

Bridging Perspectives: Unveiling Racial Dynamics in Remote Pair Programming Communication

Shandler A. Mason , Sanket Nain , and Sandeep Kaur Kuttal 

North Carolina State University, Raleigh NC 27606, USA
{samason4, snain, skuttal}@ncsu.edu

Abstract. Remote pair programming is a collaborative method in software development, yet a research gap exists concerning how race influences communication within these interactions. Our study is dedicated to unveiling communication dynamics between pairs, with a specific focus on Black developers, as their representation diminishes in the university-to-industry pipeline, where White developers constitute the majority. We recruited 24 US-based developers, both Black and White, forming 12 pairs. These pairs were equally distributed among same- and mixed-race compositions, between men and women. Employing the think-aloud method, each pair engaged in collaborative programming tasks, followed by independent retrospective interviews. Our results revealed that while mixed-race pairs were actively engaged in the task, they displayed the lowest frequency of happiness facial expressions, highlighting the intricate racial dynamics in communication. Our overarching goal is to gain insights into how communication is influenced by the perceived racial background of partners. This exploration is imperative for shaping the development of inclusive and empathetic collaborative software tools.

Keywords: Pair Programming · Race · Developers · Communication.

1 Introduction

Effective collaboration is crucial for software development teams to produce high-quality products. A well-established, agile approach to collaboration is pair programming, where two developers work together to solve programming challenges. One developer assumes the role of the driver by typing code while the other takes on the role of the navigator by contributing ideas [83]. However, fostering seamless collaboration in pair programming poses greater challenges when the software team comprises individuals from diverse backgrounds [44].

Our study is dedicated to unveiling communication dynamics between pairs, with a specific focus on Black developers, as their representation diminishes in the university (20%) to industry (6%) pipeline [64]. In this context, we selected Black and White developers based on a complex history of interracial interactions, in the United States, which has led to differences in collaborative styles [73]. Beginning with the era of slavery, this intricate history has been marked by a legacy of racial injustices [60]. Today, disparities persist in wealth [1], education [8], and the penal system [32].

Currently, there is a gap in existing research, as no study has investigated how the perceived race of developers influences communication elements such as interruptions, non-verbal cues, and prompts during pair programming.

Pair programming requires collaboration among individuals from diverse backgrounds to solve tasks. The human brain harbors implicit biases which tend to shape stereotypes within the initial moments of meeting an individual [46]. Racial stereotypes may impact communication during collaboration. To investigate how race influences communication in pair programming among both same- and mixed-race pairs, we devised three research questions:

- **RQ1: How do participants’ interruption frequencies differ between same- and mixed-race pairs during pair programming?**
- **RQ2: How do the non-verbal cues of participants in same-race pairs differ from those in mixed-race pairs during pair programming?**
- **RQ3: How do same- and mixed-race pairs exhibit different communication prompts during pair programming?**

Our study aimed to address these research questions by employing think-aloud lab studies and retrospective interviews. We created same- and mixed-race, gender-balanced pairs with 24 Black and White professional developers. We observed pairs communication as they utilized the think-aloud method to solve a programming challenge. Afterwards, we retrospectively interviewed participants to gain insight into their feelings and experience with their partner.

2 Background

2.1 Pair Programming

Pair programming is an agile methodology used by both students and professionals [33,76]. Pair programming involves two key roles: the driver and the navigator. The driver is responsible for writing code, managing the keyboard, and task execution, while the navigator focuses on error checking and strategy development [28,54,59]. Pairs can interchange roles throughout the collaborative process [84].

Remote pair programming mirrors the advantages and characteristics of in-person pair programming [35,39,47]. It utilizes software technologies to facilitate collaboration between developers across various geographical locations [81]. Given the global nature of the software development industry and the impact of COVID-19, remote work has become increasingly common in this field [4].

Pair programming offers various benefits, including heightened productivity, code quality, engagement, and communication [11,20,34,55,79]. Additionally, pair programming boosts developers’ self-efficacy, fostering resilience in solving coding challenges [7,56,58].

However, pair programming may also present difficulties for partners, such as limiting individual discovery [82], diverting pairs from their goal [51], and causing tension [38] which negatively effects pair efficiency. Other challenges may arise

such as the navigator struggling to make significant contributions to the task [61] and partners deviating from their assigned roles [16].

During pair programming communication, both professional and student pairs predominantly dedicate their time to discussing the problem [17] and implementing the code, with the navigator taking an active role in contributing to the discussion [77]. Pairs with mixed personalities exhibit enhanced communication [72], however higher conversational skill levels does not guarantee increased pair productivity [22].

2.2 Racial Dynamics in the United States

Race is a socially constructed label with a diverse range of accepted definitions [27]. Our study characterizes race as a multifaceted concept that includes an individual’s personal identity, beliefs, environment, survey responses, and physical attributes [69]. Racial classifications like “Black or African American” and “White” are labels recognized in modern US society [43].

Slavery originated the complex relationship between Black and White individuals, in the US [13]. Legal disparities persisted through discriminatory practices including segregation laws [48] and socioeconomic limitations [12]. Societal inequalities are still prevalent today with lower college graduation rates for Black adults (22.5%) compared to their White (36%) counterparts [2], and a lower loan approval rate for Black borrowers (48%) compared to White borrowers (77%) [80].

Management research pertaining to workplace collaboration amongst diverse groups has found that a negative stereotype of individuals from different racial backgrounds can adversely impact group communication, ultimately limiting productivity [19]. In the workplace, Black individuals may face challenges such as diminished trust [63], stereotype threat [78], and concerns about discrimination [45]. These factors can contribute to heightened anxiety [14] and depression [42], potentially lowering overall workplace performance [30].

2.3 The Tech Industry Pipeline Problem

Complex racial dynamics, in the US, contributes to disparities in education that impact the tech industry. The under-representation of Black developers in the tech industry manifests early in education, spanning from K-12 through college and extending into the industry. Among fourth graders in 2019, Black students encountered an 18–25 point deficit in mathematics scores compared to White students [18]. At the high school level, Black students face less access to advanced STEM courses [8]. This disparity persists into higher education, in 2019, Black individuals earned only 9% of STEM Bachelor’s degrees, despite constituting roughly 13% of the US population [18]. Within the tech industry, Black individuals hold just 7% of all computing-related occupations [65], while simultaneously facing wage disparities [25]. Strategies for fostering diversity, equity, and inclusion, at the educational and industrial levels, are essential to addressing the pipeline problem [6,52].

3 Methodology

3.1 Participants

We distributed a study flyer and description crafted without reference to race or gender to recruit participants through snowball sampling and online social media platforms including LinkedIn, Facebook, and Slack.

We employed a systematic screening questionnaire to select 24 developers. Our study participants were professionals or PhD students in a computing-related field, all possessing a minimum of a Bachelor’s degree. Our study included participants who self-identified as either man or woman, Black or White, and who were born and spent their formative years in the US. The age range of participants spanned from 18 to over 40 years old. (21/24) held professional experience within the computing sector. Participants self-reported proficiency in programming languages including Java, Python, or C#.

We deliberately created 2 Black-Black, 2 White-White, and 2 Black-White pairs for men and women. We labeled each pair (P#-X#Y#), P# for pair number, X# for gender (M-Man, W-Woman), and Y# for race (B-Black, Wh-White) of the first or second participant.

3.2 Study Design

We used Zoom, a teleconferencing tool [5], to create a controlled study environment. We requested participants to maintain both their video and audio active throughout the study. We guided participants through a series of steps, including completing the consent form, the pre self-efficacy survey, and three instructive video tutorials covering pair programming, test-driven development [9], and think-aloud [53]. Additionally, participants observed a live demonstration of Replit [24], the collaborative development environment we used in our study. Participants were directed to apply these methodologies to both programming tasks.

To enhance pair jelling [50], participants engaged in a 10 minute ‘Simple Task’ focused on checking password validation. For the ‘Main Task’, participants programmed the basic functionality of a Tic-Tac-Toe game for 40 minutes. Participants responded to post self-efficacy and pair programming questionnaires. Participants completed individual, semi-structured retrospective interviews held by a racially-matching interviewer to increase comfort levels [10]. We compensated participants with a \$30 Amazon gift card in acknowledgment of their completion of our 1.5-2 hour study.

3.3 Data Analysis

To address our research questions (RQs), we employed a combination of quantitative and qualitative analyses. For RQ1, we utilized `Pyannote.audio` [15] to examine videos recorded through Zoom’s integrated functionalities. This analysis focused on detecting overlapped speech as an indicator of interruptions,

with manual verification performed by our researchers. In addressing RQ2, the `Py-Feat` [21] library was employed to evaluate participants’ facial expressions. The detector model assessed 6 emotions on a scale of 0 to 1, every 1,000 frames, identifying the dominant emotion for each interval. Our quantitative analysis of RQ1 and RQ2 involved descriptive statistical analysis such as mean and standard deviation.

For our qualitative analysis pertaining to RQ3, we utilized Zoom’s integrated functionalities to capture participant utterances. These recordings were transcribed and segmented into individual utterances, with two researchers manually examining inaccuracies in the transcriptions. Each utterance was labeled with the corresponding timestamp, the speaker’s label, their self-reported race and gender, and their pair programming role. Table 3 specifies the code set we used to code utterances. For RQ1, RQ2, and RQ3, we implemented thematic analysis to categorize and iteratively open-code developers’ retrospective interview responses, aiming to identify relevant patterns [70]. We used the Jaccard measure to assess inter-rater reliability [41]. First, 20% of the transcripts were coded independently by two researchers, who reached agreement on 85% of the coded data. Then, for independent coding, the two researchers divided the remaining transcripts.

3.4 Limitations

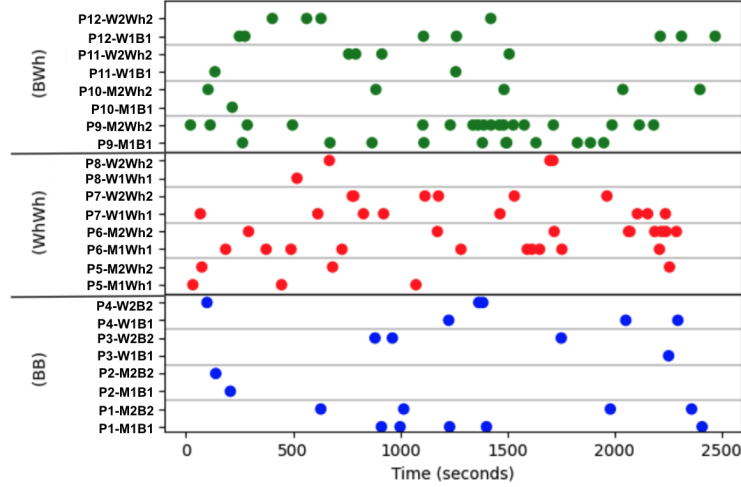
We acknowledge that our study is subject to various limitations. The sample size, consisting of 8 same-race and 4 mixed-race pairs, may be considered small for making broad assumptions or performing quantitative comparisons. However, our study marks the initial stride into understanding race’s possible influence on communication among professional developers during pair programming. Additionally, our focus was constrained to men and women, based on participants’ self-reported gender, limiting the diversity of gender representation. Recruitment challenges, particularly in reaching individuals from marginalized groups such as Black men, Black women, and White women, extended the study duration to 6 months. It is important to emphasize that race is not the only factor contributing to our results, as other factors, including skill level, unique personalities, and awareness of diversity may have played a role. The programming task in our study, implementing a Tic-Tac-Toe game, is simple in comparison to the complexity of real-world programming tasks.

4 Results

4.1 RQ1: How do participants’ interruption frequencies differ between same- and mixed-race pairs during pair programming?

An interruption is when one person begins to speak while another person’s turn is ongoing [3]. Interruptions are associated with collaborative dominance [85], conversational attentiveness [57], and engagement during the task [74]. Our quantitative analysis compared interruptions overtime (refer to Fig. 1) and interruption

Fig. 1. Participants interruptions overtime (blue) same-race (BB) vs. (red) same-race (WhWh) vs. (green) mixed-race (BWh) pairs



frequencies, using descriptive statistical analyses (refer to Table 1), across three groups (4 - same-race (BB), 4 - same-race (WhWh), and 4 - mixed-race (BWh) pairs).

Mixed-race (BWh) pairs interrupted one another more frequently than same-race (BB) and (WhWh) pairs. Mixed-race (BWh) pairs encountered the most interruptions on average, but the total frequency for each pair was dispersed (Mean = 12.5; SD = 9.95), refer to Table 1. Mixed-race (BWh) pairs generally interrupted each other during the beginning (0-500 seconds) and middle (1000-1500 seconds) of the task, refer to green in Fig. 1. (4/8) participants in mixed-race (BWh) pairs reported feeling interrupted by their partners, which had a negative impact on their workflow. P12-W1B1 discussed the pairs frequent interruptions, “*[I felt interrupted] probably every 5 minutes. It felt kind of frequent and interruptions [were] not just me saying stuff, but [when I was] thinking about things.*” This could indicate that mixed-race (BWh) pairs were engaged and wanted to be involved in task decisions [74].

Same-race (BB) pairs exhibited the lowest frequency of interruptions between partners and the standard deviation suggests that the frequencies were relatively consistent across the pairs (Mean = 5.75; SD = 3.86), refer to Table 1. Same-race (BB) pairs typically interrupted each other during the middle (1000-1500 seconds) and end (2000-2500 seconds) of the task, refer to blue in Fig. 1. (6/8) same-race (BB) pair participants reported not experiencing interruptions from their partners. When asked about his perception of interruptions, P2-M1B1 stated, “*No, [I did not feel interrupted by my partner]. He was confident, which I have no problem with. If you feel like you could tackle the answer I don’t have a personal problem with it.*” This suggests that one partner in same-race (BB) pairs was not as attentive during the tasks [57].

In mixed-race (BWh) pairs, the driver interrupted more frequently, while in same-race (BB) and (WhWh) pairs, the navigator initiated more interruptions. In (3/4) mixed-race (BWh) pairs, the driver exhibited more interruptions, whereas in (5/8) same-race (BB) and (WhWh) pairs, the navigator had a higher frequency of interruptions. P4-W2B2 (the navigator in a same-race (BB) pair) validated the observed pattern, “*The interruptions [didn’t work within the partnership] and I’m saying that as somebody that did most of the interruptions. I wish there was an easier way to do that. I feel like she was walking through the process out loud and then I would be like, ‘Oh, I have an idea.’ I don’t think that worked very well.*” The observed pattern in same-race (BB) and (WhWh) pairs aligns with a previous pair programming study, suggesting that the navigator typically initiates more interruptions [50], indicating a conversational dominance in this role.

Table 1. Mean and standard deviation (SD) of interruption frequencies for same-race (BB) vs. same-race (WhWh) vs. mixed-race (BWh) pairs

	Same-Race (BB)	Same-Race (WhWh)	Mixed-Race (BWh)
Mean	5.75	10.25	12.5
SD	3.86	5.85	9.95

4.2 RQ2: How do the non-verbal cues of participants in same-race pairs differ from those in mixed-race pairs during pair programming?

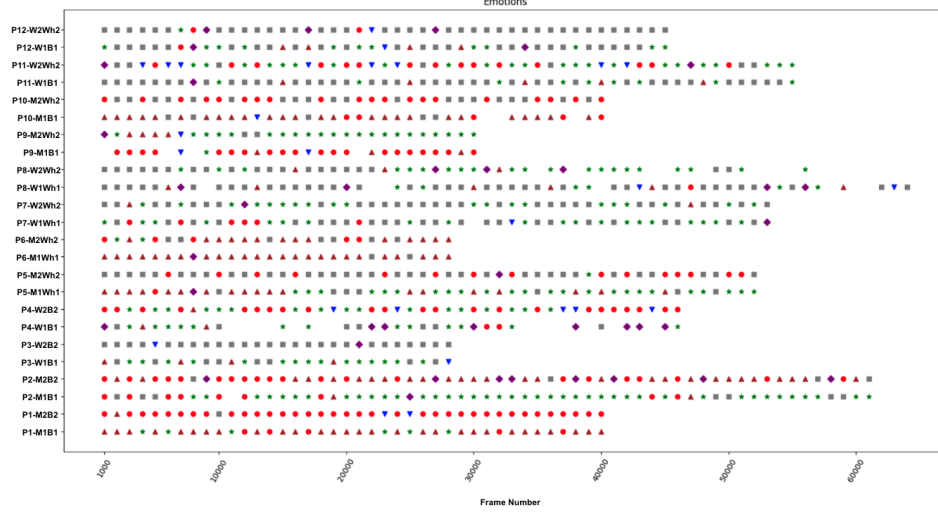
Facial expressions are a kind of non-verbal communication that can be used to relay a message or portray a feeling [31]. We quantitatively analyzed **anger**, **disgust**, **fear**, **happy**, **sad**, and **surprise** as the six standard displays of emotions [29] for each group (4 - same-race (BB), 4 - same-race (WhWh), and 4 - mixed-race (BWh) pairs).

Same-race (BB) pairs displayed the highest average frequency of anger, while same-race (WhWh) pairs displayed the lowest. Anger man-

Table 2. Mean and standard deviation (SD) of facial expression (**anger**, **disgust**, **fear**, **happy**, **sad**, **surprise**) frequencies for same-race (BB) vs. same-race (WhWh) vs. mixed-race (BWh) pairs

		Anger	Disgust	Fear	Happy	Sad	Surprise
Same-Race (BB)	Mean	12	1.1	2.25	11.5	6.6	8.4
	SD	13.2	1.7	3.6	13.8	8.5	12.5
Same-Race (WhWh)	Mean	3.6	0.4	1.5	11.8	20.4	8.6
	SD	5.3	0.7	1.3	10.4	13.8	9.5
Mixed-Race (BWh)	Mean	7.9	1.8	1.1	8.5	17.1	5.3
	SD	8.9	2.6	1.1	9.7	16	8.6

Fig. 2. Participants facial expressions overtime (every 1,000 frames) for same-race (BB) vs. same-race (WhWh) vs. mixed-race (BWh) pairs



ifests through distinct facial cues, including furrowed eyebrows, raised upper eyelids, lifted lower eyelids and tightened lips [71]. This emotional response typically arises when an individual encounters obstacles or frustrations in achieving a desired goal [62]. Same-race (BB) pairs displayed the highest frequency of **anger** (Mean = 12), while same-race (WhWh) pairs demonstrated about one-third of that frequency (Mean = 3.6), refer to Table 2 and Fig. 2. The data for same-race (BB) pairs was disperse (SD = 13.2), which is attributed to same-race (BB) men pair participants being labeled with almost triple the frequency of **anger** (Mean = 17.8), then same-race (BB) women pair participants (Mean = 6.3). This discrepancy between Black men and women may stem from stereotypes perpetuated by AI and machine learning models when analyzing the faces of Black men, suggesting that they are more likely to be perceived as ‘angry’ [49].

In (3/4) same-race (BB) pairs, the navigator exhibited more **anger** than the driver which may be attributed to frustration or confusion. P1-M2B2 (the navigator in a same-race (BB) pair) articulated a challenge the pair faced with the programming syntax, “*The thing that didn’t quite work well was when we were trying to figure out how to put in the nested for loop...it kind of put a damper on the time constraint.*” The heightened expression of **anger** may be attributed to model stereotyping or pairs encountering programming obstacles while completing the tasks, emphasizing the intricacies involved in understanding pair programming communication.

Mixed-race (BWh) pairs displayed the lowest frequency of happy, while same-race (BB) and (WhWh) pairs illustrated the highest frequency. Happiness facial cues include tightened eyes, raised cheeks, and diago-

nally lifted lip corners [66]. This expression is observed when individuals experience success or receive praise [62]. Mixed-race (BWh) pairs displayed the lowest frequency of **happy** (Mean = 8.5), whereas both same-race (BB) pairs (Mean = 11.5) and same-race (WhWh) pairs (Mean = 11.8) demonstrated a nearly equal, higher frequency (refer to Table 2 and Fig. 2). P12-W1B1 delved into possible reasons behind the lack of happiness facial expressions in mixed-race (BWh) pairs, *“I did not like it [pair programming]. It just irritated me. [I decided] I’m not trying to hurt her feelings. So I’m just gonna let this happen.”* Mixed-race (BWh) pairs possibly felt hesitant or uneasy about expressing themselves during the task.

In (5/8) same-race (BB) and (WhWh) pairs the driver exhibited a greater number of **happy** compared to the navigator. P6-M2Wh2 (the driver in a same-race (WhWh) pair) attributed his happiness to his proficiency with the programming language, *“I code daily. So I decided to be the driver. I guess it’s primarily to avoid the syntax hurdles and to be able to focus more on the problem rather than battling the code writing itself.”* Same-race (BB) and (WhWh) pairs likely experienced higher comfort levels while collaborating, with the driver displaying happiness expressions in an authoritative leadership role, while the navigator assumed a laid-back approach.

Same-race (BB) pairs exhibited the lowest frequency of sad, while same-race (WhWh) pairs exhibited the highest. Sadness is expressed on the face through down-turned eyes and drooping lip corners [66]. Sadness occurs when an individual experiences lack of an anticipated outcome [62]. Same-race (BB) pair participants exhibited the lowest frequency of **sad** (Mean = 6.6), refer to Table 2 and Fig. 2. P1-M2B2 (P1 was a same-race (BB) pair) shed light on the cause of a lower frequency of **sad**, *“I think we collaborated very well. He was very easy to talk to, very easy to work with.”* This could indicate that same-race (BB) pairs favored pair programming.

Conversely, same-race (WhWh) pairs exhibited the highest frequency of **sad** (Mean = 20.4), refer to Table 2 and Fig. 2. In (3/4) same-race (WhWh) pairs the driver had more **sad** expressions than the navigator. The same-race (WhWh) women pair participants had roughly double the frequency of **sad** (Mean = 28) compared to same-race (WhWh) men pair participants (Mean = 12.8). This suggests a gender-related effect in our results. P8-W1Wh1 offered insight to the factor contributing to her sadness, *“I don’t think I performed well during pair programming, so I don’t know that I [would] do it again.”* The elevated sadness may indicate a lower self-efficacy while pair programming.

Same-race (BB) and (WhWh) pairs had the highest number of surprise facial expressions compared to mixed-race (BWh) pairs. Surprise is visually conveyed through raised eyebrows, open eyes, and horizontal forehead wrinkles [66]. It typically arises from abrupt activity, unexpected satisfaction, or heightened awareness [62]. Same-race (BB) pairs (Mean = 8.4) and same-race (WhWh) pairs (Mean = 8.6) exhibited the highest frequency of **surprise**, whereas mixed-race (BWh) pairs showed the lowest frequency (Mean = 5.3), refer to Table 2 and Fig. 2. P12-W1B1, from a mixed-race (BWh) pair, shared

insight into her reasoning for restricting communication during the tasks, “*If I was more comfortable with her I probably would have said [more] because we all have one goal to try to get this stuff done. I feel like when working with [other] Black students they don’t see me [a certain] way and so I feel more okay to [say], hey, can we do this? Let’s try this.*” Mixed-race (BWh) pairs possibly experienced lower comfort levels, with their partners, resulting in less expressiveness during the programming tasks.

Table 3. Total frequency of communication prompts for same-race (BB) vs. same-race (WhWh) vs. mixed-race (BWh) pairs

Communication Prompts	Definitions	Same-Race (BB)	Same-Race (WhWh)	Mixed-Race (BWh)
“I” Success Statement	Attributing successes to themselves	2	0	1
“I” Mistake Statement	Taking ownership for their mistakes	7	6	5
“We” Success Statement	Attributing successes to the pair	4	1	5
“We” Mistake Statement	Attributing mistakes to the pair	4	1	3
Apology	A regretful acknowledgment of failure	20	30	22

4.3 RQ3: How do same- and mixed-race pairs exhibit different communication prompts during pair programming?

Communication prompts are used to build better connections and enhance interpersonal ties [36]. Conversational agents’ ability to comprehend dialogue can be improved by understanding communication prompts [68]. Table 3 is the code set we used to label utterances, which was adapted from Robe et al [67].

Same-race (WhWh) pairs apologized more frequently (30 instances) than same-race (BB) pairs (20 instances) and mixed-race (BWh) pairs (22 instances), refer to Table 3. Apologies serve as a strategic way to rebuild trust [26] and repair relationships [23] during conflict. Same-race (WhWh) pairs accepted responsibility for mistakes and apologized for failures, but did not credit themselves or the pair with success, refer to Table 3. For example, P6-M1Wh1 apologized to P6-M2Wh2 for his Java knowledge gap, “*Are we checking the third [character] there? I’m sorry. My Java is rusty.*” Apologies were potentially used by same-race (WhWh) pairs to build trust during collaboration.

Mixed-race (BWh) and same-race (BB) pairs attributed their success and failure to the group, while same-race (WhWh) pairs did not acknowledge their progress as a collective. “We” statements, are associated with enhanced problem-solving [75]. “We” statements were prevalent among mixed-race (BWh) pairs (5 success; 3 mistake) and same-race (BB) pairs (4 success; 4 mistake), in contrast to same-race (WhWh) pairs (1 success; 1 mistake), refer in

Table 3. P10-M1B1, from a mixed-race (BWh) pair, emphasized effective collaboration as a key factor contributing to completing the task, *“What worked well was that we had decent synergy. We sort of found out what we were both good at. I know that I’m good at developing a plan before we go ahead and try to execute versus P10-M2Wh2 was very great hopping into the code.”* Mixed-race (BWh) and same-race (BB) pairs possibly encountered heightened problem-solving during the programming tasks.

5 Discussion

We embark on the first step to bridge perspectives between Black and White developers within both same- and mixed-race pairs, aiming for an understanding of how race may influence communication in the context of pair programming. Mixed-race (BWh) pairs displayed a higher frequency of interruptions (RQ1), the lowest average of **happy** facial expressions, and were the least expressive in conveying **surprise** (RQ2). Additionally, they attributed both success and failure to the pair (RQ3). In the context of same-race (BB) pairs, they exhibited the lowest rate of interruptions (RQ1), demonstrated **happy**, had the lowest average of **sad**, and expressed **surprise** (RQ2). They attributed both success and failure to the pair (RQ3). In the context of same-race (WhWh) pairs, there was a prevalence of both **happy** and **sad**, along with a high occurrence of **surprise** (RQ2). They did not attribute success or failure to the pair, and had the highest frequency of apologies (RQ3). P3-W1B1 emphasized the critical role of facial expressions and its impact on pair programming interactions, *“If [my partner is] frowning, I’m gonna get a little bit smaller and not talk as much. If [they are] receiving and smiling I [feel like] they’re actually listening to what I’m saying and they care. I think that can also play an effect into how I navigate and lead, just based off of how [my partner] is feeling in that moment which can affect my coding output.”* Recognizing the importance of facial expressions during collaboration extends to the broader context of communication as a fundamental component within global software development teams. Communication plays a critical role in facilitating the efficient development of software products [37].

Moreover, implicit bias may impact communication during collaboration. When sharing her workplace experiences while collaborating with individuals from different racial groups, P3-W1B1 stated, *“[It’s] a little bit more difficult, because you have to feel around the vibe of the person a little bit more. You have to toggle certain things on and off whether that be slang, or how you talk, or just little inserts. So I do think if I was doing the same program with a White guy I may have not joked as much. I would have tried to refrain from maybe even thinking aloud [to] keep it more concise.”*

Our study results underscore the potential benefits of developing an empathetic facilitator agent designed to promote inclusive collaboration within both same- and mixed-race pairs. Facilitator agents are software systems crafted to aid users in identifying and resolving challenges [40]. Detectable elements such as interruptions, facial expressions, and communication prompts can be monitored

by facilitator agents. By tracking these aspects, the agent has the potential to offer encouraging feedback to individuals during collaboration, thereby enhancing the overall experience and facilitating cohesive collaboration. This implication aligns with our goal of bridging perspectives between two developers to foster greater synergy.

6 Conclusion

Our study marks an initial approach, in Human-Computer Interaction (HCI) and Software Engineering (SE), to delving into communication between same- and mixed-race pairs of developers engaged in remote pair programming. We employed a blend of quantitative and qualitative methods, to identify key findings:

- **RQ1 (Interruptions):** Mixed-race (BWh) pairs exhibited the highest frequency of interruptions, while same-race (BB) pairs demonstrated the lowest.
- **RQ2 (Non-Verbal Cues):** Mixed-race (BWh) pairs demonstrated reduced happiness, in contrast to same-race (BB) and (WhWh) pairs who showcased peak happiness and surprise.
- **RQ3 (Communication Prompts):** Mixed-race (BWh) and same-race (BB) pairs credited success to their collaboration.

Our findings highlight potential complexities of racial dynamics in pair programming communication. To address communication challenges, we aim to contribute to the creation of inclusive facilitator agents to enrich technology and collaboration.

Acknowledgements. This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-21-1-0108 and National Science Foundation under award numbers IIS-2313890 and CCF-2006977. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the view of the NSF and AFOSR. Finally, we would like to thank Raphael Phillips for his assistance in labeling the data.

References

1. https://ips-dc.org/wp-content/uploads/2019/01/IPS_RWD-Report_FINAL-1.15.19.pdf
2. <https://www.census.gov/content/dam/Census/library/publications/2016/demo/p20-578.pdf>
3. Adamczyk, P.D., Bailey, B.P.: If not now, when? the effects of interruption at different moments within task execution. In: Proceedings of the SIGCHI conference on Human factors in computing systems. pp. 271–278 (2004)
4. Ågren, P., Knoph, E., Berntsson Svensson, R.: Agile software development one year into the covid-19 pandemic. *Empirical Software Engineering* **27**(6), 121 (2022)

5. Archibald, M.M., Ambagtsheer, R.C., Casey, M.G., Lawless, M.: Using zoom videoconferencing for qualitative data collection: perceptions and experiences of researchers and participants. *International journal of qualitative methods* **18**, 1609406919874596 (2019)
6. Asare, J.G.: Google's 2019 diversity report reveals more progress must be made (Apr 2019), <https://www.forbes.com/sites/janicegassam/2019/04/07/googles-2019-diversity-report-reveals-more-progress-must-be-made/?sh=77eb81a03bef>
7. Bandura, A.: Social foundations of thought and action: a social cognitive theory. Prentice-Hall series in social learning theory, Prentice-Hall (1986), <https://books.google.com/books?id=HJhqAAAAAAAJ>
8. Barton, P.E.: Parsing the achievement gap: Baselines for tracking progress. policy information report. (2003)
9. Beck, K.: Test driven development: By example. Addison-Wesley Professional (2022)
10. Bergen, N., Labonté, R.: “everything is perfect, and we have no problems”: detecting and limiting social desirability bias in qualitative research. *Qualitative health research* **30**(5), 783–792 (2020)
11. Bipp, T., Lepper, A., Schmedding, D.: Pair programming in software development teams - an empirical study of its benefits. *Inf. Softw. Technol.* **50**(3), 231–240 (Feb 2008). <https://doi.org/10.1016/j.infsof.2007.05.006>, <http://dx.doi.org/10.1016/j.infsof.2007.05.006>
12. Bocard, N., Zenou, Y., et al.: Racial discrimination and redlining in cities. Tech. rep., Université catholique de Louvain, Center for Operations, Research & Econometrics (1999)
13. Bogen, D.S.: From racial discrimination to separate but equal: The common law impact of the thirteenth amendment. *Ohio Northern University Law Review* **38**(1), 3 (2023)
14. Bosson, J.K., Haymovitz, E.L., Pinel, E.C.: When saying and doing diverge: The effects of stereotype threat on self-reported versus non-verbal anxiety. *Journal of experimental social psychology* **40**(2), 247–255 (2004)
15. Bredin, H., Yin, R., Coria, J.M., Gelly, G., Korshunov, P., Lavechin, M., Fustes, D., Titeux, H., Bouaziz, W., Gill, M.P.: Pyannote. audio: neural building blocks for speaker diarization. In: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 7124–7128. IEEE (2020)
16. Bryant, S.: Double trouble: Mixing qualitative and quantitative methods in the study of extreme programmers. In: 2004 IEEE symposium on visual languages-human centric computing. pp. 55–61. IEEE (2004)
17. Bryant, S., Romero, P., du Boulay, B.: The collaborative nature of pair programming. In: International Conference on Extreme Programming and Agile Processes in Software Engineering. pp. 53–64. Springer (2006)
18. Burke, A., Okrent, A., Hale, K., Gough, N.: The state of us science & engineering 2022. national science board science & engineering indicators. nsb-2022-1. National Science Foundation (2022)
19. Campion, M.A., Papper, E.M., Medsker, G.J.: Relations between work team characteristics and effectiveness: A replication and extension. *Personnel psychology* **49**(2), 429–452 (1996)
20. Celepkolu, M., Boyer, K.E.: Thematic analysis of students' reflections on pair programming in cs1. In: Proceedings of the 49th ACM Technical Symposium on

- Computer Science Education. p. 771–776. SIGCSE '18, Association for Computing Machinery, New York, NY, USA (2018). <https://doi.org/10.1145/3159450.3159516>
21. Cheong, J.H., Jolly, E., Xie, T., Byrne, S., Kenney, M., Chang, L.J.: Py-feat: Python facial expression analysis toolbox. *Affective Science* pp. 1–16 (2023)
 22. Choi, S.: “better communication leads to a higher output?” an analysis of pair communication on pair programming productivity. *IEEE Transactions on Professional Communication* **64**(4), 338–353 (2021)
 23. Coombs, W.T., Holladay, S.J.: Comparing apology to equivalent crisis response strategies: Clarifying apology’s role and value in crisis communication. *Public Relations Review* **34**(3), 252–257 (2008)
 24. Cooper, S., Clinkscale, B., Williams, B., Lewis, M.: Exploring the impact of exposing cs majors to programming concepts using ide programming vs. non-ide programming in the classroom. In: *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. pp. 1422–1422 (2020)
 25. Dake, A.: 2020 state of salaries report: Salary benchmarks and talent preferences (Sep 2022), <https://hired.com/blog/highlights/2020-state-of-salaries-report/>
 26. De Greiff, P.: The role of apologies in national reconciliation processes: On making trustworthy institutions trusted. *The age of apology: Facing up to the past* pp. 120–134 (2008)
 27. Delgado, R., Stefancic, J.: *Critical race theory: An introduction*, vol. 87. NyU press (2023)
 28. Dybå, T., Arisholm, E., Sjøberg, D., Hannay, J., Shull, F.: Are two heads better than one? on the effectiveness of pair programming. *Software, IEEE* **24**, 12 – 15 (12 2007). <https://doi.org/10.1109/MS.2007.158>
 29. Ekman, P., Sorenson, E.R., Friesen, W.V.: Pan-cultural elements in facial displays of emotion. *Science* **164**(3875), 86–88 (1969)
 30. Feagin, J.R., McKinney, K.D.: *The many costs of racism*. Rowman & Littlefield Publishers (2005)
 31. Gallaher, P.E.: Individual differences in nonverbal behavior: Dimensions of style. *Journal of personality and social psychology* **63**(1), 133 (1992)
 32. Gramlich, J.: The gap between the number of blacks and whites in prison is shrinking (Apr 2019), <https://www.pewresearch.org/short-reads/2019/04/30/shrinking-gap-between-number-of-blacks-and-whites-in-prison/>
 33. Gregory, P., Lassenius, C., Wang, X., Kruchten, P.: *Agile Processes in Software Engineering and Extreme Programming: 22nd International Conference on Agile Software Development, XP 2021, Virtual Event, June 14–18, 2021, Proceedings*. Springer Nature (2021)
 34. Han, K.W., Lee, E., Lee, Y.: The impact of a peer-learning agent based on pair programming in a programming course. *Education, IEEE Transactions on* **53**, 318 – 327 (06 2010). <https://doi.org/10.1109/TE.2009.2019121>
 35. Hanks, B.: Student performance in cs1 with distributed pair programming. *ACM SIGCSE Bulletin* **37**(3), 316–320 (2005)
 36. Hartley, P.: *Interpersonal communication*. Routledge (2002)
 37. Herbsleb, J.D., Mockus, A.: An empirical study of speed and communication in globally distributed software development. *IEEE Transactions on Software Engineering* **29**(6), 481–494 (Jun 2003)
 38. Howard, E.V.: Attitudes on using pair-programming. *Journal of Educational Technology Systems* **35**(1), 89–103 (2006)

39. Hughes, J., Walshe, A., Law, B., Murphy, B.: Remote pair programming. In: 12th International Conference on Computer Supported Education. pp. 476–483. SciTePress (2020)
40. Ikeda, Y., Shiramatsu, S.: Generating questions asked by facilitator agents using preceding context in web-based discussion. In: 2017 IEEE International conference on agents (ICA). pp. 127–132. IEEE (2017)
41. Jaccard, P.: Etude de la distribution florale dans une portion des alpes et du jura. *Bulletin de la Societe Vaudoise des Sciences Naturelles* **37**, 547–579 (01 1901). <https://doi.org/10.5169/seals-266450>
42. Jackson, P.B., Thoits, P.A., Taylor, H.F.: Composition of the workplace and psychological well-being: The effects of tokenism on america’s black elite. *Social Forces* **74**(2), 543–557 (1995)
43. Jensen, E.: Measuring racial and ethnic diversity for the 2020 census (Jun 2022), <https://www.census.gov/newsroom/blogs/random-samplings/2021/08/measuring-racial-ethnic-diversity-2020-census.html>
44. Joshi, A., Roh, H.: The role of context in work team diversity research: A meta-analytic review. *Academy of management journal* **52**(3), 599–627 (2009)
45. Kaiser, C.R., Miller, C.T.: Stop complaining! the social costs of making attributions to discrimination. *Personality and Social Psychology Bulletin* **27**(2), 254–263 (2001)
46. Kaplan, S.: Scientists show how we start stereotyping the moment we see a face (May 2016), <https://www.washingtonpost.com/news/speaking-of-science/wp/2016/05/02/scientists-show-how-we-start-stereotyping-the-moment-we-see-a-face/>
47. Kaur Kuttal, S., Gerstner, K., Bejarano, A.: Remote pair programming in online cs education: Investigating through a gender lens. In: 2019 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC). pp. 75–85 (2019)
48. King, R.S.: Jim crow is alive and well in the 21st century: Felony disenfranchisement and the continuing struggle to silence the african-american voice. *Souls* **8**(2), 7–21 (2006)
49. Kyriakou, K., Kleanthous, S., Otterbacher, J., Papadopoulos, G.A.: Emotion-based stereotypes in image analysis services. In: Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization. pp. 252–259 (2020)
50. L. Jones, D., D. Fleming, S.: What use is a backseat driver? a qualitative investigation of pair programming. pp. 103–110 (Sep 2013)
51. Lemov, D.: Teach like a champion: 49 techniques that put students on the path to college (K-12). John Wiley & Sons (2010)
52. Lev-Ram, M.: Apple commits more than \$50 million to diversity efforts. *Fortune* (2015)
53. Lewis, C.H.: Using the “thinking aloud” method in cognitive interface design. RC 9265, IBM (1982)
54. Mcdowell, C., Werner, L., Bullock, H., Fernald, J.: The impact of pair programming on student performance, perception and persistence. pp. 602– 607 (06 2003). <https://doi.org/10.1109/ICSE.2003.1201243>
55. Mcdowell, C., Werner, L., Bullock, H., Fernald, J.: The effects of pair-programming on performance in an introductory programming course. vol. 34, pp. 38–42 (03 2002). <https://doi.org/10.1145/563340.563353>
56. Mcdowell, C., Werner, L., Bullock, H., Fernald, J.: Pair programming improves student retention, confidence, and program quality. *Commun. ACM* **49**, 90–95 (08 2006). <https://doi.org/10.1145/1145293>

57. Murata, K.: Intrusive or co-operative? a cross-cultural study of interruption. *Journal of Pragmatics* **21**(4), 385–400 (1994)
58. Nosek, J.: The case for collaborative programming. *Communications of the ACM* **41** (03 1998). <https://doi.org/10.1145/272287.272333>
59. Oviatt, S., Cohen, P.: Perceptual user interfaces: Multimodal interfaces that process what comes naturally. *Commun. ACM* **43**(3), 45–53 (Mar 2000). <https://doi.org/10.1145/330534.330538>
60. Pinel, E.C.: Stigma consciousness: the psychological legacy of social stereotypes. *Journal of personality and social psychology* **76**(1), 114 (1999)
61. Plonka, L., Sharp, H., Van Der Linden, J.: Disengagement in pair programming: Does it matter? In: 2012 34th international conference on software engineering (ICSE). pp. 496–506. IEEE (2012)
62. Plutchik, R.: *The emotions*. University Press of America (1991)
63. Purdie-Vaughns, V., Steele, C.M., Davies, P.G., Dittmann, R., Crosby, J.R.: Social identity contingencies: how diversity cues signal threat or safety for african americans in mainstream institutions. *Journal of personality and social psychology* **94**(4), 615 (2008)
64. Rankin, J.L.: Learning to code isn't enough (Apr 2023), <https://www.technologyreview.com/2023/04/20/1071291/learn-to-code-legacy-new-projects-education/>
65. Riccucci, N.M.: *Managing diversity in public sector workforces*. Routledge (2021)
66. Rinn, W.E.: The neuropsychology of facial expression: a review of the neurological and psychological mechanisms for producing facial expressions. *Psychological bulletin* **95**(1), 52 (1984)
67. Robe, P., Kuttal, S.K.: Designing pairbuddy—a conversational agent for pair programming. *ACM Trans. Comput.-Hum. Interact.* **29**(4) (may 2022). <https://doi.org/10.1145/3498326>, <https://doi.org/10.1145/3498326>
68. Rodríguez, F.J., Price, K.M., Boyer, K.E.: Exploring the pair programming process: Characteristics of effective collaboration. In: *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*. pp. 507–512. ACM (2017)
69. Roth, W.D.: The multiple dimensions of race. *Ethnic and Racial Studies* **39**(8), 1310–1338 (2016)
70. Seaman, C.B.: "qualitative methods in empirical studies of software engineering". In: *IEEE Transactions on Software Engineering*. vol. 25, pp. 557–572 (1999)
71. Sell, A., Cosmides, L., Tooby, J.: The human anger face evolved to enhance cues of strength. *Evolution and Human Behavior* **35**(5), 425–429 (2014)
72. Sfetsos, P., Stamelos, I., Angelis, L., Deligiannis, I.: Investigating the impact of personality types on communication and collaboration-viability in pair programming—an empirical study. In: *International conference on extreme programming and agile processes in software engineering*. pp. 43–52. Springer (2006)
73. Shelton, J.N.: Interpersonal concerns in social encounters between majority and minority group members. *Group Processes & Intergroup Relations* **6**(2), 171–185 (2003)
74. Sidner, C.L., Lee, C., Kidd, C.D., Lesh, N., Rich, C.: Explorations in engagement for humans and robots. *Artificial Intelligence* **166**(1-2), 140–164 (2005)
75. Simmons, R.A., Gordon, P.C., Chambless, D.L.: Pronouns in marital interaction: What do “you” and “i” say about marital health? *Psychological science* **16**(12), 932–936 (2005)
76. Sommerville, I.: *Software Engineering*. Addison-Wesley, Harlow, England, 9 edn. (2010)

77. Stapel, K., Knauss, E., Schneider, K., Becker, M.: Towards understanding communication structure in pair programming. In: Agile Processes in Software Engineering and Extreme Programming: 11th International Conference, XP 2010, Trondheim, Norway, June 1-4, 2010. Proceedings 11. pp. 117–131. Springer (2010)
78. Steele, C.M.: A threat in the air: How stereotypes shape intellectual identity and performance. *American psychologist* **52**(6), 613 (1997)
79. Sun, W., Marakas, G., Aguirre-Urreta, M.: The effectiveness of pair programming: Software professionals’ perceptions. *IEEE Software* **33**(4), 72–79 (2015)
80. Tolan, C., Ash, A., Marsh, R.: The nation’s largest credit union rejected more than half its black conventional mortgage applicants | cnn business (Dec 2023), <https://www.cnn.com/2023/12/14/business/navy-federal-credit-union-black-applicants-invs/index.html>
81. Tsompanoudi, D., Satratzemi, M., Xinogalos, S., Karamitopoulos, L.: An empirical study on factors related to distributed pair programming (04 2019), <https://www.learntechlib.org/p/208576>
82. Voss, J.L., Gonsalves, B.D., Federmeier, K.D., Tranel, D., Cohen, N.J.: Hippocampal brain-network coordination during volitional exploratory behavior enhances learning. *Nature neuroscience* **14**(1), 115–120 (2011)
83. Williams, L., Kessler, R.R., Cunningham, W., Jeffries, R.: Strengthening the case for pair programming. *IEEE Software* **17**(4), 19–25 (2000)
84. Williams, L.A., Kessler, R.R.: All i really need to know about pair programming i learned in kindergarten. *Commun. ACM* **43**(5), 108–114 (May 2000). <https://doi.org/10.1145/332833.332848>
85. Youngquist, J.: The effect of interruptions and dyad gender combination on perceptions of interpersonal dominance. *Communication Studies* **60**(2), 147–163 (2009)