Experiential and Expression-based Accessibility Awareness Interventions to Improve Computing Education

Yang Liu, Domenic Mangano, Krishna Prasad Neupane, Samuel Malachowsky and Daniel E. Krutz
B. Thomas Golisano College of Computing and Information Sciences
Rochester Institute of Technology
Rochester, NY, USA
Email: {yl4070, dm9965, kpn3569, samvse, dxkvse}@rit.edu

***Abstract*—This research full paper presents a study demonstrating the importance of empathy-building content in computing education to foster appreciation for accessible software. Past research identifies a lack of empathy among developers contributing to software inclusion issues. We propose that enhancing empathy can lead to more accessible and equitable software development. Despite empathy studies in other fields, integrating empathy into computing education lacks sufficient pedagogical insights and accessible resources. To address these gaps, we conducted an in-person study with 121 participants, using a randomized experimental design with two empathy-building intervention groups. We compared their responses to pre-/post-survey questions using statistical tests, examining experiential and expression-based interventions' impact on inclusiveness interest and awareness of non-inclusive software. Our research questions focused on whether these interventions increase the interest, awareness, and empathy of participants and if they are equally effective across demographics and intervention formats.**

Our findings demonstrate that empathy-building interventions effectively raise awareness and empathy among participants. Expression-based interventions significantly increase empathy and awareness across age and gender groups, with no significant differences in impact between experiential and expression-based methods. Furthermore, we provide an easily adoptable web-based educational lab to integrate empathy and inclusive development topics into diverse curricula formats, accessible via our project website: https://all.rit.edu

***Index Terms*—Experiential Learning, Empathy-building, Inclusive Software**

I. Introduction

Empathy can result in the development of software that is more accessible, inclusive, and equitable [4, 21, 22, 23]. Although experiential empathy-building interventions have been studied in several non-computing fields, such as medicine [15], and to create tolerance in social situations [6], there is a lack of a comprehensive examination of these interventions in computing education [22, 23].

The benefits of experiential interventions to build empathy in computing education have been previously established [22, 23]. Our work contributes additional insights into best practices and the potential benefits of empathy-building interventions, which was achieved through a large in-person activity with 121 real-world participants. Two separate study groups were used, documenting a diverse range of demographics to support research questions. Both evaluations used a systematically designed pre-/post-survey summation assessment. This was achieved using empathy-building interventions *experiential* and *expression-based*.

The key takeaways from our study can be summarized as follows: I) Both experiential-based and expression-based interventions exhibit non-significant effects on increasing participant interest. II) However, these interventions have been shown to be highly effective in raising awareness and fostering empathy among participants. III) Specifically, the expression-based intervention demonstrates a significant difference in increasing empathy among various age groups, while also showing a notable increase in awareness among different gender groups. IV) No statistically significant differences were observed between the two empathy building interventions in terms of their impact on increasing interest, awareness, or empathy.

II. Related Work

Empathy-building interventions have been investigated in several non-computing fields, such as medicine [15], and to create tolerance in social situations [6]. These studies have revealed the advantages of empathy-building interventions, such as improving the understanding of a patient's perspective or improving communication practices with patients. These prior initiatives differ from our research because they do not specifically address the educational computing domain.

Several recent studies have focused on the importance of empathy in educational activities and explored effective implementation practices. Motahar *et al.* [17] conducted research to understand the reasons behind the shortcomings of existing methods in fostering "'design empathy". They emphasized the need for the HCI education sector to address the unresolved challenge of teaching empathetic design to students. Oleson *et al.* [18] introduced the Critique, Imagine, Design, Expand, Repeat (CIDER) assumption elicitation technique, which aims to teach inclusive design abilities. This method helps designers recognize and address biases by critically examining

assumptions about users. The researchers conducted an 11-week mixed method case study involving 40 undergraduates enrolled in an interaction design course. They supplemented this study with subsequent interviews. The results indicated that activities based on the CIDER technique had the potential to facilitate the identification of various design biases over time and encouraged students to reflect on unconscious biases related to users.

Recent studies have examined empathy in educational activities and have explored effective implementation strategies. Motahar *et al.* [17] conducted research to understand the shortcomings of existing methods in fostering "design empathy". They highlighted the need for the HCI education sector to address the ongoing challenge of teaching empathy design to students. Oleson *et al.* [18] introduced the Critique, Imagine, Design, Expand, Repeat (CIDER) assumption elicitation technique, which aims to aid in the development of inclusive design abilities. This method helps designers recognize and address biases by examining assumptions about users. The researchers conducted an 11-week mixed method case study involving 40 undergraduates enrolled in an interaction design course, supplemented with subsequent interviews.

In previous research [3], empathy building activities have been analyzed to understand people with disabilities. However, this work reveals that unfortunately, these acts of empathy can sometimes lead to a disconnect between people with disabilities and the intended support from designers. Crabb *et al.* [7] conducted a study with 197 participants that highlighted the existence of knowledge gaps among developers in creating accessible software, despite efforts to promote accessible design. In contrast, our developed educational material serves to address and close this knowledge gap among developers, empowering them to effectively create inclusive software.

Empathy building activities have been analyzed to gain insight into the experiences of people with disabilities [3]. This study underscores that while expressions of empathy are often employed in the design process, they can inadvertently contribute to a divergence between individuals with disabilities and the expected assistance from designers. A comprehensive investigation conducted by Crabb *et al.* [7] involving 197 participants sheds light on the persistent knowledge gaps among developers in the field of creating accessible software, despite efforts to advocate for accessible design. Unlike previous research, our educational materials are strategically designed to rectify and bridge these knowledge gaps among developers. The overarching goal is to empower developers to design and build inclusive software solutions in a proficient way.

III. Background

A. Educational Demand

Recent legislation has called for software that is more inclusive and unbiased. There is also a stated educational demand for easily adoptable interventions that will support the creation of more equitable software, such as software that is created with greater empathy [2]. Increased empathy is expected to result in the creation of software that is more inclusive, equitable, and impartial [16], while also having a positive impact on developers' careers [11]. Empathic skills are increasingly being seen as a necessity due to the increasingly globalized nature of society [16]. The demand for software with these attributes will continue to grow as more interactions and tasks are performed online [19]. Preliminary observations [21, 22] have demonstrated that the proposed work has the capability to directly contribute to accomplishing these goals.

B. Experience and Expression-Based Interventions

There are two primary forms of empathy interventions, *Experience-based* and *Expression-based* interventions. Experience-based interventions often allow the perceiver to encounter a scenario through the target's perspective using either a hands-on or theoretical activity. This form of intervention has traditionally been used to build empathy through a deeper understanding of the target's thoughts and feelings [25]. Expression-based interventions teach participants to recognize the internal states of the participant and respond appropriately.

This study uses two main types of empathy interventions: *Experience-based* and *Expression-based* interventions. Experience-based interventions typically involve the perceivers immersing themselves in a scenario from the target's perspective, either through hands-on experiences or theoretical activities. This approach, rooted in the building of empathy, aims to deepen the perception of the target's thoughts and emotions, thus building empathy [25]. In contrast, expression-based interventions focus on teaching participants to recognize the internal states of the participant and respond appropriately.

C. Experiential Education

Experiential learning is commonly used on many educational topics [26] and has been routinely demonstrated to have its benefits [12]. Experiential learning provides a complete learning experience for the student, in which they both understand the concept behind an idea and interactively learn about it [5]. Compared to alternative teaching approaches, such as lectures, experiential learning has been shown to be more engaging for students [14], and supports student retention of information [10].

D. Empathy-Building Interventions

Research shows that people do not often empathize with a particular target group due to a general unwillingness to empathize [27]. Fortunately, research also suggests that empathy can be developed, frequently through experiential activities [22, 23, 25]. An identified challenge in driving people to empathize is "avoidance motives" that drive people away from empathizing [13]. An example of avoidance motive is when people think that addressing empathy-created concerns will be too costly [20] or painful [8]. Therefore, when striving to create empathy, it is imperative to demonstrate how empathy will not hinder, but rather align with the project goals [20]. In general, there are at least three related but distinct subprocesses that comprise empathy [25]. "Mentalization" is the ability

to draw inferences about a target's feelings and thoughts. "Experience sharing" is when a person vicariously experiences another person's emotional state [9]. "Empathic concern" focuses on a perceiver's desire to alleviate the target's distress [1]. These relate to several forms of empathy, including *cognitive*, *emotional*, *affective*, and *somatic* [24]. Our work will primarily focus on cognitive empathy, as it is the form that is most amiable in an environment oriented to experiential computing.

IV. Created Publicly Accessible Educational Activity

(a) Experiential Activity Pre-Colorblindness Filter, where colors are easily discernible to users, emulating the experience of non colorblind users.

(b) Experiential Activity Post-Colorblindness Filter where color variations and text are difficult to discern due to "colorblindness' emulator, emulating the experience of colorblind users.

Fig. 1: Example Intervention Screenshots where Figure 1b contains indiscernible color shapes and texts due to simulated accessibility related issues compared to the more distinguishable accessible text shown in Figure 1a.

The created empathy-building educational activities serve a dual purpose: to address the investigated research objectives outlined in Section VI and to provide a valuable educational resource for the development of empathy for computing instructors. On the activity start page, participants are given the option to participate in one of two different versions of the exercise: I) an *experience-based* activity or II) an *expression-based* version. Both variations of the activity incorporate an abbreviated and modified version of a previously created intervention that has demonstrated its educational effectiveness [22, 23]. These formats and intervention labs are freely available for educational adoption. They form the basis for our research evaluation, as outlined in Section VI. All materials are publicly available on the project website and require only a browser for adoption. This facilitates the implementation of the material for brief educational empathy-building interventions by institutions and individual learners.

A. Experience-Based Intervention Activity

Participants have the opportunity to directly experience the challenges faced by colorblind individuals when navigating inaccessible apps. For a demonstration activity, participants participate in a game involving three different colored circles. Their objective is to click the circle when it becomes red and not to click the circle when it is either of the other two colors. Although this task is straightforward for those who can distinguish colors, it becomes nearly impossible for colorblind individuals who cannot distinguish colors.

In the subsequent round of the game, participants have another opportunity to play. However, in this round, the user interface emulates the experience of color blindness. The previously distinct colors of the circle transform into shades of gray that are nearly impossible to differentiate, emulating the challenges faced by colorblind individuals and introducing a substantial barrier to achieving a high score. This exercise serves as a powerful illustration of the importance of designing web interfaces with accessibility in mind, ensuring inclusivity for colorblind individuals.

B. Expression-Based Intervention Activity

The initial phase of the expression-based empathy-building activity mirrors the experiential activity, where participants have the same objective of playing the game that involves clicking on colored circles to achieve a high score. However, a new element is introduced to enhance their understanding by having participants watch a video of a colorblind individual attempting to play the same game. The participant witnesses the color blind person's frustration in playing the game. The activity seeks to cultivate an empathetic understanding of the experiences of colorblind individuals. Participants gain insight into the challenges faced by colorblind users when engaging with digital interfaces lacking proper accessibility features. This exercise encourages participants to consider the importance of creating inclusive interfaces that meet the needs of all users.

Figure 2 shows the video of a frustrated color user who uses the inaccessible application. By incorporating this expressive empathy building component, the activity aims to foster a more empathetic appreciation for the experiences of individuals with

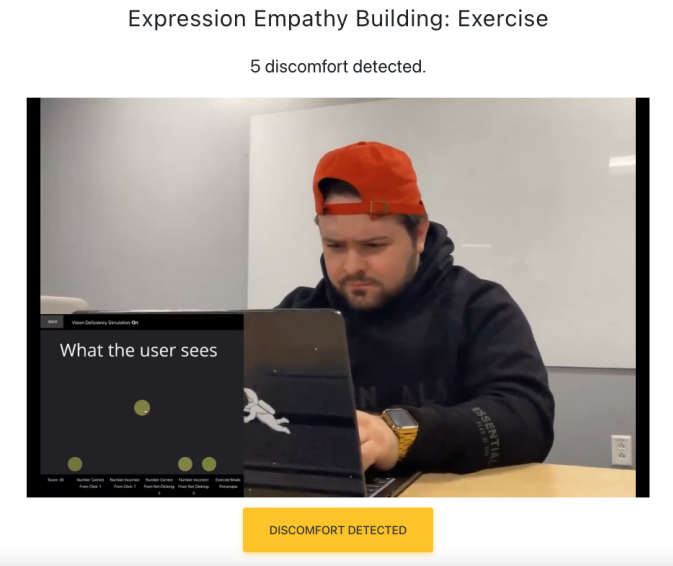

Fig. 2: Expressive intervention activity where the user is tasked with recognizing discomfort in a user experiencing the colorblindness activity.

color blindness. Participants gain valuable insights into the frustrations and difficulties faced by users with this condition when interacting with digital interfaces that lack appropriate accessibility features. This exercise encourages participants to reflect on the importance of designing inclusive interfaces that cater to the diverse needs of all users.

V. Study Design

Our research included data collected from 121 real-world participants during a local community event. IRB approval was secured before we began our study.

Our human study was conducted at RIT, a single-day event in which thousands of people from the local community visit the RIT campus and view a variety of scientific and educational venues, including robotics, software projects, and engineering activities. Visitors to the institute-wide event ranged in age from 0-51 years and were generally representative of the general population of the non-technology-oriented local population. Participants were randomly assigned one of the two interventions in order to evenly distribute the participants between the two interventions. With the help of several student workers, tables were set with laptops running the application variations described in Section IV. The participants were recruited by asking the visitors who passed by the tables if they would like to play a game that we had developed. Users were only provided vague details regarding the study (which very few of the participants likely reviewed or recalled information about our exhibit, since we were just one of the 100's of exhibitors at the event). In addition to basic technical and process-related questions, no guidance was provided to the participants. Since our IRB only covered users of at least 18 years of age, we did not retain the results of anyone younger than this age who participated in our study. We also excluded participants who did not complete the pre-post survey (Section V-A). This resulted in a total of 121 people participating in our study. As our IRB approval only covered users aged 18 and above, we did not retain the results of participants below this age who participated in our study. Furthermore, participants who did not complete the pre-post survey entirely were excluded (Section V-A). This resulted in a total of 121 individuals participating in our study.

A. Data Collected

Fig. 3: Data collection process for in-person user study

Participants were asked to fill out pre-/post-surveys, once before engaging with the application, and then again after using the application. Access to the application was contingent on completing the initial pre-survey, and participants were only allowed to proceed to the post-survey after completing the application. The initial section of the survey collected demographic information, including age and self-identified gender. This survey component measured the participant's sentiment regarding empathy intervention. Given the brevity of the large-scale in-person study, we maintained a small number of pre-/post- survey questions. The pre-/post-survey and high-level results are shown below:

Pre-Survey

1) **What Is Your Age?**
   a) 0-17 years `(n/a: excluded due to IRB)`
   b) 18-29 years `(60.33%)`
   c) 30-50 years `(18.19%)`
   d) 51+ years `(21.48%)`
2) **What Is Your Gender?**
   a) Male `(54.54%)`
   b) Female `(41.33%)`
   c) Other `(4.13%)`
3) **What demographic do you most closely identify with?**
   a) White `(61.98%)`
   b) Asian or P.I. `(22.31%)`
   c) Other `(15.71%)`

Pre-/Post-Survey

In both the pre- and post-survey phases, we solicit responses on a Likert scale to the subsequent questions from the participants.

1) **How interested are you in the topic of software inclusiveness?**
2) **I can identify with the challenges posed by inaccessible software.**
3) **I get upset when I see inequitable software.**
4) **I have empathy for people that are challenged by issues of inclusion.**
5) **I find that it is difficult to understand the impacts of non-inclusive software.**

1) Overview of Collected Data

Figure 4 and Figure 5 show an overview of group-wise comparison of collected data. In addition to the collected data that has been reported in Section V-A, we will next provide an additional breakdown of the collected data. Table I and Table II provide an overview of our collected data.

Summary of Responds for Pre-survey Questions

Qs	S. Disagree[1,2]	Disagree[1]	Neutral[1]	Agree[1]	S. Agree[1,2]
Q1	1 (0.8%)	1 (0.8%)	26 (21%)	52 (43%)	41 (34%)
Q2	9 (7.4%)	20 (17%)	36 (30%)	44 (36%)	12 (9.9%)
Q3	2 (1.7%)	4 (3.3%)	31 (26%)	66 (55%)	18 (15%)
Q4	1 (0.8%)	0 (0%)	6 (5.0%)	52 (43%)	62 (51%)
Q5	8 (6.6%)	36 (30%)	26 (21%)	36 (30%)	15 (12%)

[1]n (%)
[2]S. is abbr for Strongly

Summary of Responds for Post-survey Questions

Qs	S. Disagree[1,2]	Disagree[1]	Neutral[1]	Agree[1]	S. Agree[1,2]
Q1	1 (0.8%)	4 (3.3%)	13 (11%)	51 (42%)	52 (43%)
Q2	4 (3.3%)	11 (9.1%)	15 (12%)	48 (40%)	43 (36%)
Q3	1 (0.8%)	1 (0.8%)	17 (14%)	67 (55%)	35 (29%)
Q4	0 (0%)	0 (0%)	4 (3.3%)	41 (34%)	76 (63%)
Q5	23 (19%)	42 (35%)	14 (12%)	28 (23%)	14 (12%)

[1]n (%)
[2]S. is abbr for Strongly

TABLE I: Participant Self-identified Gender and Ethnicity

	Self-Identified Gender			
	Male	Female	Other	Total
White	44	29	2	75
Asian or P.I.	13	12	2	27
Other	9	9	1	19
Total	66	50	5	121

TABLE II: Average score for selected post-survey questions broken down by demographics

Gender	Q1	Q2	Q3	Q4	Q5	Q6
M	4.20	4.12	4.11	4.58	4.63	2.69
F	4.27	3.72	4.10	4.60	4.66	2.83
Other	4.20	3.90	4.10	4.60	4.90	2.50
Ethnicity	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**
White	4.22	4.02	4.09	4.52	4.61	2.61
Asian or P.I.	4.23	3.80	4.19	4.61	4.65	3.03
Others	4.25	3.87	4.04	4.79	4.83	2.75
Total	4.23	3.95	4.10	4.59	4.66	2.73

TABLE III: Testing results of effectiveness in terms of increasing interest for both interventions: both interventions are not significant in increasing interest.

Group	W-stats	Alternative	P-Value
Experiential-Based	92.0	two.sided	0.042
Expression-Based	162.0	two.sided	0.305

VI. Evaluation

We primarily leverage two non-parametric tests in our analysis: the Wilcoxon signed rank test (which evaluates the null hypothesis that two samples originate from identical distributions) and the Kruskal-Wallis test (which examines the null hypothesis that the population median is the same for all the groups). We choose these tests because they don't rely on assumptions about the data distribution as parametric tests do, thus providing more robust testing results. We follow the convention by choosing 5% as our significant level.

RQ1. *Are empathy-building interventions effective in terms of increasing interest?*

To answer this research question, we performed a detailed analysis to verify the significance of both interventions in changing the interests of the participants. For this, we consider the survey question: Q_1 : *How interested are you in the topic of software inclusion?* To analyze the statistical significance. We choose to use non-parametric tests, such as the Wilcoxon test, in our analysis, since these tests do not make overly strong assumptions regarding our data. The results of the Wilcoxon test for both groups of interventions are provided in Table III.

The p-value of the experiential intervention in Table III shows significance in the border range *i.e.,* close to 5%. Due to the small sample size of the experiential-based group, we test whether adding or removing one or two participants matters or not. We found a notable difference in the p-value when removing only one participant. This suggests that we need quite strong evidence to maintain the significance of this group. Hence, we conclude that it is not significant with a p-value of 0.042. For expression-based intervention, it has a large p-value of 0.305, which is greater than our chosen significant level, which means there is no increase in participant interest. Furthermore, we demonstrate the transition Likert score plot for experiential-based intervention in Figure 6 and for expression-based intervention in Figure 7. In both plots, most of the Likert scores are concentrated in the scores of 4 and 5 and have very few score changes from pre- to post-survey. Therefore, there is not enough impact to provide significance in increasing interest via both interventions.

Discussion: Additional data samples typically help detect subtle effects in data. The fact that the experiential-based group with smaller sample size shows significant effects, when comparing with our chosen significant level, while the other group doesn't, concerns us. As observed in RQ5, we couldn't find any significant difference in survey response between the two forms of interventions. This suggests that the significance we found in the test for experiential-based group is merely coincidence. Furthermore, as mentioned earlier, some data samples in the experiential-based group appear to have large leverage for the testing result, this also coincides with the conclusion that the significant testing result is a coincidence.

In summary, the key findings of this research question are:

- Both utilized experiential and expression-based interventions have no significance in increasing participant interest.
- The experiential-based group has a relatively borderline significance, but it is too data sensitive due to its small size.
- There is not enough evidence to suggest that either of the empathy-building inventions is effective in encouraging interest among participants.

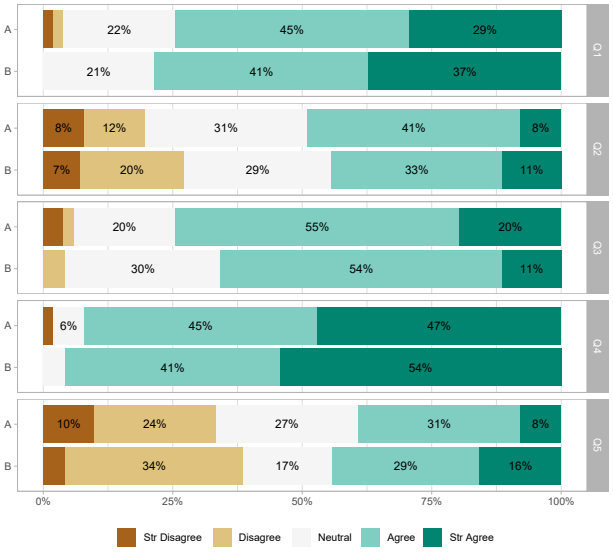

Fig. 4: Group-wise comparison of collected response to pre-survey questions

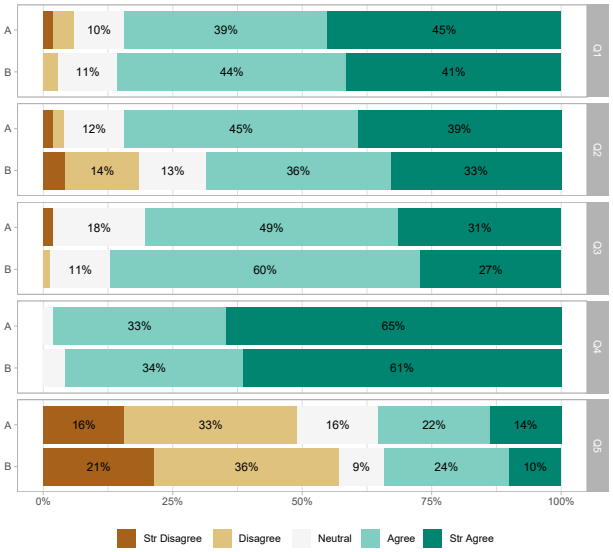

Fig. 5: Group-wise comparison of collected response to post-survey questions

RQ2. *Are empathy-building interventions effective in terms of increasing the awareness of non-inclusive software?*

To answer this research question, we use two survey questions (Q_2 and Q_5). The Q_2 is about identifying the challenges, and Q_5 is about the difficulty of understanding the impact due to inaccessible software. As mentioned, these two questions Q_2 and Q_5 have opposite meanings to each other, and hence we subtracted their Likert score (Q_2-Q_5) for both pre- and post-survey as the overall proxy measure for awareness of the participants. We used the paired Wilcoxon test to check if there is a significant difference. The test results are shown in Table IV.

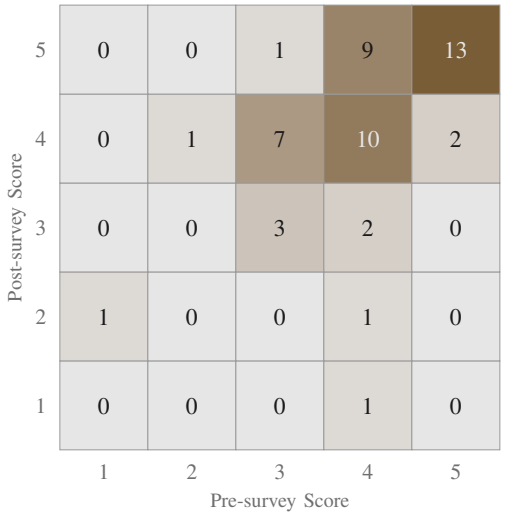

Fig. 6: Experiential-based group's transition score plot for interest, considering pre- and post-survey Likert score for survey question: Q_1.

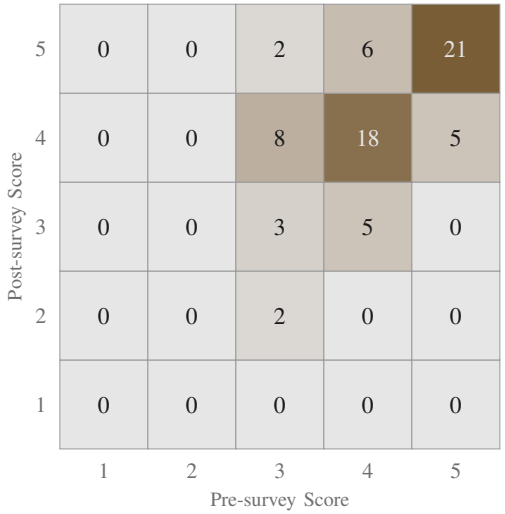

Fig. 7: Expression-based group's transition score plot for interest considering survey question: Q_1.

TABLE IV: Testing results of effectiveness in terms of increasing awareness for both interventions: both interventions increase awareness significantly.

Group	W-stats	Alternative	P-Value
Experiential-Based	52.5	two.sided	4.06×10^{-6}
Expression-Based	182.0	two.sided	4.72×10^{-6}

Table IV indicates that the tests for both interventions have p-values that are much smaller than our chosen significant level. We can reject the null hypothesis, and conclude there's no significant difference between pre- and post-survey scores.

We observed an **increase of** 1.078**, and** 1.086 **in score statistic** (Q_2-Q_5) for experiential- and expression-based group, respectively.

Discussion: Our analysis found that both empathy-building interventions are effective in terms of raising participants' awareness, in the sense that we receive a more positive response regarding identifying the difficulty of non-inclusive software after the interventions. For adopters, these results provide preliminary evidence that both empathy-building intervention formats can be effective.

In summary, the key findings of this research question are:

- Both utilized experiential-based and expression-based interventions show strong significance in increasing awareness among the participants.

RQ3. *Are empathy-building interventions effective in terms of increasing empathy?*

To address this question, we considered two survey questions Q_3 and Q_4. The survey question Q_3 concerns getting upset when there is inequitable software, and the Q_4 involves having empathy for people who are challenged by issues of inclusion. Both questions measure the level of empathy among participants, and therefore we aggregated their pre- and post-survey Likert scores as $Q_3 + Q_4$.

TABLE V: Testing results for the effects of increasing empathy-building for two different empathy-building interventions: Both experiential and expression-based interventions are effective for increasing participant empathy.

Group	W-stats	Alternative	P-Value
Experiential-Based	59.0	two.sided	$9.19e^{-4}$
Expression-Based	182.0	two.sided	$1.78e^{-4}$

From Table V, we observe that the p-values of the Wilcoxon test for both experiential and expression-based intervention are much lower than the chosen significant level, *i.e.,* 5%. Hence, we can conclude that both interventions are effective in terms of raising participants' empathy. Furthermore, we observed that on average there is **an increase of** 0.5098**, and** 0.4714 **in score statistics** for the experiential and expression-based intervention.

Discussion: We utilize two survey questions ($Q_3 + Q_4$) to reduce the variation in the response. According to our analysis, both forms of empathy-building interventions are effective in terms of raising participants' empathy. However, we also notice the increases in score statistics are rather small, it is not clear whether these differences can lead to any practical difference in terms of education outcomes. Taking this into account, the strength and duration of the intervention's effects remain unexplored.

In summary, the key findings of this research question are as follows.

- Experiential-based and expression-based interventions significantly both increase the empathy of the participants.
- Both interventions are useful to understand participants upset with inaccessible software.
- Similarly, both interventions encourage participants to provide empathy towards issues of inclusion.

RQ4. *Do empathy-building interventions uniformly impact participants from all demographics (e.g. gender, ethnicity, etc) equally?*

We performed evaluations across different demographics (*e.g.,* gender, age, etc.) to investigate whether empathy building interventions uniformly affected participants. Since only awareness and empathy effects are significant based on our previous analysis, we next tested these effects across demographics.

TABLE VI: Testing results for effects of empathy-creating intervention across different demographics: for brevity, we only present the significant results. We found that there might be a significant difference among different age groups and gender groups for empathy and awareness, respectively

Aspect	Demographic	K-stats	P-Value
Empathy	Age	7.33	0.026
Awareness	Gender	4.239	0.04

To analyze the impact of the intervention on various demographics, we employ the Kruskal-Wallis test on the changes in the survey score within each group in relation to individual demographic factors. We are unable to detect any significant difference across different demographics for participants in the experiential group. However, we detected a significant change in the expression-based group against age and gender (Table VI). Specifically, we detect a difference in empathy across different age groups, and a difference in awareness in different gender groups. From Figure 8, we observe a slight shift to the right in the score difference for male participants.

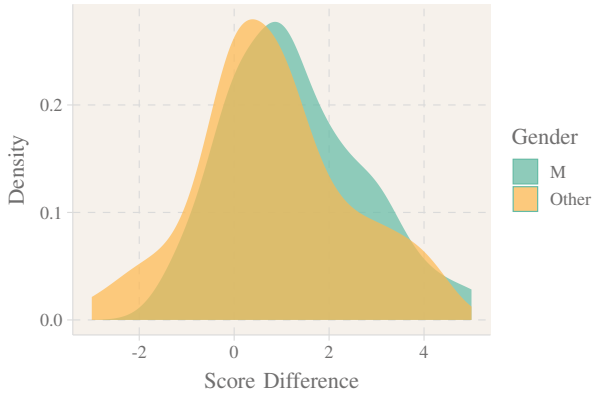

Fig. 8: Score difference for awareness between different genders: The density of score difference for male participants is slight towards the right. This shows that different gender groups may respond differently towards the interventions.

Discussion: While we observed statistically significant results solely within the expression-based group, this does not nec-

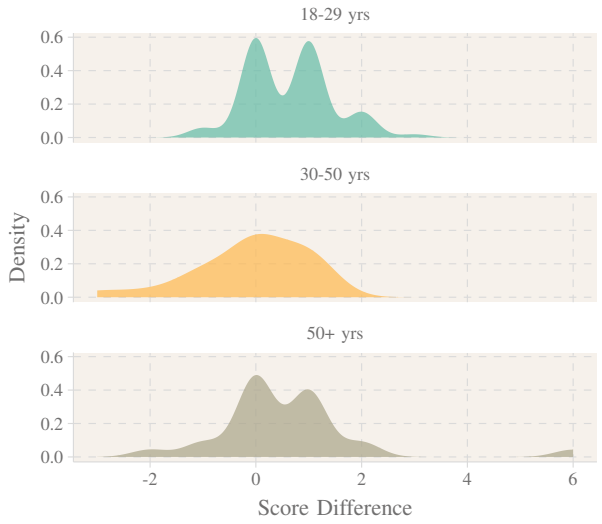

Fig. 9: Score difference for empathy among different ages: The density of score difference for participants of age 30-50 is slight towards the left. This may indicate that certain age groups are less susceptible to empathy creating interventions, which may warrant further investigation.

essarily imply two forms of interventions will have different effects (see RQ5). Our expression-based group has a larger number of participants, making it likely to detect subtle effects. Furthermore, we suggest considering our results as an indication for future experiments, not as the definitive answer to the research question, given the small size of the participants and the unreliable nature of the self-evaluated survey response. Moreover, from Figure 9, we observe that the scores for the participants in the age group of 30-50 are more concentrated around zero. Although under 5% the significant levels are significant, they still have relatively large p-values. We can also visually notice this from Figure 8 and Figure 9, where there are relatively moderate shifts in density. This indicates that individuals from different gender and age groups could exhibit different responses to interventions in awareness and empathy. More experiments are required before a more definitive conclusion can be drawn.

In summary, the key findings of this research question are:

- For the expression-based group, we found that there is a significant in the empathy effect among different age groups.
- For the expression-based group, we found that there is a significant difference in terms of raising awareness among different gender groups.
- We found no significant differences between different demographics for the experiential-based group.

RQ5. *What form of empathy-building interventions (e.g. expression-based and/or experiential-based) is more effective?*

To investigate whether there is a difference between experiential and expression-based empathy-creating interventions, we perform Wilcoxon tests. We tested against three different aspects that we studied in this paper, namely raising participants' interests, empathy, and awareness. The test results can be found in Table VII. Since the p-values for all tests are larger than our chosen significant level, we cannot reject the null hypothesis that there is no significant difference between two forms of empathy-building interventions.

TABLE VII: Results for different aspects between experiential- and expression-based groups. Given the large p-values, we cannot reject the null hypothesis that this no difference between two forms of empathy-creating interventions; *i.e.,* we found no evidence that there is a difference between two forms of empathy interventions.

Aspect	W-stats	Alternative	P-Value
interest	2,057	two sided	0.112
awareness	1,579.5	two sided	0.253
empathy	1,747	two sided	0.833

Discussion: According to our analysis on the survey responses, we are unable to detect any significant difference between the two types of empathy-building interventions. Nonetheless, it would be imprudent to ignore the difference between these two interventions. Further experiments should focus on a more thorough comparison between these two forms of inventions.

In summary, the key findings of this research question are:

- We found no difference in the effectiveness between experiential-based and expression-based interventions.

VII. Conclusion

This research contributes to the existing body of knowledge on the impact of empathy in the context of experiential and expression-based computing education. The primary findings include the following. I) Both experiential-based and expression-based interventions, although employed, do not demonstrate statistical significance in increasing interest; II) However, these interventions prove highly effective in increasing awareness and instilling empathy among participants; III) The expression-based intervention exhibits a significant impact in improving empathy among diverse age groups and increasing awareness among different gender groups; and IV) No statistically significant disparity is observed between the two empathy-building interventions with respect to increasing interest, awareness, and empathy. This work additionally provides a hosted, easy-to-adopt educational resource that may be utilized in various computing and non-computing classrooms. Complete educational material is available: https://all.rit.edu

Acknowledgements

This material is based upon work supported by the NSF under grant #1825023, #2145010, and #2111152.

References

[1] C. D. Batson. These things called empathy: eight related but distinct phenomena. *MIT press*, 2009.

[2] R. K. Bellamy, K. Dey, M. Hind, S. C. Hoffman, S. Houde, K. Kannan, P. Lohia, S. Mehta, A. Mojsilovic, and S. Nagar. Think your artificial intelligence software is fair? think again. *IEEE Software*, 36(4):76–80, 2019.

[3] C. L. Bennett and D. K. Rosner. The promise of empathy: Design, disability, and knowing the "other". In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–13, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359702.

[4] C. L. Bennett and D. K. Rosner. The promise of empathy: Design, disability, and knowing the" other". In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–13, 2019.

[5] D. Boud, R. Keogh, and D. Walker. *Reflection: Turning experience into learning*. Routledge, 2013.

[6] G. L. Clore and K. M. Jeffery. Emotional role playing, attitude change, and attraction toward a disabled person. *Journal of personality and social psychology*, 1972.

[7] M. Crabb, M. Heron, R. Jones, M. Armstrong, H. Reid, and A. Wilson. Developing accessible services: Understanding current knowledge and areas for future support. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, New York, NY, USA, 2019. Association for Computing Machinery.

[8] M. H. Davis, K. V. Mitchell, J. A. Hall, J. Lothert, T. Snapp, and M. Meyer. Empathy, expectations, and situational preferences: Personality influences on the decision to participate in volunteer helping behaviors. *Journal of personality*, 67(3):469–503, 1999.

[9] E. Hatfield, J. T. Cacioppo, and R. L. Rapson. Emotional contagion. *Current directions in psychological science*, 2(3):96–100, 1993.

[10] K. Hawtrey. Using experiential learning techniques. *The Journal of Economic Education*, 38(2):143–152, 2007.

[11] P. A. Hecker. Successful consulting engineering: a lifetime of learning. *Journal of management in engineering*, 13(6):62–65, 1997.

[12] A. Y. Kolb and D. A. Kolb. Learning styles and learning spaces: Enhancing experiential learning in higher education. *Academy of management learning & education*, 4 (2):193–212, 2005.

[13] Z. Kunda. The case for motivated reasoning. *Psychological bulletin*, 108(3):480, 1990.

[14] L. Laird and Y. Yang. Engaging software estimation education using legos: A case study. In *Proceedings of the 38th International Conference on Software Engineering Companion*, ICSE '16, pages 511–517, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-4205-6.

[15] T. Levett-Jones, R. Cant, and S. Lapkin. A systematic review of the effectiveness of empathy education for undergraduate nursing students. *Nurse education today*, 2019.

[16] M. Levy and I. Hadar. The importance of empathy for analyzing privacy requirements. In *2018 IEEE 5th International Workshop on Evolving Security & Privacy Requirements Engineering (ESPRE)*. IEEE, 2018.

[17] T. Motahar, N. Brown, E. S. Wiese, and J. Wiese. Building "design empathy" for people with disabilities: an unsolved challenge in hci education. In *Proceedings of the 5th Annual Symposium on HCI Education*, pages 68–71, 2023.

[18] A. Oleson, M. Solomon, C. Perdriau, and A. Ko. Teaching inclusive design skills with the cider assumption elicitation technique. *ACM Trans. Comput.-Hum. Interact.*, 30(1), mar 2023. ISSN 1073-0516.

[19] N. Pettijohn. *Can We Just Work From Home Forever?*, May 2020 (accessed July 3, 2020). URL https://www.forbes.com/sites/nathanpettijohn/2020/05/20/can-we-just-work-from-home-forever/#28a779e43d95.

[20] L. L. Shaw, C. D. Batson, and R. M. Todd. Empathy avoidance: Forestalling feeling for another in order to escape the motivational consequences. *Journal of Personality and Social Psychology*, 67(5):879, 1994.

[21] W. Shi, S. Khan, Y. El-Glaly, S. Malachowsky, Q. Yu, and D. E. Krutz. Experiential learning in computing accessibility education. In *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering: Companion Proceedings*, ICSE '20, page 250–251, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450371223.

[22] W. Shi, S. Malachowsky, Y. El-Glaly, Q. Yu, and D. E. Krutz. Presenting and evaluating the impact of experiential learning in computing accessibility education. In *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering: Software Engineering Education and Training*, ICSE-SEET '20, page 49–60, New York, NY, USA, 2020. Association for Computing Machinery. Distinguished Paper Award.

[23] W. Shi, H. Moses, Q. Yu, S. Malachowsky, and D. E. Krutz. All: Supporting experiential accessibility education and inclusive software development. *ACM Trans. Softw. Eng. Methodol.*, Sept 2023. ISSN 1049-331X.

[24] A. Smith. Cognitive empathy and emotional empathy in human behavior and evolution. *The Psychological Record*, 56(1):3–21, 2006.

[25] E. Weisz and J. Zaki. Empathy building interventions: A review of existing work and suggestions for future directions. *The Oxford handbook of compassion science*, pages 205–217, 2017.

[26] S. Wurdinger and P. Allison. Faculty perceptions and use of experiential learning in higher education. *Journal of e-learning and Knowledge Society*, 13(1), 2017.

[27] J. Zaki and M. Cikara. Addressing empathic failures. *Current Directions in Psychological Science*, 24 (6), 2015.