focused measures such as the scale proposed by Hess et al. [9], which emphasizes perspective-taking within specific design tasks. Because this study examines changes in user-oriented empathy within a structured design learning process, a measure that directly assesses empathy toward users in design contexts is appropriate.

EMPA-D is a recently developed and empirically validated 7-point scale designed to assess individuals' empathy toward users in design and service contexts, enabling abstract aspects of design empathy to be operationalized as measurable and comparable indicators. The EMPA-D scale was developed following established best practices in scale construction and has demonstrated strong reliability and validity across large-scale quantitative studies conducted in multiple industry contexts. The scale consists of 11 items organized into a three-factor structure, measuring Emotional Interest and Perspective-taking, Personal Experience, and Self-awareness. This structure allows EMPA-D to capture design empathy across cognitive, affective, and reflective dimensions, making it well-suited for assessing changes in empathy during design-oriented learning activities. In this study, EMPA-D data were collected at three time points: before the learning activities began (Pre), after completion of the two-week diary study (Mid), and after the Reflect phase that included facilitator-led focus group discussion (Post). This repeated-measures design enables examination of stage-specific changes in design empathy and supports analysis of overall trends across the learning process.

In addition to standard measures of central tendency and dispersion, several distributional indicators are reported to support a more detailed characterization of score distributions at each measurement stage. The interquartile range (IQR) captures the spread of the middle 50% of scores and provides a variability measure that is less sensitive to extreme values than the standard deviation. The coefficient of variation (CV), calculated as the ratio of the standard deviation to the mean, expresses dispersion relative to the central tendency, which is particularly relevant when comparing variability across time points where the mean itself shifts. Skewness and kurtosis are reported to describe the shape of each distribution: skewness reflects the degree of asymmetry, with negative values indicating a concentration of scores toward the upper end of the scale, while kurtosis captures the relative peakedness and tail weight compared to a normal reference distribution.

Paired-sample t-tests were used to evaluate within-subject changes across measurement stages, and one-way repeated-measures ANOVA was conducted for each EMPA-D subdimension to examine construct-level growth patterns. For the cross-application comparison, linear mixed models (LMM) [33] were employed to accommodate unbalanced group sizes and varying correlation structures, with likelihood ratio tests (LRT) [34] used to evaluate the Time × App interaction. Cohen's d is reported alongside significance tests to convey practical magnitude, calculated as the mean difference divided by the pooled standard deviation. Values of 0.2, 0.5, and 0.8 correspond to small, medium, and large effects, respectively, and the interpretation labels in the results tables (e.g., "Medium Effect," "Very Large Effect") are derived from these thresholds. Gain values refer to the arithmetic difference in mean scores between two stages, with positive values indicating improvement.

Results

Based on the measurement design described above, the Results section is organized into two main analyses. First, the **Overall Effectiveness Analysis** examines overall changes across all learning groups at the three measurement time points to describe aggregate changes in design empathy across the study period. Second, **App Comparison: Universal Efficacy** focuses on changes in students' design empathy across different software contexts by analyzing empathy growth curves associated with each application, in order to evaluate the consistency of the instructional mechanism across multiple application settings. Through these analyses, we evaluate the consistency of the instructional mechanism across varied software contexts.

Overall Effectiveness Analysis

Descriptive Statistics

Table 1 summarizes the descriptive statistics of overall empathy scores across the three time points. The results show a steady and monotonic increase in mean empathy scores, rising from 5.462 at the Pre stage to 5.827 at the Mid stage and further to 6.108 at the Post stage. This pattern indicates that as the diary study and subsequent reflection activities progressed, students demonstrated progressively higher levels of design empathy at the group level.

Time Point	Mean	SD	Median	IQR	Min	Max	Skewness	Kurtosis	CV
Pre	5.462	0.543	5.444	0.750	3.889	6.389	-0.592	0.347	0.099
Mid	5.827	0.493	5.889	0.583	4.444	6.889	-0.324	0.181	0.085

Post	6.108	0.467	6.056	0.639	5.000	7.000	0.114	-0.299	0.076

Table 1. Descriptive statistics of overall EMPA-D empathy scores across three measurement time points

In addition to the increase in mean scores, the distribution of empathy scores became more concentrated over time. The standard deviation decreased from 0.543 at the Pre stage to 0.467 at the Post stage, and the coefficient of variation declined from 0.099 to 0.076. These changes suggest that students' empathy scores not only increased overall but also became more tightly clustered around a higher mean. Examination of distribution shape further supports this trend. The score distribution shifted from a negatively skewed form at the Pre stage (skewness = −0.592) to an approximately symmetrical distribution at the Post stage (skewness = 0.114). This shift indicates that students who initially reported lower levels of empathy showed notable improvement over time, reducing disparities between lower and higher-scoring students and resulting in a more balanced, high-performing cohort.

Boxplot visualizations (See Figure 1) provide further insight into the distributional changes across the three stages. The median line within each box shifts stepwise from the Pre to the Post stage, consistent with the observed increase in central tendency. The boxes representing the interquartile range also move upward as a whole, and by the Post stage, the lower quartile is often comparable to or higher than the median at the Pre stage. This pattern confirms that the observed gains reflect a population-wide shift rather than improvements driven by a small subset of students.

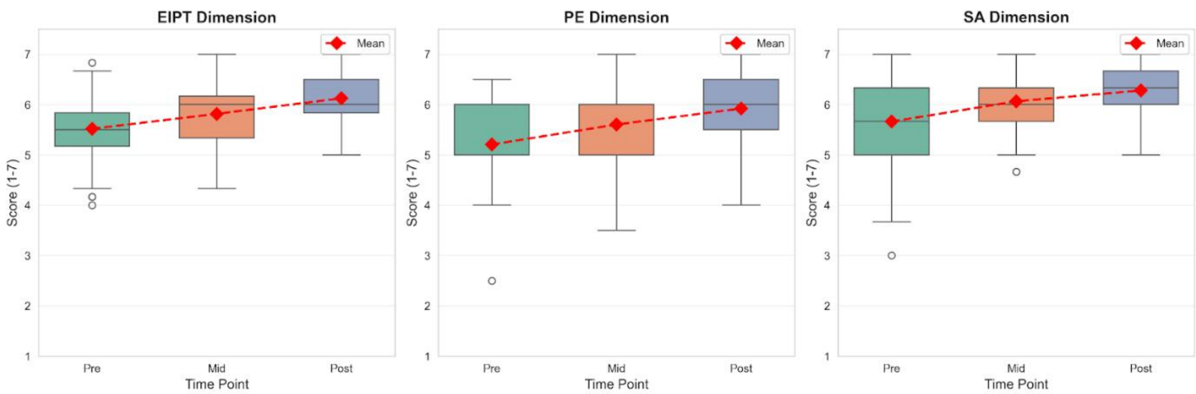

Figure 1: Distribution of empathy scores across Pre, Mid, and Post stages. The left panel shows the Emotional Interest and Perspective-taking (EIPT) dimension, the middle panel shows the Personal Experience (PE) dimension, and the right panel shows the Self-awareness (SA) dimension. Red dashed lines indicate mean scores across time points.

Significance Testing Across Time Points

To evaluate whether changes in empathy scores were statistically significant, paired-sample t tests were conducted between the three measurement stages. All comparisons revealed highly significant differences ($p < 0.0001$).

Comparison	Mean Diff.	t-value	p-value	Cohen's d	Interpretation
Pre vs Mid	+0.365	5.318	< 0.0001	0.704	Medium-Large Effect
Mid vs Post	+0.281	5.028	< 0.0001	0.586	Medium Effect
Pre vs Post	+0.646	8.761	< 0.0001	1.276	Very Large Effect

Table 2. Paired-sample t-test results for overall EMPA-D scores across three measurement stages (Pre, Mid, Post)

From the Pre to the Mid stage, empathy scores increased by an average of 0.365 ($t = 5.318$), corresponding to a medium-to-large effect size (Cohen's $d = 0.704$). This result suggests that students' empathy exhibited measurable increases even in the early phase of the diary study activity. From the Mid to the Post stage, scores increased by an additional 0.281 ($t = 5.028$, $d = 0.586$), indicating that empathy continued to develop during the Reflect phase rather than reaching an early plateau. Overall, the difference between the Pre and Post stages reached 0.646, with a very large t value ($t = 8.761$) and a substantial effect size ($d = 1.276$). This effect size exceeds the conventional threshold for a large effect, indicating that the observed changes in students' design empathy were large in magnitude and practically meaningful.

App Comparison: Universal Efficacy

Descriptive Changes Across Application Groups

Table 3 summarizes the mean empathy scores and overall gains for the seven application groups across the three measurement time points. The results show that all application groups exhibited positive gains over the two-week study period, with total increases ranging from +0.439 to +0.806. Although the magnitude of growth varied across groups, all groups demonstrated a consistent upward trend from the Pre to the Post stage, with no application group standing out as clearly superior or inferior to the others.

App	Pre Mean (SD)	Mid Mean (SD)	Post Mean (SD)	Total Gain
PlateMate	5.556 (0.486)	5.926 (0.565)	6.361 (0.349)	**+0.806**
Keep	5.363 (0.478)	5.803 (0.446)	6.128 (0.396)	+0.765
ReciMe	5.222 (0.281)	5.642 (0.430)	5.932 (0.459)	+0.710
Relive	5.478 (0.660)	5.711 (0.769)	6.144 (0.461)	+0.667
Merlin Bird ID	5.333 (0.643)	5.778 (0.386)	5.932 (0.413)	+0.599
Monster Hunter Now	5.545 (0.551)	5.939 (0.632)	6.131 (0.564)	+0.586
Pokémon Go	5.767 (0.651)	5.944 (0.383)	6.206 (0.567)	+0.439

Table 3. Mean EMPA-D scores and overall gains across application-based learning groups

A closer examination of the descriptive statistics reveals that some applications showed relatively larger raw gains on specific subdimensions. For example, the PlateMate group demonstrated a notable increase in the Personal Experience dimension, while the group from the Keep showed comparatively stronger growth in the Self-awareness dimension. However, these differences are primarily descriptive in nature and do not form stable or consistent patterns of advantage across applications. At the same time, some application groups, such as Pokémon Go, exhibited smaller overall gains.

Tool Independence Analysis

To examine whether a tool effect was present, that is, whether certain applications were more effective than others in promoting design empathy, we conducted statistical tests on the growth trajectories across application groups. The central question guiding this analysis was whether changes in students' empathy over time depended on the specific application used, or whether similar growth patterns emerged across different application contexts. We focused on the Time × App interaction effect. In the context of this study, Time represents the measurement stages across the learning process, while App represents the different software applications. The Time × App interaction tests whether the rate of empathy growth varies by application. A significant interaction would indicate that empathy growth trajectories differ across application groups, with some applications producing faster or slower gains. In contrast, a non-significant interaction indicates that empathy growth rates are statistically consistent across applications.

Table 4 summarizes the growth patterns across the seven application groups for the three EMPA-D subdimensions, along with the results of the global interaction tests. Across Emotional Interest and Perspective-taking, Personal Experience, and Self-awareness, the LRT p-values for the Time × App interaction were all well above the significance threshold, specifically $p = 0.999$, $p = 0.232$, and $p = 0.985$. These results indicate that differences in empathy growth trajectories across application groups were not statistically significant.

App Group	EIPT Gain	PE Gain	SA Gain
Keep (China Group)	+0.59	+0.81	+0.90
Merlin Bird ID	+0.59	+0.61	+0.59
Monster Hunter Now	+0.59	+0.59	+0.58
PlateMate	+0.61	+1.25	+0.56
Pokémon Go	+0.72	+0.20	+0.40
ReciMe	+0.57	+1.00	+0.56
Relive	+0.53	+0.80	+0.67
LRT p-value	**p = 0.999**	**p = 0.232**	**p = 0.985**
Result	**No Diff**	**No Diff**	**No Diff**

Table 4. Empathy growth across EMPA-D subdimensions and global Time × App interaction test results

Although descriptive results showed that some applications exhibited larger raw gains on specific subdimensions, such as the relatively strong increase in Personal Experience observed in the PlateMate group, statistical testing indicated that these differences fell within the range of random variation and did not constitute systematic application-level advantages. Such patterns may be attributable to smaller group sizes or differences in baseline scores, rather than to the inherent effectiveness of particular applications. The results provide no evidence of a significant tool effect. Regardless of the application around which the diary study was conducted, students demonstrated similar rates of growth in design empathy.

Discussion

Empathy as a Scaffolded Development: Interpreting Changes in Design Contexts

Whether empathy can be systematically supported in educational settings has long been a topic of debate [15], [35]. In computer science education, which is often centered on the development of technical competencies, course objectives commonly emphasize algorithms, systems, and programming skills, while aspects related to understanding users, such as empathy, may receive less systematic attention within regular coursework. Although prior research commonly positions empathy as an important component of design competence [4], [36], its development in educational contexts is often associated with long-term experience accumulation or highly contextualized practice, rather than with learning processes that can be readily observed within regular coursework. Against this backdrop, how empathy can be observed, measured, and understood within everyday instructional settings remains an open question in engineering and computing education.

This study did not include a control group and therefore does not allow for strict causal claims regarding the relationship between instructional activities and changes in empathy. Nevertheless, data collected using the EMPA-D scale at three measurement time points reveal a clear and consistent pattern of change. Across the learning process, students' design empathy exhibited a steady upward trend at the cohort level. This pattern was not driven by a small number of students or a single learning group but appeared consistently across groups. These results suggest that, within the course context examined in this study, students' empathy levels were not static but showed measurable variation over time.

A closer examination of the temporal distribution of these changes further highlights a structured developmental pattern. Rather than showing a sudden increase only at the end of the course, empathy scores rose gradually across multiple time points, reflecting a progressive trajectory. This observation suggests that the development of design empathy may not take the form of a one-time cognitive shift but instead unfolds through sustained engagement with usage contexts and repeated opportunities for reflection. Although the present study cannot attribute these changes to any specific instructional component, the temporal pattern of empathy growth aligns with the sustained presence of reflective activities throughout the learning process.

In this sense, the findings of this study are not intended to demonstrate that a particular instructional intervention “increased” students’ empathy. Instead, they provide empirical evidence for the variability of design empathy within computer science education contexts. The results indicate that in learning environments that include ongoing experience-based activities and structured opportunities for reflection, students’ design empathy can be consistently observed, measured, and tracked over time. This observation offers an empirical starting point for future research to further examine how empathy may be supported, shaped, and assessed in engineering and computing education.

From Individual Reflection to Collective Construction: A Structure-Driven Learning Mechanism

This study integrates diary study and focus group discussion to construct a learning mechanism that supports students’ understanding of user experience and user needs through a progression from individual reflection to collective construction. From an instructional design perspective, diary study and focus group discussion serve distinct yet interconnected roles within this learning process. Diary study provides students with a reflection structure grounded in personal experience, enabling them to continuously document their interactions, emotional responses, operational challenges, and design perceptions in real usage contexts. Through this process, abstract design concepts are translated into concrete experiences that can be examined and reflected upon, allowing students to approach technical systems from the standpoint of a user.

At the same time, reflection based solely on individual experience has inherent limitations. When documenting their own use, students may implicitly treat their personal context as representative, overlooking the diversity of user situations and perspectives. To address this limitation, the study introduced facilitator-led focus group discussions during the Reflect phase, embedding individual reflections within a collective exchange. In these discussions, students brought their diary entries into a shared conversational space, where their interpretations were compared, questioned, and refined through interaction with peers. Facilitators played an active role by managing the flow of discussion, balancing participation, and posing guiding questions that directed attention toward overlooked usage scenarios and alternative user perspectives.

Viewed as an integrated process, diary study and focus group discussion together form a progressive reflection mechanism. Repeated diary entries allow students to revisit and adjust their understanding over time, while subsequent group discussions provide opportunities for contrast and reinterpretation of these individual reflections. This ongoing movement between personal reflection and collective dialogue helps students recognize the partial nature of their own experiences and supports the gradual expansion of their perspective toward broader user contexts.

This reflection-centered learning mechanism does not rely on the functional characteristics of specific tools or applications. Regardless of which application students examined through diary study, the learning process followed a similar reflective trajectory, beginning with individual experience and extending through structured recording and collective discussion. This consistency helps explain the tool-independence observed in the results: learning outcomes appear to be shaped more by how reflection activities are organized than by differences among the applications themselves. In this sense, the mechanism that links individual reflection with collective construction provides a coherent structure for supporting students' understanding of user experience across diverse technical contexts.

Conclusion & Future Work

This study examined changes in design empathy among undergraduate computer science students through a PCAR (Plan–Collect–Analyze–Reflect) diary study, with focus group discussions incorporated during the Reflection phase. Using the EMPA-D scale across three measurement points, the results revealed a stable pattern of change in students' design empathy over the course of the learning activity, with this pattern remaining consistent across different application contexts. These findings suggest that organizing learning activities around authentic software use and structured reflection is associated with observable changes in students' design empathy over time.

From an instructional design perspective, the results further indicate that learning outcomes were not strongly dependent on the functional characteristics of specific applications but were more closely associated with the learning mechanism formed by diary study and structured reflection activities. By combining sustained documentation of real usage experiences with collective discussion during the Reflection phase, the PCAR framework provides students with a learning pathway that begins with individual experience and gradually expands toward a broader understanding of users. This highlights the role of instructional structure in supporting user-centered thinking within engineering and computing education.

This study adopted a design within an authentic course context. Accordingly, the findings are intended to characterize patterns of change in students' empathy within a specific instructional structure, rather than to establish causal comparisons between different pedagogical interventions. Future work may build on this foundation by introducing additional instructional conditions or reflection strategies to further explore how diary studies function across diverse engineering and computing education settings.

Reference

[1] R. W. Veryzer and B. Borja de Mozota, *"The impact of user-oriented design on new product development: An examination of fundamental relationships",* Journal of Product

Innovation Management, vol. 22, no. 2, pp. 128–143, 2005.

[2] M. Hassenzahl, *Experience Design: Technology for All the Right Reasons*, vol. 8. San Rafael, CA: Morgan & Claypool Publishers, 2010.

[3] A. Oleson, M. Solomon, and A. J. Ko, *"Computing students' learning difficulties in HCI education",* in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020.

[4] E. Schmitt et al., "The importance of incorporating designer empathy in senior capstone design courses", in 2016 ASEE Annual Conference & Exposition, 2016.

[5] J. Fan and S. McCrickard, "Diary studies as a pedagogical tool for human-computer interaction education", in Proceedings of the 2025 ACM Conference on International Computing Education Research, vol. 2, 2025.

[6] M. M. Hennink, Focus Group Discussions. Oxford: Oxford University Press, 2013.

[7] L. Drouet, K. Bongard-Blanchy, and C. Lallemand, "Development of the Empathy in Design Scale: Measuring employees' empathy toward users in service design", Interacting with Computers, vol. 36, no. 5, pp. 313–334, 2024.

[8] M. Kouprie and F. S. Visser, "A framework for empathy in design: Stepping into and out of the user's life", Journal of Engineering Design, vol. 20, no. 5, pp. 437–448, 2009.

[9] J. L. Hess, N. D. Fila, E. Kim, and S. Purzer, "Measuring empathy for users in engineering design", International Journal of Engineering Education, vol. 37, no. 3, pp. 733–743, 2021.

[10] E. Kim, R. S. Adams, and S. Purzer, "A design team's perspective-mixing for shared empathic understandings", CoDesign, vol. 20, no. 4, pp. 621–644, 2024.

[11] N. Fila, J. Hess, E. Sanders, and C. Schimpf, "Toward an integrated framework of empathy for users among engineering student designers", in ASEE Annual Conference Exposition, 2024.

[12] J. L. Hess, J. Strobel, and A. O. Brightman, "The development of empathic perspective-taking in an engineering ethics course", Journal of Engineering Education, vol. 106, no. 4, pp. 534–563, 2017.

[13] J. James, V. Svihla, C. Qiu, and C. Riley, "Using design challenges to develop empathy in first-year courses", in Proceedings of the ASEE 125th Annual Conference and Exhibition, 2018.

[14] L. Mitchell and L. Light, "Increasing student empathy through immersive user empathy experiences in first-year design education", in 2018 ASEE Annual Conference & Exposition

[15] D. Shah, B. Morkos, and X. Yang, "Can empathy be taught? The results of an assignment targeted at improving empathy in engineering design", in 2020 ASEE Virtual Annual Conference Content Access, 2020.

[16] S. Carter and J. Mankoff, "When participants do the capturing: The role of media in

diary studies", in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2005.

[17] N. Bolger, A. Davis, and E. Rafaeli, "Diary methods: Capturing life as it is lived", Annual Review of Psychology, vol. 54, no. 1, pp. 579–616, 2003..

[18] D. Olorunfemi, "Diary studies in research: More than a research method", International Journal of Market Research, vol. 66, no. 4, pp. 410–427, 2024.

[19] R. F. Belli, F. P. Stafford, and D. F. Alwin, Calendar and Time Diary Methods in Life Course Research. Sage, 2009.

[20] J. Li, C. He, J. Hu, B. Jia, A. Y. Halevy, and X. Ma, "DiaryHelper: Exploring the use of an automatic contextual information recording agent for elicitation diary study", in Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, pp. 1–16, 2024.

[21] S. Oh, J. Zhao, C. Russo, and M. Bolmer Jr., "Boosting diary study outcomes with a fine-tuned large language model", in Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, pp. 1–7, 2025.

[22] J. Fan, M. Saaty, and D. S. McCrickard, "Education in HCI outdoors: A diary study approach", in Proceedings of the 6th Annual Symposium on HCI Education, pp. 1–10, 2024.

[23] J. Fan, D. Haqq, M. Saaty, W.-L. Wang, and S. McCrickard, "Diary study as an educational tool: An experience report from an HCI course", in Proceedings of the 56th ACM Technical Symposium on Computer Science Education, vol. 1, pp. 311–317, 2025.

[24] S. Rosenbaum, G. Cockton, K. Coyne, M. Muller, and T. Rauch, "Focus groups in HCI: Wealth of information or waste of resources?", in CHI '02 Extended Abstracts on Human Factors in Computing Systems, pp. 702–703, 2002.

[25] J. L. Ridgeway, M. Albertie, E. Pantoja, D. Prescott, X. Zhu, and C. R. Breitkopf, "Understanding diverse perspectives on genetic research through focus group talk", International Journal of Qualitative Methods, vol. 18, p. 1609406919892476, 2019.

[26] W. Boateng, "Evaluating the efficacy of focus group discussion (FGD) in qualitative social research", International Journal of Business and Social Science, vol. 3, no. 7, 2012.

[27] P. J. O'Donnell, G. Scobie, and I. Baxter, "The use of focus groups as an evaluation technique in HCI", in People and Computers VI, pp. 211–224, 1991.

[28] E. Hastuti, "The effectiveness of focus group discussions (collaborative learning techniques) to teach English as viewed from the students' motivation", in Conference on English Language Teaching, pp. 259–268, 2024.

[29] H. B. Memduhoglu, N. Kotluk, and A. Yayla, "The effect of focus group discussions on pre-service teachers' teaching experiences and practices: A mixed methods study", International Journal of Instruction, vol. 10, no. 4, pp. 273–292, 2017.

[30] T. Anderson and R. Soden, "Peer interaction and the learning of critical thinking skills",

Psychology Learning & Teaching, vol. 1, no. 1, pp. 37–40, 2001.

[31] D. C.-L. Fung and T. W. Liang, Fostering critical thinking through collaborative group work. Springer, 2019.

[32] M. H. Davis, “Measuring individual differences in empathy: Evidence for a multidimensional approach”, Journal of Personality and Social Psychology, vol. 44, no. 1, p. 113, 1983.

[33] B. M. Bolker, “Linear and generalized linear mixed models”, in Ecological Statistics: Contemporary Theory and Application, pp. 309–333, 2015.

[34] Q. H. Vuong, “Likelihood ratio tests for model selection and non-nested hypotheses”, Econometrica, pp. 307–333, 1989.

[35] M. S. W. Cuzzo, M. R. Larson, L. M. Mattsson, and T. D. McGlasson, “How do you effectively teach empathy to students?”, New Directions for Teaching and Learning, vol. 2017, no. 151, pp. 61–78, 2017.

[36] X. Li, J. Chen, and H. Fu, “The roles of empathy and motivation in creativity in design thinking”, International Journal of Technology and Design Education, vol. 34, no. 4, pp. 1305–1324, 2024.