Evaluating Gender Significance Within a Pair Programming Context

Kyungsub S. Choi
School of Management, Rhode Island College, Providence, RI
kchoi@ric.edu

Abstract
This study reports findings of the gender differences within a pair programming context. A large pool of university computer programming course undergraduate and graduate students was paired into three distinct pair categories: male-male, male-female, and female-female. All pairs performed pair programming tasks under a controlled lab environment. The pairs’ final outputs in the categories of code productivity and code design were quantitatively measured, and the post-experiment questionnaire was also measured. The results revealed that the male-male pairs had the highest scores and female-female pairs had the lowest scores in both code categories, but not significantly. In the post-experiment questionnaire, the communication and compatibility levels showed significant differences between the three different pair categories. The male-male, female-female pairs or same gender pairs had significantly higher levels than the male-female or mixed gender pairs. Additionally, the female participants particularly voiced gender-biased concerns about collaborating with male partners in doing the pair programming task.

1. Introduction

Gender is, and likely will continue to be, one of most debated topics among researchers. It remains a controversial subject of research inquiry across disciplines. Despite a vast amount of studies, we are still grappling with fully understanding its exact effects and meaning.

On the surface, the main debate on the effect of gender is between ‘substantive difference’ versus ‘no substantive difference’. The problem is that, while every research study claims to be objective and unbiased, as it were, we find copious studies that support either argument. Complicating this further, research communities are expected to provide the answers to even more complex gender implied questions.

From an organizational perspective, a gender related question becomes particularly relevant within a team-play setting; higher team productivity, performance, quality, communication, cohesion, efficiency, and other attributes come to focus. There has been a sharp increase of women in science and engineering due to the steadfast effort by both government and private sectors. Pragmatically, a clear answer to the on-going gender question and gender-related team-play questions would shed light on more efficient and effective use of human resources for organizations and management. It would be rare in today’s global and distributed business ecosystem to find a team that doesn’t employ collaborative processes and operations.

In an IS environment, the gender question is equally significant. In the highly specialized IS areas, such as software engineering or system analysis, team-play is often emphasized. The ability to manage human resources effectively would greatly enhance the overall organization’s IS functions.

This study probes gender differences using a pair programming lens. Briefly explained, pair programming (PP) is an unorthodox programming practice where two individuals share one keyboard and one monitor simultaneously, while undertaking group programming [3]. Programming can be considered a task that allows one’s full manifestation of problem solving and logical ability [45]. When two individuals are harnessed to unify their programming processes, it engenders a series of cognitive collisions and problem solving conflicts [19]. The PP underlying psychosocial impacts bring a number of obstacles to the paired individuals in managing the conflicts and reaching compromises [7], [8].

Using the PP lens, this study aims to understand how mixed gender groups program together, what programming output level the mixed gender group generates, and how well they interact – in terms of compatibility and communication – during a programming session. This study’s result is expected to assist programming shop managers better understand the gender issue and expedite their group programming schedules more effectively.

2. Literature review

2013 46th Hawaii International Conference on System Sciences

1530-1605/12 $26.00 © 2012 IEEE
DOI 10.1109/HICSS.2013.209
There are various different research approaches to the gender variable. Pinker and Spelke [30] talk about ‘nature vs. nurture’ that focuses on the developmental aspect rather than biological nature. They point to the fact that gender is only an issue during infancy and its significance depends on how a person is raised. The developmental environment could heavily influence one’s cognitive behavior. Rubin [35] couples the gender variable with the social construction of an individual. Rubin states that the blend of developmental and social environments systematically shapes the lives of men and women in our society. Bern [5] echoes Rubin’s perspective, emphasizing that gender-related experiences limit women and men to their subservient gender roles. Ridley [32] mentions the significance of human genes. Naturally, human genes set males and females apart in the beginning, but the genes also nurture males and females to be different or similar depending on individual behaviors or environment; Ridley claims that our genes are “both cause and consequence of our actions” (page 6).

Others remain unconvinced of any gender difference. Hyde and Lindberg [26] have reviewed a large number of various gender meta-analysis and concluded that these analyses are questionable because of the methodological issues, the individual differences, as well as the researcher’s beliefs and inclinations.

To corral all these polarized views and gain a broader perspective across different disciplines, this study conducts a literature review of the different domains.

2.1. Gender difference in learning competency

Many studies report a gender difference in the learning style between men and women [23], [28], [43]. The results of the studies initially report that men perform significantly “better” than women in analytical studies, such as science, technology, engineering, and mathematics [28]; however, the authors point out that the pedagogical instruments and the style of instruction – the instructor’s knowledge “transmission” – favors men over women. Barrett [4] delineates this by claiming that men tend to respond well to a conventional pedagogical style, an authoritarian teaching style, whereas women identify more with a conventional pedagogical style, whereas women identify more with a facilitating style.

Analyzing the emotional differences in gender in a learning environment, Stoilescu and McDougall [40] report that women demonstrate higher anxiety, a lack of confidence, and underachievement in the categories of equity with instruction in both computer science and in opportunities to foster computer culture. In the experiment, women showed significantly higher levels of anxiety in completing the assigned computer task as they experienced difficulties when the process stalled. This research complements Durnell and Haag’s [14] study, which reported that men scored higher in computer self-efficacy, lower in anxiety, and maintained more positive attitudes toward the use of the Internet.

Gender also plays a role in attitude, perception, and judgment [16], [17], [21]. Women are found to be more concerned about others and tend to conform more to societal norms; in contrast, men are driven more by individual attitudes and objectives. The general consensus is that women prefer feeling-based judgment and manage well with ambiguity, whereas men prefer facts, an analytic-based judgment, and a logical and rational approach.

2.2. Gender difference in brain anatomy

We found two biological studies on gender difference particularly significant. These investigated gender differences by reviewing functional magnetic resonance imaging (fMRI) captured brain activity. Riedl, et al. [31] investigated the brain activity of 20 participants – 10 males and 10 females – on how they processed their decisions on the trustworthiness of eBay offers. The results revealed that there are “several gender-specific trust regions, but there are also two gender independent trust areas” (pg. 419) and that “women activated more brain areas than did men” (pg. 415). Furthermore, the study result confirms gender differences in neural information processing modes.

Similarly, Dimoka [13] studied the cerebral brain activity of 15 right-handed subjects - 9 males and 6 females – while they carried out given tasks. The result revealed that women reported higher brain activity in a certain brain area (the limbic system), while the men reported higher brain activity from a different brain area (the prefrontal area) under the same circumstances. The report concludes that “the emotional responses in the brain are more salient for women, implying that there are differences between genders” (pg. 388). With these facts, it is clear that certain neurological and physiological differences exist between men and women [20]. Further, trust and neural information processing are closely related to pair programming as both programmers must trust and bond to each other when focusing on and processing complicated programming sequence information [3].

2.3. No gender difference view

A more radical gender view is that there is no gender difference. A number of items – methodological problems, individual differences,
inference and generalization problems, and the researcher’s beliefs and inclination - are identified as underlying reasons [25], [26], [38]. Eccles et al [15] presents the expectancy-value model. Briefly, the model posits that males and females have different expectations and values of success in different areas. And because of these differences, males and females select and choose differently, thus further emphasizing their gender differences. 

Hyde and Lindberg [26] recommend that researchers not only focus on finding gender differences but that they also search for gender similarities; in sum, they call for a more balanced gender study that would enrich the gender research endeavor.

2.4. Gender difference in Information Systems

While other major disciplines remain cautious about boldly asserting a gender gap [12], the IS domain putatively presumes the gender difference [1]. Many IS studies that deal with the gender variable tend to fall under one of these two questions: 1) how is computing technology perceived and consumed differently by men and women; and 2) why is there such a low women population in the computing technology sector, what are the causes, and what can we do about it.

The gender perspective in IS generally begins with the notion that computing technology knowledge acquisition favors men [1], [23], [24]. In other words, men exhibit a significantly higher level of mastery of computing technology knowledge over women. For the reasons behind this claim, one may point to knowledge method and delivery [11], [18], [44], social and structural issues [2], and psycho-social concerns [22].

In the knowledge method and delivery, Gefen and Straub [18] place an emphasis on the communication used with men and women in introducing technology. They report that men tend to “focus discourse on hierarchy and independence, while women focus on intimacy and solidarity” (pg. 389). Cooper and Weaver [11] and Barrett [4] claim that knowledge method and delivery are more conducive for men than for women. For the social and structural issues, Ahuja [2] and Cooper and Weaver [11] list: social expectations, occupational culture, peers and parental support and influence. They surmise that, regardless of one’s effort and attitude, our society and its social structure is not particularly conducive for women to enjoy computing technology, unlike men. Psycho-social concerns provide another explanation. The major finding is that women exhibit a significantly higher level of anxiety than men in adopting computing technology [22].

A number of IS studies are committed to seeking solutions to redress this gender gap. Given the reports that, currently, computing technology better fits a masculine demographic, the consensus in remedying this gender gap is to gear computing technology knowledge transfer method and delivery in pro-feminine ways [1], [24]. Some recommend shunning today’s authoritarian pedagogical style [4] and embracing more women participants into environments designed to support the development of their skills [11], [33]. Some even suggest women-only learning environments [11]. One prudent approach for both women and men would be lowering the identified “barriers” [29], initiating feminist positivism [1], mixing gender within group collaborations [37], and encouraging and supporting with effective persuasive messages [33].

Werner and Denning [46] have used PP to understand the women shortage problem in science and engineering disciplines. Attributing this lack to a competitive, masculinity culture and to individual women’s low confidence in problem solving on the computer, the women shortage problem in those disciplines was microscopically examined through a group of middle school female students engaged in PP.

A total of one hundred twenty-six girls were paired in groups of two, and given a computer programming task. Each group was instructed to make decisions en route to completing the given task. The PP sessions were audio-recorded.

The experiment results revealed that the girls lacked the computer programming task-required persistence, succumbing to the difficulties of computer programming. But through the PP sessions, the girls complemented and supported each other to stay with the given task.

The girls exercised a metacognitive analysis in: identifying and isolating a problem, determining and performing the necessary work, and testing the solution. These metacognitive steps were materialized through the PP and team collaboration. This study underpins a notion that if female individuals are supported by their peers and others, and given a safe-net facilitation, they can be productive in computer programming task.

3. Hypothesis development

Given the gender duality of male and female, the available combinations are male-female, male-male, and female-female. In other words, we have a homogeneous or same gender pair category, and a heterogeneous or mixed gender pair category.

The PP would require a uniquely high level of friendly group collaboration. One’s cognitive and psychological processes are not that easy to share nor do they necessarily enable collaboration with another
individual [10], [45], [47]. Sharing a keyboard and a computer monitor with another individual of the opposite gender can be more challenging [8], [19]. Given these circumstances, the aim is to secure high coding output, pair communication, and pair compatibility.

According to published reports, males are expected to have an easier time with coding output given their technology knowledge [1], [24], [27], [28]. Females, on the other hand, are expected 1) to provide a higher level of pair communication and compatibility [46]; 2) to provide pair closeness, solidarity, and support [18], [46]. Reports indicate that women exhibit a higher level of communication, freely using more words and body gestures [39], [42]. With these facts, the diversity of male and female pairs may achieve the objectives [6]. We propose the following hypotheses:

H1: The male-female [MF] pairs would significantly achieve higher level of code activity than the same gender pairs, [MM] and [FF].

H1.1: code productivity
H1.2: code design

H2: The male-female [MF] pairs would significantly achieve higher level of compatibility than the same gender pairs, [MM] and [FF].

H3: The male-female [MF] pairs would significantly achieve higher level of communication than the same gender pairs, [MM] and [FF].

4. Administering the experiment

A pool of 128 university computer programming course undergraduate and graduate students – 93 males and 35 females - participated in the experiment.

For the pairing process, each participant was carefully screened for any possible covariates, such as course level, class academic performance, and programming experience. Given the academic backdrop, no participants had any substantive professional programming experience. For the pairs, we had 23 male-female pairs, 35 male-male pairs, and 6 female-female pairs.

A set of two PP problems was devised. The difficulty and parity levels of the problems were reviewed and fine-tuned by three professional programmers. Through a few pilot trials, 45 minutes per problem was found to be adequate for the purpose of this experiment. The experiment was conducted at a controlled PC lab where noise and distractions were completely excluded or, at least, controlled.

Each pair had two PP sessions. They were ordered to exercise the PP protocol, cognitive sharing, and collaboration in completing the assigned computer programming problems. The navigator role was exercised and swapped in a timely manner by the experiment facilitator to expose both programmers equally to the role. The outputs were evaluated and scored in the categories of code design and code productivity by a panel of judges.

The judging of code design and code productivity is a subjective evaluation. For the code design, the code efficiency and code readability were considered. For the code productivity, the code lines and accuracy were considered. Although this is a subjective evaluation, the evaluation had to be reasonably agreeable to the other equally competent and experienced programming experts. A series of meetings were conducted to encourage the transparency and validation of the judging process and decisions.

After a complete review of the submitted printouts by the subjects, the judges were instructed to give each a score for code productivity and code design, from a scale of 0 to 10.

Relying on the judges’ programming expertise and subjectivity, the judges’ inter-reliability is controlled by bivariate correlation analysis. The two judges had highly agreeable values 0.963-code productivity and 0.967-code design.

A post-experiment questionnaire [7] was devised to capture the participants’ PP experience (see Appendix). The factor analysis and Cronbach reliability measurement were used for instrument validation [9], [34], [41]. The factor analysis was used by the means of principal components analysis with varimax rotation. Cronbach reliability measurement indicates how well a set of items measure a single latent construct. Typically, 0.700 is viewed as the acceptable minimal value and the higher the value, the more reliable the set is [9]. We have the Cronbach value of 0.8794 for PP communication, and 0.7084 for PP compatibility.

5. Analysis and results

Table 1 shows the general description of the code design and code productivity analysis. By the mean values, the MM pairs in both code categories show the highest scores.
Table 1. Code evaluation descriptive statistics

<table>
<thead>
<tr>
<th>Pair</th>
<th>N</th>
<th>Mean</th>
<th>S. D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>code prod.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MM</td>
<td>67</td>
<td>5.60</td>
<td>2.95</td>
</tr>
<tr>
<td>MF</td>
<td>44</td>
<td>5.44</td>
<td>3.03</td>
</tr>
<tr>
<td>FF</td>
<td>12</td>
<td>4.63</td>
<td>3.04</td>
</tr>
<tr>
<td>code design</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MM</td>
<td>67</td>
<td>5.60</td>
<td>2.88</td>
</tr>
<tr>
<td>MF</td>
<td>44</td>
<td>5.42</td>
<td>3.01</td>
</tr>
<tr>
<td>FF</td>
<td>12</td>
<td>4.50</td>
<td>2.92</td>
</tr>
</tbody>
</table>

The data was checked for normal distribution [34]. Besides the popular normality tests, histogram and normal Q-Q plots, for more accuracy, we considered two normality tests: the Kolmogorov-Smirnov and Shapiro-Wilk tests. We used the Kolmogorov-Smirnov test (Table 2). This test compares the set of scores in the sample to a normally distributed set of scores with the same mean and standard deviation. If the test result is non-significant (p > 0.05), the sample data is normally distributed. Both code productivity and code design showed significant values (p = 0.000), thus the data failed the normality test.

The normal distribution test was also administered on the questionnaire result. The non-parametric test, the Kolmogorov-Smirnov test was also conducted (Table 4). The result failed to meet the normal distribution requirements: PP communication (p = 0.000) and PP compatibility (p = 0.000).

Table 2. Coding output Kolmogorov-Smirnov test

<table>
<thead>
<tr>
<th></th>
<th>statistic</th>
<th>df</th>
<th>sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>code prod.</td>
<td>.145</td>
<td>123</td>
<td>.000</td>
</tr>
<tr>
<td>code design</td>
<td>.118</td>
<td>123</td>
<td>.000</td>
</tr>
</tbody>
</table>

We additionally exercised the data transformation option, but it yielded the same failed result. The verdict is that the data is unfit for parametric tests.

Given these circumstances, a parametric test, the Kruskal-Wallis test was used to test any significant differences among the three pair categories. The Kruskal-Wallis test is an extension of the Mann-Whitney U test and is a nonparametric analog of a one-way analysis of variance (ANOVA), which detects differences in distribution location.

The Kruskal-Wallis test result is shown in Table 3. Both code productivity (0.579, p > 0.05) and code design (0.443, p > 0.05) show there are no significant differences. In other words, gender had no significant impact on the pair output. Consequently, no further group contrasting analysis was performed.

Table 3. Coding output Kruskal-Wallis test

<table>
<thead>
<tr>
<th></th>
<th>chi-square</th>
<th>df</th>
<th>sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>code prod.</td>
<td>1.093</td>
<td>2</td>
<td>0.579</td>
</tr>
<tr>
<td>code design</td>
<td>1.630</td>
<td>2</td>
<td>0.443</td>
</tr>
</tbody>
</table>

Given the abnormal data distribution, the Kruskal-Wallis test was performed (Table 5), and PP communication (p < 0.05), PP compatibility (p < 0.05) show significant differences among the three pair categories.

Table 4. Questionnaire Kolmogorov-Smirnov test

<table>
<thead>
<tr>
<th></th>
<th>statistic</th>
<th>df</th>
<th>sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP comm.</td>
<td>.112</td>
<td>210</td>
<td>.000</td>
</tr>
<tr>
<td>PP compat.</td>
<td>.121</td>
<td>210</td>
<td>.000</td>
</tr>
</tbody>
</table>

Table 5. Questionnaire Kruskal-Wallis test

<table>
<thead>
<tr>
<th></th>
<th>chi-square</th>
<th>df</th>
<th>sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP comm.</td>
<td>12.105</td>
<td>2</td>
<td>.002</td>
</tr>
<tr>
<td>PP compat.</td>
<td>19.209</td>
<td>2</td>
<td>.000</td>
</tr>
</tbody>
</table>

To assess which pair category significantly differs from which category, the Mann-Whitney test was executed for each construct: PP communication (table 6) and PP compatibility (table 7). Table 6 indicates that [MM] shows a significantly higher mean (p < 0.01) than [MF], [FF] shows a significantly higher mean (p < 0.01) than [MF], and no significance in means (p = 0.144) between [MM] and [FF].

Table 6. PP communication Mann-Whitney test

<table>
<thead>
<tr>
<th></th>
<th>[MM] vs. [MF]</th>
<th>[MF] vs. [FF]</th>
<th>[MM] vs. [FF]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>3596.5</td>
<td>444.0</td>
<td>958.5</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>6836.5</td>
<td>3684.0</td>
<td>8098.5</td>
</tr>
<tr>
<td>Z</td>
<td>-2.921</td>
<td>-2.808</td>
<td>-1.063</td>
</tr>
<tr>
<td>sig.</td>
<td>.002</td>
<td>.003</td>
<td>.144</td>
</tr>
</tbody>
</table>

Table 7 shows that [MM] exhibits a significantly higher mean (p < 0.01) than [MF], [FF] shows a significantly higher mean (p < 0.01) than [MF], but no significant difference between [FF] and [MM].
Table 7. PP compatibility Mann-Whitney test

<table>
<thead>
<tr>
<th></th>
<th>[MM] vs. [MF]</th>
<th>[MF] vs. [FF]</th>
<th>[MM] vs. [FF]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>3255.0</td>
<td>482.000</td>
<td>1136.000</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>6576.0</td>
<td>3803.000</td>
<td>8639.000</td>
</tr>
<tr>
<td>Z</td>
<td>-4.123</td>
<td>-2.797</td>
<td>-0.495</td>
</tr>
<tr>
<td>sig.</td>
<td>0.000</td>
<td>0.003</td>
<td>0.210</td>
</tr>
</tbody>
</table>

6. Discussion

All the hypotheses outlined in this paper are unsupported. Gender diversity does not necessarily induce significant or positive higher group productivity. Table 1 shows that the [MM] pair category had the highest mean value in both code productivity and code design categories, [MF] the second, and [FF] the last with least values. With this result, it appears that the males are superior in computer programming tasks and females less so. But statistically, this result is not significant. In other words, this result can be different in other instances.

This indicates that there is no “significant” gap in cognitive dispositions between men and women in computer programming.

Another intriguing observation is that female presence increases as the mean value gets lower, where [FF] pair category managed the lowest mean value. This “lower-but-not-significantly-lower” result on the female gender variable may provide an ecological fallacy to gender-biased debate. Moreover, this fallacy is confirmed by previous studies [11], [46]. Many researchers despise the current classroom setting because it is a male-biased environment. Through changes and enhancements in the classroom, alternative pedagogical approaches, and knowledge deliveries, females would show equal performance and capacity in areas, such as computer programming.

This gender difference in the questionnaire result demonstrates a more significant development than the quantitative analysis. Table 5 showed the significant differences in the levels of pair communication and pair compatibility. To elucidate this, Table 6 reports both [MM] and [FF] show a significantly higher mean (p < 0.01) than [MF], and no significance in means (p = 0.144) between [MM] and [FF].

Similarly, Table 7 reports both [MM] and [FF] show a significantly higher mean (p < 0.05) than [MF], and [FF] do not show a significant difference from [MM].

These results indicate that the homogeneous pairs find it easier working with one another, whereas the heterogeneous pairs have a more difficult time collaborating with a partner of the opposite gender during pair programming.

We also received supplementary comments from the participants, mostly from the female participants. They offered a subjective perspective about their personal experience within this mixed-gender PP experiment. Here are few examples of comments from the female participants in [FF] pairs;

“My partner was of the same sex (we are both female). I think this really helped us because it made us feel more comfortable around each other. I really think that if I had worked with a guy, there would have been some initial awkwardness. Whether it is counter-productive or not, this would be depend on the person who comes with a background that doesn’t respect woman’s opinion”.

“I think it would have been more difficult if I were paired with a male instead. Some men like to monopolize a project like this. Also, some males make assumptions about your level before they get to see first-hand so I was happy to get paired with a female”.

Following are a few more examples of comments from the female participants in [MF] pairs;

“my partner was male, and he tended not to explain what he was thinking or doing which was sometimes frustrating.”

“I think gender caused us to be a little more reserved around each other. If I programmed with a female I probably would have been more open.”

No such comments were issued from any of the male participants in all of the pair categories.

These comments and the qualitative results duplicate the perpetual gender-confusion, “different-but-not-so-different”. Men and women are not significantly different in their cognitive problem solving ability, but they are significantly different in their psychological experience and in their perceptions in the gender-specific cognitive collaboration.

In a situation like PP, most females experience, a high level of anxiety, a lack of confidence, and underachievement [22], [40]. The “difficulties” that the female participants expressed were overtly amplified with the male partners, whereas they registered absolutely no difficulty at all with the female partners.
One plausible explanation is that the two female pair partners sympathize and support each other, leading them to leverage the problem [46]. This reciprocal support may be implicit and naturally understandable to each partner, without the need for explicit explanation and communication, as with a male partner. Additionally, this intimacy and solidarity may only occur with the female pairs and not with female-male or male-male pairs [1], [6], [24], [33], [46].

With the female-male pair, the demanding programming environment [8] makes it difficult for the female to comfortably convey or implicitly manifest her feelings to the male partner, because of a fear of disapproval. This self-created “self-inferiorism” may hinder the positive pair programming experience and, further, perpetuate the endogenous negative attitude toward future pairing with a male partner.

With the male-male pair, their high-level of perceived communication and compatibility may largely depend on their similar analytical and administrative conducts [16], [17], [21]. They are not necessarily sharing or sympathizing with their partner’s psychological or emotional experience or feeling, but rather agreeing with each other on quick, decisive, and logical programming decision-making without much discussion or change in decisions [36]. With the female-male pair, this may not be accomplishable. Moreover, the males are much less expressive (e.g. no comments from any male participants) even though they had a significantly more positive experience with same gender partner.

All these results direct us to another round of research. In the next round of study, our explanations must be tested and validated. One question to ask is whether or not females’ predisposed gender-biased views affect their pair programming effort and experience.

This study makes significant research findings: 1) there is no significant gender difference in cognitive disposition under a pair programming context; 2) there are significant differences between same gender pairs and mixed gender pairs in PP communication and compatibility; 3) The female participants subjectively express more negative gender-biased views about computer programming or other similar technology-related tasks.

In terms of the limitations of this study, we point to: 1) the scant sample of female-female pairs, and 2) the participants’ lack of substantive professional programming experience. Having just 6 pairs of [FF] is not a substantial sample size. This jeopardizes the validity and conclusion of the experiment’s findings on female-female pair category. Also, the participants’ lack of substantive professional programming experience also risks the generalizability of this experiment’s result. The participant’s limited programming knowledge and skills may have interfered with or undermined the experiment’s objective. Considering these limitations, one must be cautious in interpreting and applying the experiment’s result.

7. Conclusion

The study reports that there are no significant gender differences in computer programming or problem solving ability, but there are significant gender differences in pair communication and compatibility. The same gender pairs, male-male and female-female, showed significantly higher pair communication and compatibility levels. A notable finding is that many female participants’ had commented about the difficulties and perceived difficulties in pairing with male partners while undertaking pair programming.

For the research implications of this study, it would be worthwhile to investigate the overall group impact. The peculiar findings of this study may play a certain degree of underlying influence on the overall group cohesiveness.

For the practical implications, this study presents a useful guideline in scheduling pair programming. If possible, achieving a higher level of pair programming output and team morale is preferred. Also certain software development sub-tasks may require good team chemistry in order to achieve a goal. This study’s result may serve as a basis for such a task.

Besides the productivity issue, if group harmony and morale are significant factors, then gender differences must be considered and appropriately managed. Gender related studies would enlighten agile software development, which cherishes disparate human values.

8. References


