





# Exploring Gender Bias in Remote Pair Programming among Software Engineering Students: The `twincode` Original Study and First External Replication

Amador Durán<sup>1</sup> , Pablo Fernández<sup>1</sup> , Beatriz Bernárdez<sup>1</sup> , Nathaniel Weinman<sup>2</sup>,  
Aslıhan Akalın<sup>2</sup>, and Armando Fox<sup>2</sup> 

<sup>1</sup> SCORE Lab, Universidad de Sevilla (US), Seville, Spain

<sup>2</sup> Computer Science Division, University of California, Berkeley (UCB), USA  
{amador, pablofm, beat}@us.es,  
{nweinman, asliakalin, fox}@berkeley.edu

**Keywords:** Gender Bias · Pair Programming · Remote Pair Programming · Software Engineering Education · Experiment Replication

**Published in:** Empirical Software Engineering, Vol. 29, No. 40, pp. 1–44, 2024

**Impact Factor:** JCR 2022 4.1 - Q2 - Position: 29/108 - Area: CS/SE

**DOI:** <http://dx.doi.org/10.1007/s10664-023-10416-6>

**Abstract.** *Context.* Women have historically been underrepresented in Software Engineering (SE), partly due to the gender bias that men outperform women at programming. Pair programming (PP) is widely used and has been shown to increase student interest in SE, particularly among women. However, if those same gender biases are also present in PP, its potential for attracting women to the field could be thwarted. *Objective.* We aim to explore the effects of gender bias in PP. Specifically, in a remote setting in which students cannot directly observe the gender of their peers, we study whether the perception of the partner, the behavior during programming, or the style of communication differ depending on the perceived gender of their remote partner. To our knowledge, this is the first study on the impact of gender stereotypes and bias *within* pairs in PP. *Method.* We have developed an online pair-programming platform (`twincode`) that provides a collaborative editing window and a chat pane, both of which are heavily instrumented. Students in the control group had no information about their partner's gender, whereas students in the treatment group could see a gendered avatar representing the other participant as a man or as a woman. The gender of the avatar was swapped between programming tasks to analyze 45 variables related to the collaborative coding behavior, chat utterances, and questionnaire responses of 46 pairs in the original study at the US, and 23 pairs in the external replication at the UCB. *Results.* We did not observe any statistically significant effect of the gender bias treatment, nor any interaction between the perceived partner's gender and subject's gender, in any of the 45 response variables measured in the original study. In the external replication, we observed statistically significant effects with moderate to large sizes in four of the 45 dependent variables within the experimental group, comparing how subjects acted when their partners were represented as a man or a woman. *Conclusions.* The results in the original study do not show any clear effect of the treatment in remote PP among current Software Engineering students. In the external replication, it seems that students delete more source code characters when they have a woman partner, and communicate using more informal utterances, reflections and yes/no questions when they have a man partner, although these results must be considered inconclusive because of the small number of subjects in the replication, and because when multiple test corrections are applied, only the result about informal utterances remains significant. In any case, more mixed methods replications are needed in order to confirm or refute the results in the same and other SE students populations.

