

Measuring and Promoting Empathic Formation in a Multidisciplinary Engineering Design Course

Justin L Hess (Assistant Professor)

Dr. Justin L Hess is an assistant professor in the School of Engineering Education at Purdue University. His vision is to inspire change in engineering culture to become more socially responsive, environmentally friendly, and inclusive, thereby providing opportunities for all current and prospective engineers to reach their maximum potential. Dr. Hess's research focuses on empathy, equity, and ethics in engineering education. He received his PhD from Purdue University's School of Engineering Education, as well as a Master of Science and Bachelor of Science from Purdue University's School of Civil Engineering. He is the 2022 division chair for the ASEE Liberal Education/Engineering and Society division; deputy director of research for the National Institute of Engineering Ethics; and the editorial board chair for the Online Ethics Center.

Elizabeth Sanders

Elizabeth A. Sanders is an Engineering Education Ph.D. student at Purdue University. She holds a B.S. in Chemical and Biomolecular Engineering (University of Illinois Urbana-Champaign, 2018) and an M.A. in Higher Education (University of Michigan, 2020).

Nicholas D. Fila (Research Assistant Professor)

Measuring and Promoting Empathic Formation in a Multidisciplinary Engineering Design Course

Abstract

Empathy is an important skill and disposition in engineering education but measuring and assessing empathy in specific engineering contexts is a novel domain of research. In this study, we iterated on a measure of empathy in engineering design. In this refined instrument, we measured and compared responses to the same set of survey items in different configurations. In the first configuration, we measured *Cognitive Empathy* and *Affective Empathy* across three design phases. In the second configuration, we retained the focus on *Cognitive Empathy* and *Affective Empathy* and variation across three design phases, but we also differentiated between self- and other- orientated empathy. An example construct in this second configuration is *Imagine-Other Cognitive Empathy in Needfinding*. To provide evidence of the trustworthiness of constructs, we computed Cronbach's alpha as a measure of internal consistency reliability and identified Spearman correlations with four extant empathy constructs as a means of external validity. All constructs in the first configuration were reliable but several constructs in the second configuration were unreliable. However, many constructs in both configurations exhibited moderate to large correlations with four existing constructs. We found students exhibited significant changes in *Cognitive Empathy in Needfinding*, but students did not exhibit changes in affective or cognitive empathy in other design phases. However, by employing the second configuration, we found that students demonstrated significant and positive changes in *Imagine-Other Cognitive Empathy* in two design phases (Concept Generation and Solution Evaluation) while exhibiting no changes in *Imagine-Self Cognitive Empathy*. We also analyzed students' written responses to an open-ended question pre/post-course. This analysis revealed that, after participating in this course, students: (1) situated users as the primary rationale for design work, (2) understood addressing users' needs as critical to design work, and (3) exhibited broadened definitions about who (or what) constitutes a user. This work provides instructors with a means to assess students' empathy with and for users in design and to more purposefully target students' empathic development whilst accounting for engineering design phases.

Keywords: Empathy; Engineering design; Assessment; Psychometrics

Introduction

Recent research on empathy in engineering has proliferated [1-3], but assessing empathic formation remains a challenge due to the complexity of empathy as a phenomenon [4]. One of the key challenges for measuring empathy is that it has numerous unique operationalizations [5, 6]. To this end, much extant research focuses on discrete empathy concepts [5] and dimensions [4]. While these concepts and dimensions are often related, they are theoretically distinct [5] and, thus, distinctly measurable [7]. Moreover, researchers outside of engineering have argued for the importance of contextual measures of empathy [4]. This argument has led to the design of many different measures and conceptualizations on how empathy manifests uniquely in specific contexts, such as health professionals [8], nursing [9], and social work [10].

These studies suggest there is a need for more coherent frameworks and tools for studying empathy in engineering, including a more concerted focus on specific engineering contexts. To this end, our team previously developed a measure of empathy in engineering design, which accounted for ways that empathy distinctly manifested across three design phases: needfinding, concept generation, and evaluation [11]. Our guiding theory in this work was that empathy may manifest in engineering in ways that are related but distinct when compared with other disciplinary contexts, across engineering contexts, or even within different phases of a single engineering context, such as design.

This study aims to continue this line of work and includes two parts. First, we iterate on a psychometric measure of empathic tendencies as well as a measure of empathy in engineering design. Second, we employ the refined instrument to identify quantitative and qualitative evidence of students' empathic development in a junior-level multidisciplinary engineering design course. We address the following research questions (RQs):

- RQ1:** To what extent are a set of revised empathy constructs for measuring empathy in engineering design reliable (based on correlation coefficients) and valid (as aligned with extant empathy constructs)?
- RQ2:** To what extent do students exhibit changes in empathy in engineering design constructs between the start and end of the course?

Background and Motivation

Conceptualizing Empathy

The question, "What is Empathy?" will likely generate different responses from different individuals and audiences. Batson [5] provided a list of eight overlapping but distinct phenomena or "concepts" called empathy in the context of social psychology. These empathy concepts varied in numerous ways, including their emphasis on identifying other's internal states, understanding other's perspectives, and matching other's emotions. Hess et al. [3] offered naming conventions for each of Batson's empathy concepts, including: (1) empathic accuracy or theory of mind; (2) motor mimicry; (3) emotional contagion; (4) projection: imagine-self within another's position; (5) perspective-taking – imagine other; (6) perspective-taking; imagine-self-as-if-self was the other; (7) empathic distress; and (8) empathic concern or sympathy.

While complex, Batson's [5] list is not comprehensive. For example, Cuff et al. [6] identified 43 distinct definitions of empathy. They suggested that definitions of empathy vary by eight themes, including whether empathy is cognitive or affective, a congruent or incongruent feeling, a trait or a state, self- or other-oriented, inclusive of behavior or simply a motivator of behavior, controlled or automatic, how it is affected by various stimuli, and its relationship to other concepts. This research shows there is a lack of consensus regarding what empathy is, exactly, and supports the need for authors to make explicit the concepts and dimensions that they are studying when using the term "empathy."

As one example of the contentions in conceptualizing empathy, sympathy, tenderness, and compassion are sometimes identified as empathy but often identified as "related concepts" [6, p. 145]. For example, sympathy is sometimes considered as "feeling for" another, whereas empathy is described as "feeling as" another [6, p. 145]. Here, "feeling for" another entails feeling concern for another person, whereas "feeling as" another entails feeling the *same* emotion, or at least feeling a congruent emotion, as another [6]. According to Cuff [6], distinctions between tenderness and compassion are more nuanced. Compassion generally involves experiencing an urge to help another after observing their emotion(s), while tenderness emphasizes the vulnerability of the individual being observed [6]. These distinctions highlight the import of scholar's being explicit with their definitions of the empathy concepts or dimensions they study, including concepts that they situate as other-than empathy.

Clark et al. [4] performed a systematic review on organizational definitions of empathy. Based on their analysis, they argued that there are three primary existing empathy "dimensions": cognitive, affective, and behavioral. In terms of Batson's [4] framing, we identify empathic accuracy or theory of mind, projection, and perspective-taking as distinct *cognitive empathy* concepts. Conversely, we identify emotional contagion, empathic distress, and empathic concern or sympathy as *affective empathy* concepts. Finally, of Batson's eight empathy concepts, we previously identified motor mimicry (i.e., the immediate reaction to another's' observable behaviors) as a type of affective empathy but we note that Clark et al. [4] identified this empathy concept as a type of behavioral empathy. As Cuff indicated [6], there is not consensus in the literature on whether empathy is a behavior or an outcome of a behavior, but Clark et al. included behavior as an empathy dimension and thus made this categorization possible [6]. At a minimum, many authors suggest that affective and cognitive empathy lead to empathic behaviors [4, 12, 13] – whether behavior constitutes empathy itself is an ongoing debate [6].

Building off these ideas, Hess and Fila [3] developed a four-part framework for conceptualizing empathy that varies in terms of affect/cognition and self/other orientation. This framework aligned with ways of viewing empathy as an "overarching category" that contains "associated concepts" [6]. Figure 1 provides a graphical depiction of four empathy concepts mapped to self/other orientation and cognitive/affective dimensions. While this list does not include all empathy concepts, it includes four concepts that we believe are salient for engineering education instructors and researchers. We employed this four-part framework in our initial instrument on empathy in engineering design [11] and we iterate on this framework in this study.

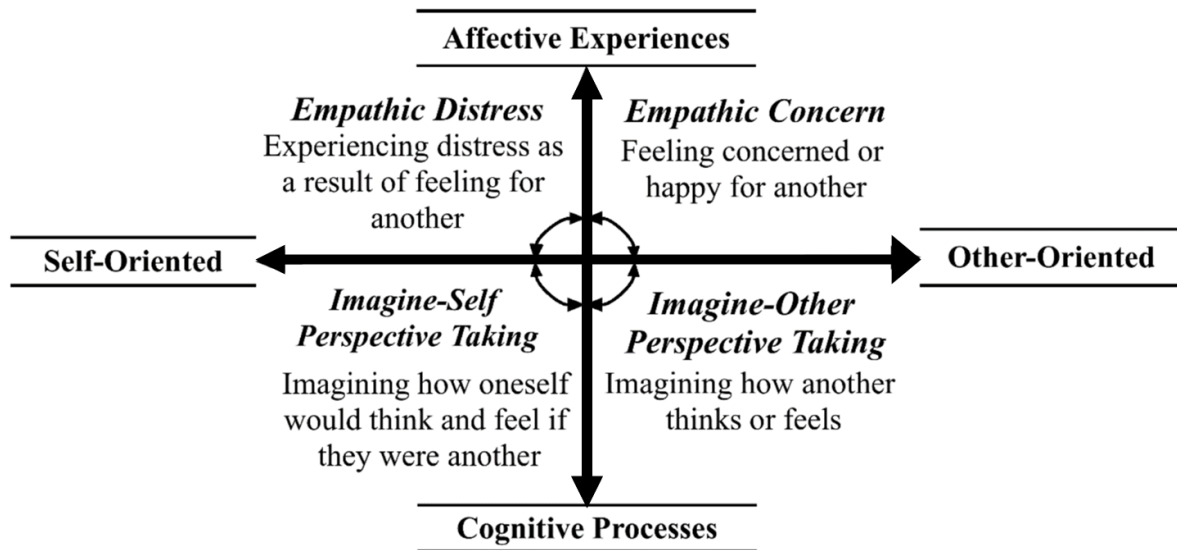


Figure 1. Conceptualizing Empathy (framework from [3])

Measuring Empathy in General

There exist many prominent quantitative measures of empathy, such as the Interpersonal Reactivity Index or IRI [7], the Empathy Quotient [14, 15], and the Questionnaire Measure of Emotional Empathy [16], to name a few. However, “the organizational literature lacks consensus on how empathy should be conceptualized, measured, and studied” [4, p. 167]. Thus, the variation in conceptualizations leads to various strategies to measure empathy. Using these three instruments as an example, the IRI is a multi-dimensional instrument that measures empathic tendencies along four constructs that vary by affect/cognition and self/other orientation (similar but not identical to the concepts described in Figure 1); the Empathy Quotient includes affective and cognitive items but combines these responses into a single empathy score; and the Questionnaire Measure of Emotional Empathy focuses on self/other-oriented *affective empathy* types and their resultant vicarious behaviors, thus excluding cognitive empathy.

The Interpersonal Reactivity Index or IRI [7] is commonly used to measure empathy both in organizations [6, 17] and in engineering education [17-21]. The IRI includes four constructs: (1) *Perspective-Taking*, or the tendency to take another’s perspective; (2) *Empathic Concern*, or the tendency to feel feelings of concern for another; (3) *Fantasy*, or the tendency to imagine one’s self in the place of characters in books, movies, or plays; and (4) *Personal Distress*, or the tendency to become tense or anxious as a result of stressful encounters. *Perspective-taking* and *empathic concern* are often deemed the two best empathy concepts offered by the IRI [4, 14, 22]. Based on this literature, we use these two constructs from the IRI in this study.

While Clark et al. [4] found that organizations tended to use extant instruments for measuring empathy (such as the IRI), several organizations designed novel strategies that also took into account the context. Based on this distinction, Clark et al. categorized measures of empathy as focusing on empathic “traits” versus empathic “states.” Trait-focused studies measure empathic tendencies (i.e., how one *tends* to act) whereas state-focused studies identify how empathy manifests in specific contexts (i.e., how one acts or *tends* to act in a specific context). Their first

recommendation for research on empathy was to recognize “empathy as a construct with both trait and state components” (p. 178). By embracing both conceptions, researchers can identify how individuals may employ empathy, in general, versus in their specific professional contexts.

Measuring Empathy in Engineering and Engineering Design

Based on the state/trait distinction offered by Clark et al. [4], multiple scholars have conjectured that empathy manifests in specific contexts in different ways. For example, in the context of engineering, Walther et al. [2] argued that we can conceive of empathy in both similar and distinct ways of other contexts. As they indicated, while many traditional empathy concepts are important to engineers, some aspects of empathy are uniquely important to engineers, such as holistic service to society. Likewise, Surma-Aho et al. [23] provided a model of empathy in design based on extant empathy conceptualizations and argued that five empathy concepts are core to design: “empathic understanding, empathic design research, empathic design action, empathic orientation, and empathic mental processes.”

While Clark et al. [4] focused on quantitative measures of empathy used in organizations, researchers have also used qualitative procedures to measure or assess empathy, including in the context of engineering. For example, Walther et al. [24] used student reflection data to identify engineering students’ empathic formation and guided by an empathy in engineering model [2] and Sochacka et al. [25] used reflection data to identify how empathic communication is influenced by mental models. Other scholars have focused on assessing empathy’s manifestation in engineering design. For example, Fila, et al. [26] identified empathic approaches to design by using a think-aloud protocol and thematic coding procedures guided by a prior empathy model and Guanés et al. [27] used interview procedures in capstone a design course to review how students’ empathic approaches were affected by their beliefs about empathy and engineering. These qualitative studies often focus on factors that influence empathy’s manifestation and the outcomes of empathy’s manifestation rather than focusing on empathic processes alone.

There are many studies in engineering wherein empathy is not the direct focus but an important aspect of a study’s findings. These findings can also provide unique guidance for measuring and assessing empathy. For example, while not seeking to directly measure empathy, Zoltowski et al. [28] conducted interviews to understand the different ways students experience human-centered design. They found that *empathic design* was the most comprehensive way of doing so, which was supported by students’ demonstration of advanced design strategies and deep understanding of the users they were designing for. While the “nested hierarchy” [28, p. 48] of ways that students experienced human-centered design might not all entail empathic practices, their findings can illuminate practices that precede empathetic engagement with users that may be useful to measure to better understand empathic formation in the context of engineering design.

Methodology

Prior Instrument Design and Rationale for Revisions

Our prior instrument design built on the ideas presented in the background and aimed to measure empathy concepts in engineering design. We employed a measure of empathy which had students reflect on how empathic they were toward users in a specific context (i.e., design). More specifically, survey respondents reflected on how empathic they were with users during three

engineering design phases of their course project: (1) needfinding, (2) concept generation, and (3) solution evaluation. We found empathy to manifest to varying levels across these three design phases [11]. We employ and slightly expand this set of survey items and constructs in this study with the overall objective of developing a more robust measure of empathy in engineering. In this section, we identify our instrument iterations and rationale for iterations.

Like our prior study [11], we hypothesize that empathy may manifest differently across three engineering design phases. For example, we theorized that how self-oriented affective empathy manifests in the needfinding phase may differ from how self-oriented affective empathy manifests in the concept generation and solution evaluation phases of design. This study offers additional insights into this theory with a small sample of students, in a single course context (i.e., a multidisciplinary design course), and with a slightly refined instrument. We identify correlations between constructs (research Phase 1) and student changes exhibited in this course (research Phase 2). More details on the course and research design are provided in a later section.

While our prior instrument provided structural validity evidence, we had ongoing concerns, especially those associated with content validity [29]. These primary challenges and resultant iterations included (1) iterations on the empathy conceptualization and construct design to generate a more representative instrument for assessing empathy in engineering design; (2) additional items aligned with this re-conceptualization grounded in extant views of empathy in engineering design, and (3) modified prefaces to survey items to boost confidence in comparing pre/post response and better ensure that students are considering how empathy manifested uniquely across distinct design phases.

Iteration 1: Empathy Conceptualizations and Associated Constructs

First, we designed items aligned with a four-part empathy model that differentiated between self/other orientation and cognitive/affective, thus resulting in four empathy constructs. In our instrument iteration, we aimed to measure four empathy constructs per design phase. First, we measured self-oriented affective empathy (which includes concepts like empathic distress and emotional congruence [3] and aligns with *feeling with*) separately from other-oriented affective empathy types (including concepts like empathic concern or sympathy which align with *feeling for* [3]). Figure 2 provides our re-conceptualization of the four empathy constructs.

Most notably, Figure 2 represents a *broadened* conceptualization compared to Figure 1. In Figure 2, we emphasize the empathy quadrant rather than specific empathy concepts. For example, while empathic distress is an “empathy concept,” this is but one concept among several that is aligned with self-oriented affective empathy. The broadened framework enabled us to consider other empathy concepts ostensibly aligned with self-oriented affective empathy, such as emotional congruence or motor mimicry. This realization came to us during our process of creating new items associated with (originally) empathic distress, and also helped us realize that this broadening led us to a more representative view of empathy by opening the door to other empathy concepts. In turn, we reframed *Perspective-Taking* constructs in Figure 1 as *Cognitive Empathy* constructs in Figure 2, thus broadening these constructs to general cognitive acts that involve *thinking as* others (self-oriented) and *thinking of* others (other-oriented). We used this broadening to help us consider cognitive empathy concepts beyond perspective-taking alone.

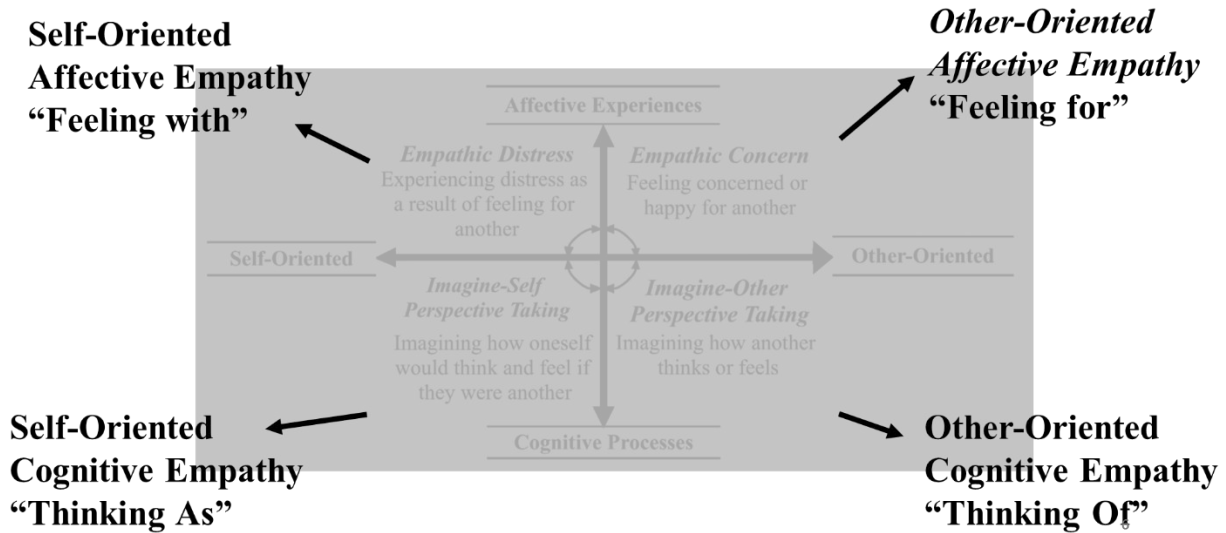


Figure 2. Revision of prior model and overview of four empathy constructs

While Figures 1 and 2 distinguish between self/other affective empathy, in our previous instrument design [11] we generated items associated with a single “Affective Empathy” construct, thus foregoing separate self/other-oriented affective empathy constructs. This prior approach was based on the challenge of measuring affective empathy types distinctly [4]. However, this separation ran counter to Batson [5] and the prior model (see Figure 1), which identified affective empathy as theoretically varying between affective self/other orientations. In short, researchers agree that *feeling for* (self-oriented) and *feeling as* (other-oriented) are distinct phenomena [4, 5], but we omitted this separation in our prior instrument [11]. In our iteration, we sought to measure these phenomena distinctly.

Iteration 2: Generation of New Survey Items

Our modified survey included 38 items, as shown in the Appendix. We included at least three items per empathy type and per design phase. Thus, we included a minimum of 12 items per design phase and 9 items per empathy type across the three design phases. In the concept generation section, we included four items on imagine-other cognitive empathy and four items on imagine-self cognitive empathy. All items were collected on a six-point Likert type-scale where responses were collected on a scale from 1 = “strongly disagree” to 6 = “strongly agree.”

Our shifts in conceptualization (i.e., Figure 1 to Figure 2) led to the need for the creation of new items. The need for more affective empathy types was most apparent, since the prior instrument only included five affective empathy items total (one in needfinding, two in concept generation, and two in solution evaluation). As a guide, we sought at least three items per construct to enable us to compute reliability statistics for discrete constructs. Thus, we needed a minimum of nine self-oriented affective empathy items and nine-other oriented affective empathy items. In other words, we needed at least 13 additional affective empathy items.

In our item creation process, we also purposefully generated more positive items in addition to our tendency towards the negative in the prior survey [11]. Positive items may be associated with feeling happy or relieved, whereas negative items may emphasize emotions such as feeling guilt or shame. For example, for the Self-Oriented Affective Empathy construct within Needfinding, we added the negative question, “I felt guilty if I was unable to understand users' perspectives,” and the positive question, “I felt relieved when I was able to identify users' needs.”

The Appendix also includes a summary of descriptive statistics for each item. While we do not include a focus on descriptive statistics in this study, we recognize that sharing these results can enable readers to interpret the outcomes of Phases 1 and 2 in this study with more clarity. Moreover, a review of these results suggests that an item-analysis can guide future revisions and implementations of this survey, especially considering Phase 1 and 2 results. Such future work might seek a larger sample of student respondents and analyze other statistics, such as skewness, kurtosis, inter-item correlations, and confirmatory factor analyses.

Iteration 3: Survey Preface Questions

As in the initial survey [11], in the pre-test, we asked students how empathy manifested in their previous design experiences. The question preceding each Design Phase section was, “While responding to these questions, consider how you thought, felt, and behaved during your prior design experiences.” Thus, students answered numerous Likert-type questions with this framing in mind. In our iteration, we included two additional questions in the pre-test to better understand the extent, nature, and self-perceptions of students’ prior design experiences: (1) In the context of design, what is a user?; and (2) What prior design experiences have you had? These preface questions enable us

Like the previous survey implementation [11], in the post-test, we asked students to reflect on how empathy manifested in their design project. Specifically, we prompted students to consider, “While responding to these questions, consider how you thought or felt during your [course name] design project experiences.” We also included two qualitative questions this iteration, wherein the first question was the same in both pre/post tests and the second was course specific.

- In the context of design, what is a user? (*students addressed this both pre and post course*)
- Who was the primary user that you designed for in [this course] this semester? If there are multiple primary users you considered throughout your project, please list them. In the following sections, we ask that you consider your primary user(s) as you respond to items. (*students addressed this question only post-course*)

In this iteration, while we retained the general overarching question for the survey, we altered the guiding question for each design phase. Specifically, the revisions included altering the guiding question in the Needfinding section from, “While reading or hearing about the design scenario...” to, “When IDENTIFYING design problems or challenges...” Similarly, we altered the guiding question in the Concept Generation section from, “While generating my design ideas...” to “When GENERATING design ideas or criteria...” Finally, we altered the guiding question in the Solution Evaluation section from, “While evaluating my ideas...” to “When EVALUATING

design solutions...” While subtle, this framing theoretically prompted respondents to consider how empathy manifested in slightly distinct ways when compared to the prior survey.

Course Overview

To test the revised instrument, we collected data with a small sample of students from a course that focused on engineering design methodologies and the needs of diverse stakeholders. The course integrates topics drawn from a variety of disciplines (including anthropology, psychology, human-computer interaction, and engineering) to introduce students to interdisciplinary engineering approaches, including: (1) design tools, systems, and environments to support cognitive processes, (2) human-centered design techniques, and (3) critical reflection experiences. The semester design project provided a context for students to apply these approaches and emphasizes human-centered design.

Assignments in the course included weekly reading reflections (example texts include *The Design of Everyday Things* [30]; “What Google Learned From Its Quest to Build the Perfect Team [31]), a design journal with guided prompts (example prompts are provided below), end-of-semester self-reflections on students’ personal design philosophies, a group user needfinding assignment, a resource sharing assignment where students identified and presented on design techniques, and a group design project report and presentation. Students completed the group design project in teams of four, the needfinding assignment in a team of two and the resource sharing individually or with a partner (based on student choice).

We incorporated ‘mindsets’ into this iteration of the course for the first time. The mindsets spanned each week and aligned with the course assignments, readings, and weekly topics. Mindsets included Autonomy, Embrace Vulnerability, Beginner’s Mind [32, 33], Empathize with Peers, Embrace Ambiguity & Leave your Comfort Zone [34], Challenge Assumptions & Biases, Empathize with Users, Optimism [35], Entertain Wild Ideas [36], Create to Destroy [37], Iterative Thinking, Feedback is a Gift [38, 39], and Design is a Story. We designed weekly journal prompts in alignment with these mindsets. Example journal prompts include:

- As you finalize Milestone 3, reflect on your ability to empathize with users. For example, in what ways do you feel like you were able to understand and relate to user perspectives and experiences?
- How will you empathize with peers this semester? What benefits do you think this will have on your teamwork experiences? What benefits do you think this will have on your design process?
- How will you practice vulnerability during your design project this semester? What benefits do you think this will have on your teamwork experiences? What benefits do you think this will have on your design process?

Students provided interim reports on their design work through a series of project milestones. Throughout the course, we emphasized that the final design product (i.e., prototype deliverable) itself was not as important as the process of design *and* students’ collaboration. Thus, in addition to the milestones, we conducted two self-team evaluations. We met with students and teams who experienced challenges to generate ideas for improving team collaboration.

Participant Overview

We performed quantitative analysis with a sample size of 12 students who completed both pre and post-tests. While all students did not complete both surveys, all 16 students completed the pre-test and provided consent to participate in this study. In Table 1, we provide pseudonyms and demographic data for only the 12 students who completed both surveys. Due to the small sample size, we omitted demographic information to help protect students' identities. Example demographic information that we possess but did not provide in Table 1 include gender, concentration in multidisciplinary engineering or interdisciplinary engineering studies, and extracurricular and professional activities. Students pursued concentrations in computer science, engineering management, general engineering, nano-engineering, theatre, and visual design.

Table 1. Participant Overview

Pseudonym	Year	Major
Alex	Junior	Multidisciplinary Engineering
Blake	5 th Year	Multidisciplinary Engineering
Cameron	Junior	Multidisciplinary Engineering
Dakota	Junior	Multidisciplinary Engineering
Emmett	Senior	Multidisciplinary Engineering
Frances	Junior	Multidisciplinary Engineering
Grey	Senior	Multidisciplinary Engineering
Harper	Junior	Interdisciplinary Engineering Studies
Indigo	Junior	Multidisciplinary Engineering
Jordan	Junior	Multidisciplinary Engineering
Kai	Senior	Multidisciplinary Engineering
Lee	Senior	Multidisciplinary Engineering

Research Phases

In this study, we addressed two research questions. Each research question included two phases, as shown in Figure 1. We describe each research phase in the following sub-sections.

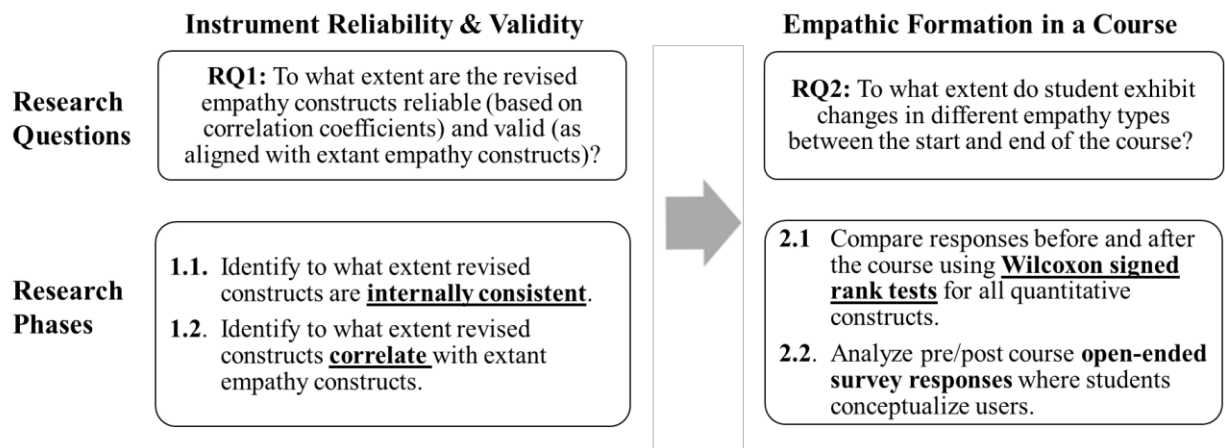


Figure 1. Research Design Overview

Phase 1.1. Reliability Testing

Reliability is “the degree to which scores in a particular sample are precise” [40, p. 90]. Our reporting of reliability statistics is a form of structural validity [29], thus identifying whether the items collectively represent the same phenomenon. The “reliability coefficient reported most often in the literature is coefficient alpha, also called Cronbach’s alpha” (p. 91). Cronbach’s alpha is a measure of internal consistency reliability and identifies how consistent responses are to a set of items. Internal consistency reliability tends to be “higher as there are more items or the average interitem correlation increases” (p. 91). DeVellis [41] provided suggestions for designing scales as well as thresholds for interpreting Cronbach’s alpha results. Following DeVellis, we interpreted Cronbach’s alpha above .70 to be reliable, an alpha above .60 as minimally reliable, and alpha scores below .60 as unreliable.

Phase 1.2. Correlation Analysis

In addition to this refined instrument, we employed two constructs from the Interpersonal Reactivity Index [7] - *Empathic Concern* and *Perspective-Taking* - and two constructs from a recent validation study in engineering - *Interpersonal Self-Efficacy* and *Emotion Regulation* [42]. This correlation analysis is a form of external validity [29], as it identifies if our novel constructs align with extant empathy constructs in expected ways. *Empathic Concern* can be considered an other-oriented affective empathy concept and thus should theoretically correlate with our affective empathy in design constructs. Similarly, *Perspective-Taking* is a cognitive empathy concept which primarily includes other-oriented items. Thus, we anticipated significant positive correlations between this construct and the cognitive empathy in design constructs. *Interpersonal Self-Efficacy* theoretically aligns with behavioral empathy, specifically empathic communication [4], which is often considered an outcome of affective and cognitive empathy concepts [6, 7]. Thus, we expected to see positive relationships between both affective and cognitive empathy constructs and *Interpersonal Self-Efficacy*. Finally, *Emotion Regulation* is considered essential to controlling one’s emotions and thus may align with affective empathy in design constructs.

We investigated the relationship between these four empathy constructs and the refined empathy in engineering design constructs (see Appendix). We checked scatterplots of variables to check for normality. Many relationships appeared monotonic but non-normal. Thus, we computed Spearman’s rank-order correlation rather than Pearson’s product-moment correlation. As we tested relationships between numerous constructs, a Bonferonni correction would have us interpret significance at a more conservative value rather than .05 (i.e., the p-value divided by number of hypotheses tested [43]). In our analysis, we still report significance at p-values of .05, 0.01, and .001 but we emphasize that as we test 125 hypotheses, using a Bonferonni correction would have us interpret significance at $p < .0004$. Significance testing alone has limitations, so we also emphasize effect size using threshold suggested by Rea and Parker [44].

Phase 2.1. Wilcoxon Signed Rank Tests (RQ2)

Due to the small sample size, we compared pre- and post-course responses using Wilcoxon signed ranks test. This comparative testing is a form of consequential validity [29], as it enables

us to identify if students are exhibiting empathic growth in positive direction and, if not, it can inform the need for foci in future course offerings. The Wilcoxon Signed Rank compares the median difference between related samples (or, in the context of this study, pre-course and post-course responses) [45, 46]. We computed effect sizes using suggestions from Pallant [47], where effect size (r) is computed by dividing z by the square root of the number of paired responses. We used effect size threshold suggestions from Cohen [48] for r where less than 0.30 is a small effect, between 0.30 and 0.50 is a moderate effect, and greater than 0.50 is a large effect.

Phase 2.2. Thematic Analysis

We analyzed students' responses to the newly introduced open-ended, written response questions across the pre- and post-course surveys through thematic analysis. Like the correlation analysis, this thematic analysis is a form of external validity [29], as it provides evidence regarding whether student quantitative changes evident in 2.1 are aligned with or substantiated by qualitative patterns. Thematic analysis is a method used to analyze qualitative data and organize insights into a set of reportable themes or patterns [49]. This method involved engaging with the raw data, constructing themes to describe the data, and continuously returning to prior data to ensure that the set of themes generated by the researchers appropriately represented the data set. In our analysis, one author initiated the analysis, and the themes were reviewed and critiqued by the other authors based on the data and the framing of emergent themes.

Results

Phase 1.1. Reliability Testing

We computed internal consistency reliability at the dimension level (e.g., *Cognitive Empathy* and *Affective Empathy*) in each design phase, as we as by differentiating between self/other (Table 2).

At the dimension level, most constructs were reliable, but the following constructs were minimally reliable (1) pre-course responses to *Cognitive Empathy in Needfinding* ($\alpha = .66$) and (2) post-course response to *Affective Empathy in Concept Generation* ($\alpha = .61$). However, we found several constructs that accounted for self/other orientation were unreliable. These included (1) the pre-score for *Imagine-Other Cognitive Empathy in Needfinding* ($\alpha = .46$), the post-score for *Imagine-Self Cognitive Empathy in Needfinding* ($\alpha = .03$), and the post-score for *Self-Oriented Affective Empathy in Needfinding* ($\alpha = .46$); and (2) both the pre- and post-scores for *Self-Oriented Affective Empathy in Concept Generation* ($\alpha = .48$ and $\alpha = .37$, respectively) and the post-score for *Other-Oriented Affective Empathy in Concept Generation* ($\alpha = .51$). All empathy constructs that differentiated between self/other orientation within Solution Evaluation were minimally reliable.

The wide differences between reliability statistics for pre and post-tests may suggest that students were considering vastly different design contexts between pre/post. For example, in the pre-survey, students were considering *any* design experience, whereas in the post-test, students were considering the specific course design project. Future research might collect both sets of data simultaneously to identify in greater detail how this framing might influence construct reliability. Moreover, the relatively low scores on several *Affective Empathy* constructs suggests the need for improvement of these constructs, in particular, in future instrument iterations.

Table 2: Reliability Coefficients (Cronbach's Alpha)

Construct / Sub-Construct	Pre-Course α	Post-Course α
Need Finding		
Cognitive Empathy	.66	.79
Imagine-Other	.46*	.81
Imagine-Self	.65	.03*
Affective Empathy	.79	.82
Other-Oriented	.66	.76
Self-Oriented	.66	.46*
Concept Generation		
Cognitive Empathy	.86	.91
Imagine-Other	.71	.74
Imagine-Self	.71	.85
Affective Empathy	.77	.61
Other-Oriented	.76	.51*
Self-Oriented	.48*	.37*
Solution Evaluation		
Cognitive Empathy	.90	.88
Imagine-Other	.88	.69
Imagine-Self	.93	.83
Affective Empathy	.85	.85
Other-Oriented	.85	.89
Self-Oriented	.67	.62

Phase 1.2. Correlation Analysis

We computed Spearman's correlation (r_s) to assess the relationship between empathy constructs across three design phases with four extant empathy constructs: *Empathic Concern*, *Perspective-Taking*, *Interpersonal Self-Efficacy*, and *Emotion Regulation* (see Table 3). There was a significant correlation at $p < .001$ between *Affective Empathy in Needfinding* and *Empathic Concern*. While there were no other significant correlations at $p < .001$, many relationships were significant at $p < .01$ (see Table 3). Notably, as we tested numerous hypotheses, a Bonferroni correction would lead us to employ a threshold of .05 divided by 125 (total number of hypotheses tested) or $p < .0004$. We do not see any values significant at this threshold. Nonetheless, many parameters exhibited strong (i.e., $r_s > .60$) and very strong (i.e., $r_s > .80$) effect sizes [44].

When analyzing relationships by self/other orientation, no correlations exhibited significant relationships at $p < .001$. However, there were three very strong correlations and several strong correlations. Very strong correlations included: (1) the post-score for *Other-Oriented Affective-Empathy in Needfinding* and *Empathic Concern* ($r_s = .80$) and (2) the post-score for *Other-Oriented Cognitive Empathy in Solution Evaluation* and *Empathic Concern* ($r_s = .80$), and (3) *Other-Oriented Cognitive Empathy in Needfinding* and *Interpersonal Self-Efficacy* ($r_s = .80$).

While there were few very strong relationships, most strong and several “relatively” strong positive correlations (i.e., $r_s > .40$) were exhibited on the *post-test* compared to relatively few on the *pre-test*, especially with respect to the construct, *Empathic Concern*. These findings suggest that students felt they exhibited more empathic concern in this course compared to their prior design experiences. While there were fewer relatively strong relationships between *Perspective-Taking* and the empathy in design constructs across the design phases on the *post-test*, all responses were stronger on the *post-test* than the *pre-test*.

Table 3: Correlation Coefficients (Spearman’s r)

	Empathic Concern		Perspective-Taking		Interpersonal Self-Efficacy		Emotion Regulation	
	<i>Pre</i>	<i>Post</i>	<i>Pre</i>	<i>Post</i>	<i>Pre</i>	<i>Post</i>	<i>Pre</i>	<i>Post</i>
NEEDFINDING								
<i>Cognitive Empathy</i>	.51	.71**	.25	.58*	.62*	.69*	.63*	.27
Other-Oriented	-	.52	-	.67*	-	.80**	-	.13
Self-Oriented	.43	-	.29	.25	-	.22	.53	-
<i>Affective Empathy</i>	.36	.83***	.18	.48	.40	.53	.21	.17
Other-Oriented	.35	.80**	.26	.42	.56	.63*	.17	.13
Self-Oriented	.41	.79**	.07	.47	.49	.49	.43	.17
CONCEPT GENERATION								
<i>Cognitive Empathy</i>	.57	.57	.29	.37	.52	.44	.54	.19
Other-Oriented	.50	.70*	.23	.48	.58*	.49	.52	.17
Self-Oriented	.58*	.52	.27	.36	.51	.44	.48	.19
<i>Affective Empathy</i>	.56	.64*	.32	.50	.57	.52	.53	.27
Other-Oriented	.54	-	.41	-	.50	-	.45	-
Self-Oriented	-	-	-	-	-	-	-	-
SOLUTION EVALUATION								
<i>Cognitive Empathy</i>	.12	.74**	-.09	.34	.52	.34	.37	.20
Other-Oriented	.37	.81**	.04	.47	.62*	.43	.50	.25
Self-Oriented	.02	.56	-.04	.24	.46	.28	.34	.17
<i>Affective Empathy</i>	.34	.73**	-.07	.44	.47	.41	.15	.16
Other-Oriented	.34	.73**	.14	.42	.41	.55	.09	.12
Self-Oriented	.36	.56	-.05	.28	.48	.27	.07	.15

* $p < .05$; ** $p < .01$; *** $p < .001$

When we review trends between the empathy in design constructs, *Interpersonal Self-Efficacy* and *Emotion Regulation*, the pre-score correlations tended to be higher than post-course correlations. Moreover, we see significant correlations in Needfinding, but few in the later design phases. These results highlight design phase is important to account for in addition to the design project when instructors prompt students to consider how they empathized with users. Moreover, these findings indicate that the course may have played a role in students’ cognitive and affective empathic development, which we test in Phase 2.1.

Phase 2.1. Wilcoxon Signed Ranks Test

We computed Wilcoxon Signed Rank Tests in two configurations. The first included *Cognitive Empathy* and *Affective Empathy* without differentiating by self/other orientation (Table 4) and the second accounted for self/other orientation (Table 5).

Configuration 1 – Cognitive and Affective Empathy without Self/Other Orientation

In our first testing configuration, Wilcoxon signed rank tests showed that students exhibited statistically significant changes on *Cognitive Empathy in Needfinding* ($Z = 2.990$, $p = .001$), with a strong effect size ($r = .76$, note: we use Cohen [48] for effect size thresholds). While no other changes were significant, multiple other changes exhibited moderate effect sizes, including *Affective Empathy in Needfinding* ($r = .43$); *Cognitive Empathy in Concept Generation* ($r = .39$), and *Cognitive Empathy in Solution Evaluation* ($r = .36$).

Table 4: Wilcoxon Signed Rank Tests – Empathy at the Dimension Level

Construct / Sub-Construct	+	-	0	Z	p	r
<i>NEEDFINDING</i>						
Cognitive Empathy	11	1	0	2.99	.001*	.76
Affective Empathy	8	4	0	1.50	.142	.43
<i>CONCEPT GENERATION</i>						
Cognitive Empathy	7	3	2	1.35	.195	.39
Affective Empathy	6	6	0	0.39	.733	-.05
<i>SOLUTION EVALUATION</i>						
Cognitive Empathy	8	3	1	1.26	.229	.36
Affective Empathy	4	6	2	-0.36	.742	-.10

* $p < .05$; ** $p < .01$; *** $p < .001$

These results suggest that students exhibited the greatest gains in Cognitive Empathy overall, with the greatest gains in the Needfinding design phase and moderate gains in Concept Generation and Solution Evaluation design phases. Similarly, students exhibited moderate gains on Affective Empathy constructs in Needfinding, but slight reductions in Concept Generation and Solution Evaluation. These results indicate that a greater focus on Affective Empathy and the latter two design phases in future course implementations is important to focus on.

Configuration 2 – Cognitive and Affective Empathy with Self/Other Orientation

We computed Wilcoxon Signed Rank tests (Table 5) by differentiating between self and other orientation. We did not compute Wilcoxon Signed Rank tests for unreliable constructs. We found significant gains in *Imagine-Other Cognitive Empathy in Concept Generation* ($Z = 1.708$, $p < .05$) with a large effect size ($r = .90$). However, we found a slightly negative change in *Imagine-Self Cognitive Empathy in Concept Generation* ($Z = -0.161$, $p = .931$). Similarly, we found a significant increase in *Imagine-Other Cognitive Empathy in Solution Evaluation* ($Z = 1.708$, $p < .05$), with a moderate effect size ($r = .49$), but no significant change in *Imagine-Self Cognitive Empathy* in the same phase ($Z = .913$, $p = .385$). Finally, while not significant,

students exhibited a slight gain in effect on *Other-Oriented Affective Empathy in Solution Evaluation* ($r = .15$), but a negative change in *Self-Oriented Affective Empathy* in this phase.

These findings align with our theory that empathy types manifest differently across design contexts. Moreover, the results suggest that students generally exhibited greater gains in other-oriented empathy types compared to self-oriented empathy types, *especially* other-oriented cognitive empathy. These findings suggest a greater focus on *self-oriented* empathy may be important in future qualitative analysis and, potentially, future course offerings.

Table 5: Wilcoxon Signed Rank Tests – Empathy at the Type Level

Construct	+	-	0	Z	p	r
NEEDFINDING						
Imagine-Other Cognitive Empathy	pre score not reliable					
Imagine-Self Cognitive Empathy	post score not reliable					
Other-Oriented Affective Empathy	7	3	2	0.717	.441	.21
Self-Oriented Affective Empathy	post score not reliable					
CONCEPT GENERATION						
Imagine-Other Cognitive Empathy	8	1	3	2.432	.015*	.90
Imagine-Self Cognitive Empathy	5	6	1	-0.161	.931	-.05
Other-Oriented Affective Empathy	post score not reliable					
Self-Oriented Affective Empathy	pre and post scores not reliable					
SOLUTION EVALUATION						
Imagine-Other Cognitive Empathy	7	2	3	1.708	.010*	.49
Imagine-Self Cognitive Empathy	7	3	2	0.913	.385	.26
Other-Oriented Affective Empathy	6	4	2	0.514	.643	.15
Self-Oriented Affective Empathy	4	6	2	-0.594	.580	-.17

* $p < .05$; ** $p < .01$; *** $p < .001$

Changes in Four Extant Empathy Constructs

We next computed Wilcoxon Signed-Rank tests for the four existing empathy constructs. Our results indicated that students exhibited significant positive gains in *Perspective-Taking* ($Z = 2.613$, $p < .01$) with a large effect size ($r = .53$) and *Interpersonal Self-Efficacy* ($Z = 2.283$, $p < .05$), with a moderate effect size ($r = .47$). While not significant, students also exhibited moderate changes in *Empathic Concern* ($r = .39$) and *Emotion Regulation* ($r = .37$).

Table 6: Wilcoxon Signed Rank Tests – Prior Empathy Constructs

Construct / Sub-Construct	+	-	0	Z	p	r
Empathic Concern	7	5	0	1.893	.058	.39
Perspective-Taking	10	2	0	2.613	.009**	.53
Interpersonal Self-Efficacy	9	2	1	2.283	.022*	.47
Emotion Regulation	8	3	1	1.816	.077	.37

* $p < .05$; ** $p < .01$; *** $p < .001$

These results indicated that students' general empathic tendencies and related skills all improved at least to a moderate extent. In the spirit of Clark et al. (2019), results from Table 6 may be interpreted as students' trait-gains, whereas results from Tables 4 and 5 can be interpreted as students' state-gains. Thus, while results in Table 6 show students' overall empathic gains, results in Tables 4 and 5 identify numerous areas where the instructors may focus on in future iterations of the course to improve students' use of empathy across engineering design phases.

Phase 2.2. Qualitative Analysis

In this section, we asked, "In what ways do student reflections support the quantitative changes?" As described above, this course encouraged students to cultivate a deep understanding of their user, including user needs, characteristics, perspectives, and values.

The pre/post survey prompted students to reflect on how empathy manifested in previous design experiences (pre-course survey) and in their semester-long course project (post-course survey). Here, we examined the common question between the pre- and post- survey: "In the context of design, what is a user?" Some students provided pre- and post- survey responses that demonstrated little shift in conceptualization of the user (e.g., "*The person who utilizes the design in any form.*" → "*Anyone who utilizes the thing being designed.*"). However, some students demonstrated shifts in their conceptualization of users. We identified three primary themes which represent differences between pre- and post-course responses: (1) Designing FOR Users; (2) Users have Needs; and (3) Broadened Definitions of Users and their Needs.

Theme 1: Designing FOR a User

Students tended to leave the course operationalizing design as something someone does *for* a user, rather than a thing (e.g., an artifact or design product) that a user will use. Student language provides evidence of a shift in the students' perceived role of the user or users in design. This shift indicates students viewed users as more significant by positioning users as the reason or motivators of a design, rather than a passive beneficiary. For example, in the pre-test, students described users as individuals affected by design. In this conceptualization, the design artifact effects users, positively or negatively. In contrast, in the post-test, students articulated that the goal of design is to develop something for a user, situating the user in the more predominate role.

Table 7: Theme 1 – Designing FOR Users

Pseudonym	Pre-Test Response	Post-Test Response
Harper	A user is the person or the company that will be dealing with the product being designed.	In the context of design, a user is the person, or group of similar people, that interact with the design and have a need that needs to be helped.
Dakota	A user is the one that is able to without any effort, mold to any idea and utilize it.	A user is the one who uses the design.

This shift in conceptualization is exemplified by Harper (see Table 7). Dakota also shows a shift that assigns users more independence in the design process, as they originally describe the user

as someone who will “mold” a design, suggesting that the design may require the user to undergo a change to ensure usability. In the post-course response, these students omitted this step, suggesting the user will be able to use the design without changing something about themselves or their interaction with the design artifact. Therefore, students situated design artifacts in more of an active role and having an effect upon the user. However, in the post-test, students demonstrated an understanding that the goal of design is something to help the user.

Theme 2: Users have Needs

The pre- and post-course survey responses suggested that many students started the course perceiving design to be an activity intended to create a product that users will use. However, at the end of the course, three students elaborated on this, suggesting that designing *for* a user requires understanding the users’ needs and that design activities should be designed to ensure that the designer can effectively meet those needs (see Table 8). This is exemplified by Ollie whose definition of users shifted from a focus on a user as someone who design ‘directly affects’ to someone ‘whose need the product is designed to cater to.’

Table 8: Theme 2 – Users have a Need

Pseudonym	Pre-Test Response	Post-Test Response
Harper	A user is the person or the company that will be dealing with the product being designed.	In the context of design, a user is the person, or group of similar people, that interact with the design and have a need that needs to be helped.
Noel	In the context of design, a user is someone who will be using the designed product.	a user is someone who has needs that need to be solved
Ollie	The person that the product being designed directly affects.	The person or group of people whose need the product is designed to cater to.

Theme 3: Broadened Definitions of Users and Their Needs

Finally, students broadened their definitions of who (or what) could be a user, including how user groups may vary. In pre-course responses, students generally viewed users as a single individual. However, in the post-course reflections, some students shifted to describe users as groups of people or multiple people who will interact or be affected by the product of design work (see Table 9).

Indigo shared that “a user is a ‘customer’ to that final [design] product;” however, at the end of the course, they recognized that users could be considered anyone who is affected by a design (which could or could not include those beyond outside of “customers”). Emmett provided similar changes, wherein their post-course response evidenced an updated understanding about who could be designed for and what could be designed. The student suggests that in addition to “person(s),” designers can also design for groups and objects. Moreover, in addition to products, designers can also design services and processes to alleviate challenges experienced by their users. Interestingly, this student was part of the one group in the course that designed for a non-human user group.

Table 9: Theme 3 – Broadened Definitions

Pseudonym	Pre-Test Response	Post-Test Response
Harper	A user is the person or the company that will be dealing with the product being designed.	In the context of design, a user is the person, or group of similar people, that interact with the design and have a need that needs to be helped.
Emmett	A user is the person for which a product is designed. They are the people who will be using a product the most.	A user is the person/group/object for which the product/service/process is being designed.
Indigo	I think of design as yielding a final product. To me, the user is a "customer" to that final product. They are the ones most effected by the design.	The user is anyone effected by a design.
Kai	an individual that interacts with the product and is usually the one who is the one to be consider the intended person for in which something is designed.	Any individual that experiences or interacts with a product or system.

Discussion

We use Clarke et al.'s [4] four recommendations for “designing empathy research” to guide our discussion (p. 182). Specifically, Clarke et al. suggested organizations conduct (a) experimental research designs, (b) observational research designs, (c) longitudinal research designs, and (d) multi-level research designs. Accordingly, we discuss each of Clark et al.'s recommendations in turn, while reflecting on these recommendations in the context of this study, the course, and recommendations for assessing empathetic formation in engineering design education.

First, Clark et al. suggested that organizations should conduct *experimental research designs* which compare outcomes of empathy due to an intervention. We argue that many (if not all) engineering education contexts provide an opportunity to promote and assess students' empathic formation. To conduct experimental educational research on empathy, we encourage educators to consider using existing tools or crafting new tools to support students' empathetic development based on the empathy concepts they perceive as most important to their context. Researchers can apply tools created by other researchers or even iterate on these instruments based on their specific contexts or specific sought changes.

In the context of this course, we aimed to further iterate on a survey designed to assess empathy, while also understanding how students' empathetic development changed over the course of a multidisciplinary, human-centered design course with a semester-long design component. We found that students experienced development in different constructs across different design phases. Returning to the first recommendation (experimental designs), this assessment does not indicate specific elements of the course that supported students' empathic development but does suggest that students exhibited changes in different phases. Moving forward, future research can provide more information by using controlled research designs or collecting data to better understand students' progression at specific time points in a course.

Second, Clark et al. suggested organizations conduct ***observational research designs*** which view how empathy, especially *behavioral* empathy, manifests in context. As previously mentioned, Clark et al. [4] identified that several organizations sought novel ways of assessing empathy that took into account context. Surveys provide a useful tool to assess development across a time-bound experience, such as a course. However, observational research can provide organizations with insights specifically related to the context and provide deeper insights that quantitative research may omit. We suggest that employing and analyzing artifacts based on empathy use at multiple timepoints throughout a course should be more closely coupled to observing and documenting how students tend to conceive of and engage users when employing these design techniques. Developing processes to assess empathic formation through observational research techniques can include classroom observations but can also involve out-of-class activities, such as employing think-aloud interview protocols or encouraging community interactions.

In this study, while the survey provides us with evidence that students perceived empathic development across multiple (but not all) constructs and design phases, the survey does not provide specific insights into the experiences that helped promote that development. Throughout the course, the instructional team sought to engage students in empathic practices, such as taking a perspective of the user (e.g., university students experiencing sleep deprivation, dog shelter volunteers); however, the instructional team did not conduct a formal analysis of students' artifacts to supplement the survey in assessing students' empathic formation. To this end, instructors can leverage existing artifacts (such as student reflections [24]) in tandem with survey findings to identify where and what aspects of a course promote or inhibit students' empathic formation.

Third, Clark et al. suggested organizations conduct ***longitudinal research designs*** which examine empathic development over time. Longitudinal research is a critical need in engineering education, given documented declines in students' concern for social welfare across multiple universities [50]. However, longitudinal research studies are time and resource intensive. Thus, while we encourage instructors to document empathic development throughout single design courses, we also encourage researchers to consider the empathic development of students across multiple courses or throughout programs. Such studies could aim to triangulate annual survey responses with qualitative data obtained near the culmination of a students' undergraduate tenure, such as through exit interviews or reflection data in capstone course experiences.

Finally, Clark et al. suggested organizations employ ***multilevel research designs*** that focus on how empathy manifests as a state versus a trait. The addition of general empathic tendencies alongside state-specific measures of empathy is a novel approach to studying empathic development. Our study found overall gains in trait perspective-taking by using the IRI, but we also found that gains in Cognitive Empathy were evident in only a few constructs. Thus, these results support the need for (and differences resultant from) the use of trait-versus-state measures of empathy. Collecting and comparing such findings can help educators ensure empathically-inclined students have ample opportunity to translate their inclinations into practice.

Limitations & Future Work

We hope that this paper will encourage others to iterate on existing instruments for assessing empathy in their specific course contexts. We often find that individuals are reluctant to modify

existing instruments due to validity concerns. However, iterations may be necessary to measure a complex phenomenon, such as empathy, in specific contexts (such as engineering design) in valid ways. Thus, we provide our iterations to provide others with courage to iterate on the measures they employ in their own research and teaching. As we did, they may modify instruments and use these alongside extant validated instruments. Importantly, our study includes a small sample size and a single context. Thus, in recognition that validation is an ongoing process [51], we consider these findings as preliminary insights that can inform future studies on measuring empathy in engineering design.

Conclusion

In this study, we iterated on a psychometric measure of empathy in engineering design and situated the refined measure in a multidisciplinary engineering design course. In our first phase of analysis, we conducted reliability testing on refined constructs. We found that all constructs were reliable at the *Cognitive Empathy* and *Affective Empathy* levels, but we found several constructs were unreliable when we accounted for self/other orientation in addition to the cognitive/affect dimensions. Next, we identified how constructs correlated with four empathy constructs, including two from the Interpersonal Reactivity Index (empathic concern and perspective-taking) and two from prior studies in engineering (interpersonal self-efficacy and emotion regulation). We found many strong relationships, especially between the empathy in design constructs and empathic concern. Third, we employed the refined instrument to identify quantitative and qualitative evidence of students' empathic development in design. We found significant changes in *Cognitive Empathy* in *Needfinding*, as well as *Other-Oriented Cognitive Empathy* in *Concept Generation* and *Solution Evaluation*. These findings revealed the import of accounting for self/other orientation and also suggested that students exhibited gains in empathic use primarily in the front-end of design. Finally, we accounted for changes in students' perceptions of users before and after the course. This qualitative analysis suggest that students developed broadened definitions of users. Taken together, these results can guide others as they seek to measure and assess empathy in their unique contexts.

Acknowledgements

This study was made possible by the research participants, including the students involved in the multidisciplinary engineering design course. We thank you for your time and energy throughout the course and your willingness to candidly share your experiences with our research team via myriad data collection strategies.

References

- [1] N. W. Sochacka, D. A. Delaine, T. G. Shepard, and J. Walther, "Empathy Instruction through the Propagation Paradigm: A synthesis of developer and adopter accounts," *Advances in Engineering Education*, vol. Spring 2021, 2020.
- [2] J. Walther, S. E. Miller, and N. W. Sochacka, "A model of empathy in engineering as a core skill, practice orientation, and professional way of being," *Journal of Engineering Education*, vol. 106, no. 1, pp. 123-148, 2017, doi: <http://dx.doi.org/10.1002/jee.20159>.
- [3] J. L. Hess and N. D. Fila, "The development and growth of empathy among engineering students," presented at the American Society for Engineering Education Annual Conference, New Orleans, LA, 2016.
- [4] M. A. Clark, M. M. Robertson, and S. Young, "'I feel your pain': A critical review of organizational research on empathy," *Journal of Organizational Behavior*, vol. 40, no. 2, pp. 166-192, 2019.
- [5] C. D. Batson, "These things called empathy: Eight related but distinct phenomenon," in *The Social Neuroscience of Empathy*, J. Decety and W. Ickes Eds. Cambridge, MA: MIT Press, 2009, ch. 1, pp. 3-15.
- [6] B. M. Cuff, S. J. Brown, L. Taylor, and D. J. Howat, "Empathy: A review of the concept," *Emotion Review*, vol. 8, no. 2, pp. 144-153, 2016.
- [7] M. H. Davis, "Measuring individual differences in empathy: Evidence for a multidimensional approach," *Journal of Personality and Social Psychology*, vol. 44, no. 1, pp. 113-126, 1983.
- [8] P. Irving and D. Dickson, "Empathy: towards a conceptual framework for health professionals," (in English), *International Journal of Health Care Quality Assurance*, vol. 17, no. 4/5, pp. 212-220, 2004.
- [9] D. Kunyk and J. K. Olson, "Clarification of conceptualizations of empathy," *Journal of Advanced Nursing*, vol. 35, no. 3, pp. 317-325, 2001.
- [10] S. King Jr. and M. J. Holosko, "The development and initial validation of the empathy scale for social workers," *Research on Social Work Practice*, vol. 22, no. 2, pp. 174-185, 2012, doi: 10.1177/1049731511417136.
- [11] J. L. Hess, N. D. Fila, E. Kim, and S. e. Purzer, "Measuring empathy for users in engineering design," *International Journal of Engineering Education*, vol. 37, no. 3, pp. 733-743, 2021.
- [12] M. L. Hoffman, *Empathy and moral development: Implications for caring and justice*. Cambridge, UK: Cambridge University Press, 2000.
- [13] M. H. Davis, *Empathy: A social psychological approach* (Social Psychology Series). Boulder, CO: Westview Press, 1996.
- [14] E. J. Lawrence, P. Shaw, D. Baker, S. Baron-Cohen, and A. S. David, "Measuring empathy: Reliability and validity of the Empathy Quotient," *Psychological Medicine*, vol. 34, no. 5, pp. 911-920, 2004, doi: <https://doi.org/10.1017/S0033291703001624>.
- [15] S. Baron-Cohen and S. Wheelwright, "The Empathy Quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences," *Journal of Autism and Developmental Disorders*, vol. 34, no. 2, pp. 163-175, 2004.
- [16] A. Mehrabian and N. Epstein, "A measure of emotional empathy," *Journal of Personality*, vol. 40, no. 4, pp. 525-543, 1972.

- [17] C. Rasoal, H. Danielsson, and T. Jungert, "Empathy among students in engineering programmes," *European Journal of Engineering Education*, vol. 37, no. 5, pp. 427-435, 2012.
- [18] J. L. Hess, J. Beever, C. B. Zoltowski, L. Kisselburgh, and A. O. Brightman, "Enhancing engineering students' ethical reasoning: Situating reflexive principlism within the SIRA framework," *Journal of Engineering Education*, vol. 108, no. 1, pp. 82-102, 2019.
- [19] J. L. Hess, N. D. Fila, and S. Purzer, "The relationship between empathic and innovative tendencies among engineering students," *International Journal of Engineering Education*, vol. 32, no. 3A, pp. 1236-1249, 2016.
- [20] A. O. Surma-aho, T. A. Bjorklund, and K. Holttä-Otto, "Assessing the development of empathy and innovation attitudes in a project-based engineering design course," presented at the ASEE Annual Conference & Exposition, Salt Lake City, UT, 2018.
- [21] J. L. Hess, G. A. Fore, B. H. Sorge, M. A. Coleman, M. F. Price, and T. W. Hahn, "Exploring ethical development from standard instruction in the contexts of biomedical engineering and earth science," presented at the ASEE 2019 Annual Conference, Tampa, FL, 2019.
- [22] J. Decety and C. Lamm, "Empathy versus Personal Distress," in *The Social Neuroscience of Empathy*, J. Decety and W. Ickes Eds. Cambridge, MA: MIT Press, 2009, ch. 15, pp. 199-214.
- [23] A. Surma-aho and K. Hölttä-Otto, "Conceptualization and operationalization of empathy in design research," *Design Studies*, vol. 78, p. 101075, 2022.
- [24] J. Walther, M. A. Brewer, N. W. Sochacka, and S. E. Miller, "Empathy and engineering formation," *Journal of Engineering Education*, vol. 109, no. 1, pp. 11-33, 2020.
- [25] N. W. Sochacka, K. M. Youngblood, J. Walther, and S. E. Miller, "A qualitative study of how mental models impact engineering students' engagement with empathic communication exercises," *Australasian Journal of Engineering Education*, pp. 1-12, 2020.
- [26] N. D. Fila, J. L. Hess, Ş. Purzer, and E. Dringenberg, "Engineering students' utilization of empathy during a non-immersive conceptual design task," *International Journal of Engineering Education*, vol. 32, no. 3B, pp. 1336-1348, 2016.
- [27] G. Guaness, L. Wang, D. A. Delaine, and E. Dringenberg, "Empathic approaches in engineering capstone design projects: student beliefs and reported behaviour," *European Journal of Engineering Education*, pp. 1-17, 2021.
- [28] C. B. Zoltowski, W. C. Oakes, and M. E. Cardella, "Students' ways of experiencing human-centered design," *Journal of Engineering Education*, vol. 101, no. 1, pp. 28-59, 2012.
- [29] S. Messick, "Standards of validity and the validity of standards in performance assessment," *Educational Measurement: Issues and Practice*, vol. 14, no. 4, pp. 5-8, 1995.
- [30] D. A. Norman, *The design of everyday things*. Basic books, 2002.
- [31] C. Duhigg. (2016) What Google learned from its quest to build the perfect team. *The New York Times Magazine*. Available: <https://www.nytimes.com/2016/02/28/magazine/what-google-learned-from-its-quest-to-build-the-perfect-team.html>
- [32] S. Suzuki, *Zen mind, Beginner's mind*. Boulder, CO: Shambhala Publications, 2020.
- [33] J. R. Morgan. "Beginner's mind." <https://uxdesign.cc/beginners-mind-in-ux-design-with-shunryu-suzuki-11a00787c8a9> (accessed May 13, 2022).

- [34] P. Martin. "Embrace ambiguity." IDEO. <https://www.designkit.org/mindsets/5> (accessed January 19, 2021).
- [35] J. Bielenberg. "Optimism." IDEO. <https://www.designkit.org/mindsets/6> (accessed January 19, 2021).
- [36] R. F. Dam and T. Y. Siang. "Learn how to use the best ideation methods: Brainstorming, braindumping, brainwriting, and brainwalking." Interaction Design Foundation. <https://www.interaction-design.org/literature/article/learn-how-to-use-the-best-ideation-methods-brainstorming-braindumping-brainwriting-and-brainwalking> (accessed January 19, 2021).
- [37] D. S. Tenbrink and J. Chizek, "A lecture on prototyping," ed, 2020.
- [38] M. Kuechenmeister. "Feedback is a gift." UX Collective. <https://uxdesign.cc/feedback-is-a-gift-4227d1202b32> (accessed January 19, 2021).
- [39] D. S. Tenbrink and J. Chizek, "A lecture on critique and feedback," ed, 2020.
- [40] R. B. Kline, *Principles and practice of structural equation modeling*, 4th ed. New York: The Guilford Press, 2016.
- [41] R. F. DeVellis, *Scale development: Theory and applications*. Los Angeles, CA: SAGE Publications, Inc., 2011.
- [42] J. L. Hess, A. Chase, G. A. Fore, and B. Sorge, "Quantifying interpersonal tendencies of engineering and science students: A validation study," *International Journal of Engineering Education*, vol. 34, no. 6, pp. 1754-1767, 2018.
- [43] B. G. Tabachnick and L. S. Fidell, *Using multivariate statistics*, 6th ed. Harlow: Pearson, 2014.
- [44] L. M. Rea and R. A. Parker, *Designing and conducting survey research: A comprehensive guide*. San Francisco, CA: Jossey-Bass, 2014.
- [45] D. C. Howell, *Fundamental statistics for the behavioral sciences*, 7 ed. Belmont, CA: Cengage Learning Inc., 2010.
- [46] E. McCrum-Gardner, "Which is the correct statistical test to use?," *British Journal of Oral and Maxillofacial Surgery*, vol. 46, no. 1, pp. 38-41, 2008.
- [47] J. Pallant, *SPSS survival manual: A step by step guide to data analysis using SPSS*, 4th ed. New York, NY: McGraw-Hill 2010.
- [48] J. Cohen, "A power primer," *Psychological Bulletin*, vol. 112, no. 1, pp. 155-159, 1992, doi: 10.1037/0033-2909.112.1.155.
- [49] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77-101, 2006.
- [50] E. A. Cech, "Culture of disengagement in engineering education?," *Science, Technology, & Human Values*, vol. 39, no. 1, pp. 42-72, 2014.
- [51] K. A. Douglas and Ş. Purzer, "Validity: Meaning and relevancy in assessment for engineering education research," *Journal of Engineering Education*, vol. 104, no. 2, pp. 108-118, 2015, doi: 10.1002/jee.20070.

Appendix. Overview of Survey Items and Descriptive Statistics of Items ($n = 12$)

Construct	Survey Item	Pre		Post	
NEEDFINDING: When IDENTIFYING design problems or challenges..		M	SD	M	SD
Imagine-Other Cognitive Empathy	I imagined a user's everyday activities in their real-life context.	4.25	0.75	4.75	0.97
	I imagined how a user feels when they experience a problem with existing designs.	4.33	0.78	5.17	0.58
	I thought of ways that users might interact with designs.	4.75	0.62	5.25	0.75
Imagine-Self Cognitive Empathy	I imagined how I felt when I experienced a problem with an existing design.	4.50	0.80	5.25	0.75
	I was able to relate to the challenges that users experience in their everyday life.	4.00	0.74	4.83	1.03
	I imagined challenges that I would experience if I were a user to help identify problems.	4.75	0.62	5.25	0.62
Other-Oriented Affective Empathy	I felt sorry for users who experience problems due to existing designs.	4.83	0.58	5.08	0.79
	I felt good for users who experience existing designs in positive ways.	4.75	0.62	5.17	0.58
	I felt concern for users who face challenges due to poor design.	5.00	0.74	4.75	1.06
Self-Oriented Affective Empathy	I felt a desire to identify ways to improve users' experiences with existing designs.	4.67	0.89	4.83	0.83
	I felt relieved when I was able to identify users' needs.	5.17	0.72	5.33	0.65
	I felt guilty if I was unable to understand users' perspectives.	4.17	1.11	4.42	1.31
CONCEPT GENERATION: When GENERATING design ideas or criteria...		M	SD	M	SD
Imagine-Other Cognitive Empathy	I tried to imagine myself in users' shoes to generate design criteria.	4.67	0.65	5.17	0.94
	I thought from users' perspectives to generate design ideas.	4.83	0.83	5.00	0.95
	I thought of ways to improve how users experience existing designs.	4.92	0.67	5.25	0.62
	I attempted to address design criteria that are important to users.	5.00	0.43	5.25	0.62
Imagine-Self Cognitive Empathy	I imagined user challenges to generate design ideas.	4.75	0.87	5.17	0.72
	I generated ideas by imagining that I was a user.	4.83	0.94	5.00	0.60
	I thought of positive aspects of designs that I have experienced in the past.	5.25	0.75	4.83	0.83
	I thought of what design criteria would be important to me if I were a user.	5.33	0.49	5.17	0.72
Other-Oriented Affective Empathy	I hoped that my ideas would be useful for users.	5.50	0.67	5.33	0.78
	I prioritized user safety due to concern for users.	4.67	0.89	5.08	0.79
	I hoped that my ideas would address user needs.	5.33	0.65	5.58	0.67
Self-Oriented Affective Empathy	I felt happy when I generated ideas that could be helpful to users.	5.33	0.65	5.50	0.52
	I got upset when I could not think of ideas that would be helpful for users.	3.92	0.79	3.33	1.07
	I felt guilty if I was unable to generate ideas that may have positive user outcomes.	3.92	1.51	4.33	1.30
SOLUTION EVALUATION: When EVALUATING design solutions..		M	SD	M	SD
Imagine-Other Cognitive Empathy	I imagined what aspects of my design solution(s) users would enjoy.	4.75	0.87	5.33	0.78
	I imagined why users might dislike my design solution(s).	4.50	0.90	4.75	0.75
	I imagined what aspects of my ideas users would find enjoyable.	5.00	0.95	5.17	0.72
Imagine-Self Cognitive Empathy	I imagined how I would use my design solution(s) if I were a user.	5.08	0.67	4.92	0.67
	I imagined what problems I would have with my design solution(s) if I were a user.	4.83	0.83	5.25	0.75
	I evaluated my design solution(s) by imagining that I am a user.	5.00	0.74	5.08	0.67
Other-Oriented Affective Empathy	I felt concern when my design solution(s) were unlikely to meet the needs of users.	4.92	0.90	5.08	1.00
	I felt good when my design solution(s) were likely to help users.	5.42	0.67	5.42	0.79
	I felt a desire to ensure that users experience my design in the best way possible.	5.17	0.72	5.08	1.00
Self-Oriented Affective Empathy	I felt upset when my design solutions were unlikely to meet the needs of users.	4.58	1.31	4.25	1.14
	I feel relieved when my design solution(s) were likely to help users.	5.33	0.49	5.25	0.75
	I felt guilty if users were likely to be unable to experience my design solution in the best way possible.	4.50	1.00	4.58	1.31