Empirical Software Engineering (2024) 29:161
https://doi.org/10.1007/s10664-024-10550-9

Understanding and effectively mitigating code review anxiety

Carol S. Lee[1] · **Catherine M. Hicks**[1]

Accepted: 23 September 2024 / Published online: 5 October 2024

Abstract
Anxiety about giving and receiving code reviews has been documented as a common occurrence that leads to developers avoiding code reviews by procrastinating and limiting their cognitive engagement with them. This avoidance not only increases anxiety in the long term, but also prevents developers, their teams, and their organizations from accessing the technical and sociocognitive benefits of effective and efficient code reviews. However, software research has not yet empirically examined code review anxiety, and from this, tractable intervention targets and strategies for mitigating code review anxiety. In this study, we present an empirical framework for understanding the factors maintaining and exacerbating code review anxiety. Utilizing a randomized waitlist control trial, we also tested the effectiveness of a novel single-session cognitive-behavioral workshop intervention. Our results show evidence that positive impact can be obtained from a brief intervention and suggest code review anxiety can be successfully mitigated by targeting developers' cost bias, anxiety self-efficacy, and self-compassion.

Keywords Code review · Anxiety · Software engineering · Developer experience

1 Introduction

Healthy and efficient code review practices are central to modern technology development (Kudrjavets and Rastogi 2024). Code reviews are a software engineering practice, during which developers manually inspect for defects and provide feedback on peers' code (Ackerman et al. 1984, 1989). This process has not only been shown to increase defect finding, thus increasing code quality and codebase security (Bacchelli and Bird 2013; Shull and

Communicated by Meiyappan Nagappan.

✉ Carol S. Lee
carol-lee@pluralsight.com

Catherine M. Hicks
cat-hicks@pluralsight.com

[1] Developer Success Lab, Pluralsight, Draper, UT, USA

Seaman 2008), but has also been shown to provide long-term sociocognitive benefits to software teams, such as moments of learning and knowledge transfer (Bacchelli and Bird 2013; Bosu and Carver 2013; Cunha et al. 2021), collaborative and creative problem solving (Bacchelli and Bird 2013; Wurzel Gonçalves et al. 2023), trust (Bosu and Carver 2013), and community building (Bosu and Carver 2013). To reap these benefits, research indicates that developers must actively participate in code reviews and commit to reviewing code with speed and accuracy. As such, a significant body of software research has been devoted to identifying empirically supported methods of increasing the speed and accuracy of code reviews, such as testing code locally, using appropriate code review tools (Söderberg et al. 2022; Wurzel Gonçalves et al. 2023), creating smaller pull requests (Baysal et al. 2016), having a positive relationship with peers (Wurzel Gonçalves et al. 2023), or integrating more automated or human-AI feedback loops to reduce the cognitive load and time burden placed on reviewers (Bird et al. 2022; Li et al. 2022).

However, less attention has been given to the psychological factors influencing developers' active participation in code reviews. One such factor is code review anxiety, characterized by a fear of judgment, criticism, and negative evaluation while giving or receiving code reviews. Although code review anxiety has not yet been empirically examined, it has been acknowledged as a common occurrence in industry commentary that leads to the avoidance of code reviews (Calafell 2019; Cohen et al. 2013; Moss 2018; Piper 2023). Additionally, anxiety about social and evaluative processes like code reviews have been frequently referenced as a blocker for developers in advancing their careers (Darmuki et al. 2017; Hazzan and Har-Shai 2014; Nazligul et al. 2017; Peters and Moreno 2015; Vitasari et al. 2011). As giving code reviews includes the obligation to uphold code quality for reviewers and also the pressure to provide timely and expert feedback (Kudrjavets and Rastogi 2024), and frequently raises the possibility of interpersonal conflict for both givers and receivers of code reviews (Wurzel Gonçalves et al. 2023), anxiety may cause difficulty no matter which role a developer currently occupies in a given code review.

Given that anxiety has been consistently shown to increase avoidance, inhibit learning and knowledge transfer, and reduce effective collaboration, these findings are not only unsurprising, but also point to the importance of addressing code review anxiety to improve the technical and sociocognitive outcomes of developers and their organizations (Barlow 2002). Despite this, the literature to date has neither empirically examined how code review anxiety is maintained and exacerbated, nor provided guidance on how to reduce code review anxiety in an evidence-based way. One key reason for this gap in the research may be the difficulty of studying clinical topics within professional software teams, as ethically studying experiences such as anxiety may require integrating clinical expertise, health and research ethics, and human subjects research standards. However, the pervasive reporting of code review anxiety highlights the need to integrate clinical perspectives and standards into software research, so that we can improve developers' experiences in a holistic and human-centered way.

To address this gap, the present study takes a clinical perspective on code review anxiety. Given the shared symptoms of a fear of judgment, criticism, and negative evaluation between code review anxiety and social anxiety (Heimberg and Becker 2002; Moss 2018) the present study's purpose was to: (1) adapt empirically validated models of social anxiety to develop a model of code review anxiety and (2) adapt scientific best practices in clinical research to a software development context to first develop, then assess the effectiveness

of a novel evidence-based code review anxiety workshop intervention using a randomized waitlist control trial.

The major contributions of this study are: (1) a novel model of code review anxiety that bridges software and clinical research to provide a deeper understanding of how software practitioners and researchers can better target and mitigate code review anxiety and (2) A single-session code review anxiety intervention that has been developed based on clinical best practices and empirically tested for its effectiveness that software practitioners and researchers can use to reduce code review anxiety. [1]

2 Related work

2.1 Code review anxiety

From a clinical perspective, the symptoms and experiences of code review anxiety are similar to social anxiety, in that they both share a set of core symptoms: a fear of judgment, criticism, and negative evaluation (Heimberg and Becker 2002; Moss 2018). As such, one approach to developing an empirical model of code review anxiety may be to use current models of social anxiety as a guide. According to these models, social anxiety is maintained and exacerbated by negative feedback loops that reinforce biased thinking and avoidance in social or performance situations (Clark and Wells 1995; Hofmann 2007; Rapee and Heimberg 1997). In particular, individuals experiencing social anxiety are less likely to believe that they have the skills to manage their anxiety (low anxiety self-efficacy) and are more likely to overestimate the cost and probability of situations ending poorly (high cost and probability bias) – all of which contribute to greater avoidance behaviors such as procrastinating, mentally "checking out," or prematurely leaving social and evaluative situations.

In the context of code review anxiety, developers experiencing code review anxiety may similarly experience low anxiety self-efficacy, high cost and probability biases, and high avoidance, as these symptoms are often alluded to in the grey literature (Clark and Wells 1995; Hofmann 2007; Moss 2018; Rapee and Heimberg 1997). For example, they may believe that they are unable to manage their code review anxiety (low anxiety self-efficacy), believe that they are likely to break production (probability bias), and believe that this will be the end of their career as an engineer (cost bias), all of which increases their code review anxiety. Developers may subsequently procrastinate on code reviews and limit their cognitive engagement, collaboration and receptiveness to feedback (e.g. for code review reviewers by "rubber stamping" changes, and for recipients of code reviews by skimming through feedback quickly instead of thinking about how they can learn from the feedback) as they "check out" to reduce their anxiety in the moment (avoidance).

Clinical anxiety disorder research has argued that effective change can occur from either directly modifying irrational beliefs, or from deactivating these irrational beliefs while making other, more positive beliefs more available to an individual experiencing anxiety (Hofmann 2007). For instance, a high probability bias may lead a developer to assume that negative outcomes are more likely than positive outcomes from a code review, and assume that other developers are inherently critical of anyone submitting a code review and are likely to view them negatively, an irrational belief which is heightened by the overall

1 Pre−registered hypotheses, methods, and materials are available at: https://osf.io/dq4fs/.

ways in which social anxiety heightens competitive elements of social interactions, while diminishing people's awareness of cooperation and support (Leary and Kowalski 1995; Trower and Gilbert 1989). High estimated social cost may emerge for a developer serving as a code reviewer in the form of irrational beliefs that others expect them to understand every technological domain area, a distressing expectation that may be increased by the way that social anxiety diminishes the perception of one's social skills and abilities (Hofmann 2007). In such a state, even developers with high technical proficiencies may doubt their ability to communicate well, maintain social bonds, and productively resolve conflict, a necessary part of making progress in collaborative code work—indeed, clinical research on anxiety has argued that performance *capability* does not serve as a key factor in many anxiety disorders (Barlow 2002). Based on the evidence across anxiety disorders in other settings, developers facing code review anxiety are likely to have the same range of skill-based competencies as developers without code review anxiety but find that anxiety inhibits their execution of those skills; developers then appraise their own performance more negatively than developers without code review anxiety, leading to further negative outcomes (Hofmann 2007).

Code review anxiety is also likely to be a persistent disorder, maintained and exacerbated by rumination, avoidance and "safety behaviors," behaviors which are intended to reduce distress during an anxious experience and serve to obscure that distress from others (Alden and Bieling 1998; Hofmann 2007; Barlow 2002). For instance, a developer with code review anxiety may engage in post-event rumination, mentally reviewing their past code reviews in detail while centering on negative self-perceptions. In this rumination cycle, code review anxiety may cause a developer to continually cast their own capability as a developer and experience with others on their team as more negative than reality and engage in negative anticipation of their next code review.

2.2 Single-session cognitive-behavioral interventions

To break the negative feedback cycle of factors maintaining and exacerbating social anxiety, current "gold standard" clinical interventions for social anxiety emphasize taking a cognitive-behavioral approach (David et al. 2018; Heimberg and Becker 2002). This approach emphasizes teaching individuals about the model for how anxiety occurs and is maintained and introducing techniques for cognitive restructuring such as identifying negative cognitions, known as automatic thoughts. Given the similarities between code review and social anxiety, it is likely that a cognitive-behavioral intervention would be similarly effective in reducing code review anxiety. For instance, a developer may practice recognizing that before every code review, they experience the negative automatic thought that their teammates are looking for reasons to fire them based on the quality of their code.

Cognitive-behavioral approaches elicit change through three key processes: (1) increasing awareness of internal experiences via psychoeducation, self-monitoring, and relaxation techniques, (2) reducing biased thinking and increasing constructive and compassionate thinking through a process called cognitive restructuring, and (3) reducing avoidance via exposure to feared stimuli (Craske 2017).

Notably, single-session interventions targeting even just one of these three key processes have been shown to have positive outcomes (Bertuzzi et al. 2021; Cougle et al. 2020; Hindo and Gonzalez-Prendes 2011; Hayes-Skelton and Lee 2020). Single-session interventions

have also been shown to be effective at scale, particularly when implemented as a single-session workshop-style intervention - a method that reduces barriers to access by allowing larger groups of participants to meet at once. Additionally, single-session workshop interventions for anxiety have been shown to have positive outcomes when delivered in-person, virtually, and asynchronously (Brown et al. 2011; Cukrowicz and Joiner 2007; Danitz and Orsillo 2014; Eustis et al. 2017; Lee et al. 2022; Muhomba et al. 2017; Panepinto et al. 2015; Rizvi and Steffel 2014; Uliaszek et al. 2016; Ward-Ciesielski et al. 2016).

As a measurable outcome for these interventions, alongside the decrease of cost bias and probability bias, treatment of anxiety has frequently focused on measuring increases in anxiety self-efficacy (Lee et al. 2022). In psychological science, self-efficacy refers to an individual's belief that they have the ability to achieve a goal, or complete a certain task (Bandura 1977). Anxiety self-efficacy assesses people's perceived ability that they can manage and tolerate their anxiety and is considered a powerful mediating psychological factor which can re-contextualize an individual's subjective experience. For instance, a developer experiencing code review anxiety but bringing anxiety self-efficacy to the forefront might recontextualize fearing difficult negative feedback on their code from a colleague as something that they have the skill and experience to manage and tolerate. In both clinical science and education research, enhancing individuals' self-efficacy is a well-studied intervention target, and there is large-scale evidence that such enhancement increases individuals' resilience and persistence across multiple positive outcomes, including academic achievement (Komarraju and Nadler 2013; Lent et al. 1984; Schwarzer and Luszczynska 2023). This large body of evidence positions self-efficacy as a productive preliminary outcome to target for mitigating code review anxiety.

Recent work on cultural competency in the treatment of clinical anxiety disorders has further emphasized that targeting self-compassion along with self-efficacy can validate people's authentic experiences with negative social interactions (Iwamasa and Hays 2019; Martinez et al. 2020). Self-compassion is a psychological construct which refers to a healthy way of relating to oneself during difficult and distressing circumstances, marked by multiple elements such as care towards self, recognition of shared humanity, mindfulness in the present moment, and reduction in self-judgment (Neff et al. 2021). For instance, a developer experiencing code review anxiety but bringing self-compassion to the forefront might recontextualize the difficulty of ensuring code quality as a shared goal on their team, bring greater mindful awareness to their own anxiety, and consider how to take a compassionate stance toward others during the code review collaboration. Greater self-compassion has been linked to greater overall well-being and positive outcomes such as optimism, as well as associated with lower rates of clinical depression (Neff et al. 2021; Zessin et al. 2015), and has been studied as a potential outcome measure for interventions (Neff and Germer 2013). This emerging body of evidence positions self-compassion as an important preliminary outcome to target for mitigating code review anxiety.

3 Methods

3.1 Hypotheses

Despite the prevalence and clear negative impact of code review anxiety in software practitioners' accounts of their experience, software teams currently lack an empirical understanding of code review anxiety that would allow them to identify intervention targets and strategies for mitigating it. Given the similarities between code review anxiety and social anxiety, it is possible that, like social anxiety, code review anxiety is maintained and exacerbated by low self-efficacy, high probability and cost biases, all of which maintain and exacerbate avoidance during code reviews. As such, we hypothesized that:

- H1. Self-efficacy, cost bias, and probability bias would maintain and exacerbate code review anxiety, with lower anxiety self-efficacy, higher cost bias, and higher probability bias being associated with higher code review anxiety.
- H2. Self-efficacy, cost bias, probability bias, and code review anxiety would maintain and exacerbate avoidance, with lower anxiety self-efficacy, higher cost bias, higher probability bias, and higher code review anxiety being associated with greater avoidance during code reviews.

Additionally, given the positive outcomes of single-session cognitive-behavioral interventions for anxiety overall, it is possible that a single-session cognitive-behavioral intervention for code review anxiety is an effective way to reduce code review anxiety and its associated mechanisms (self-efficacy, probability bias, and cost bias) in software developers. Further, given that cognitive behavioral interventions elicit change by increasing compassionate thinking (Heimberg and Becker 2002), it is likely that they are also an effective way to increase self-compassion. Using a randomized waitlist control design, we assessed the effectiveness of a single-session cognitive-behavioral workshop intervention for code review anxiety delivered synchronously online by a clinically trained facilitator compared to a waitlist control condition and hypothesized that:

- H3. Controlling for baseline measures, participants in the workshop group would experience lower code review anxiety, greater anxiety self-efficacy, lower cost bias, lower probability bias, and greater self-compassion, compared to those in the waitlist control group at follow-up.

3.2 Recruitment and participants

Prior to recruitment, all study protocols and materials were reviewed for ethical standards in human subjects research and approved by an IRB. The study was advertised as a "Code Review Anxiety Workshop Research Study." We recruited participants by posting publicly on various social media (e.g. X, facebook, mastodon, linkedin, and reddit) from researchers' personal social media accounts and via direct emails and slack messages to professional listservs of interest to developers. The study was open to all professional developers who experience code review anxiety. For this study, we defined this eligibility as: a developer

experiences code review anxiety to the extent that they (1) have a subjective experience of anxiety in response to thinking about or engaging in a code review either as the giver or receiver of a code review (2) are motivated to alleviate that subjective experience by understanding and attempting to practice new skills to manage their experience with code reviews.

Prior to enrolling in the study, participants were screened for code review anxiety using the self-report Subjective Units of Distress Scale (SUDS; Wolpe 1990). Participants were invited to complete baseline measures and sign up for a workshop time slot if they endorsed at least minimal code review anxiety or distress (SUDS≥2; see Fig. 1). Time slots were randomly assigned to the workshop or waitlist control conditions. A detailed description of each condition can be seen in Sect. 3.2 and 3.5. Of the 109 participants that completed baseline measures, 50 were not able to make one of the workshop times. Our final sample consisted of 59 participants (workshop=30, waitlist control=29). As previous work has noted that software engineering research frequently fails to consider generalizability in its sampling (Baltes and Ralph 2022), effort was taken to ensure that developers recruited to the study represented many demographic categories and firmographic contexts. In order to reduce participant burden and decrease the possibility of selection bias against participants who did not wish to volunteer this information, however, participants were given the option to not report on any demographic or firmographic item. Per human subjects research and social justice and equity science standards around the necessity for transparency on sample characteristics as a means to assess representative sampling and generalizability (Baltes and Ralph 2022; Call et al. 2023; Roberts et al. 2020), the full demographic and firmographic characteristics of our sample can be seen in Tables 1 and 2.

3.3 Code review anxiety workshop

The 120-minute code review anxiety workshop was delivered synchronously online by the first author through Zoom. We offered six workshops in total over the course of four weeks, spread across morning, afternoon, and evening times to increase the availability of workshop times across time zones. To maximize participant engagement, groups were also limited to 25 participants per group and consisted of both didactic presentation and hands-on practice. Group sizes for each session ranged from 8 to 14 participants (M=10.83, SD=2.56).

We developed the workshop by adapting content from the Cognitive-Behavioral Group Therapy (CBGT) for social anxiety disorder (Heimberg and Becker 2002) and Dialectical Behavior Therapy (DBT) manuals (Linehan 2014). All workshop materials can be found at https://osf.io/dq4fs. CBGT and DBT are both empirically supported cognitive behavioral approaches. Notably, CBGT is considered to be the "gold standard" treatment for social anxiety disorder, and DBT is considered to be a "gold standard" protocol for increasing

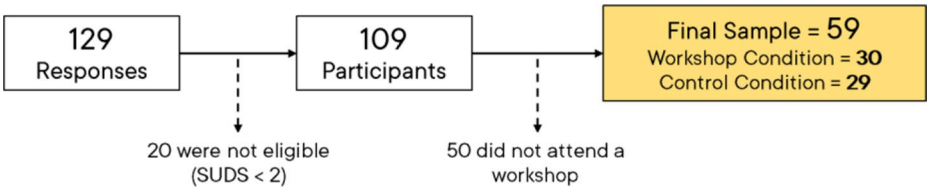

Fig. 1 Flowchart of participant drop out and final sample size

Table 1 Demographic characteristics

Demographic	Statistic		
	Control (n=29)	Workshop (n=30)	Overall (N=59)
Gender (n=59)			
Female	9 (31.0%)	9 (30.0%)	18 (30.5%)
Male	19 (65.5%)	18 (60.0%)	37 (62.7%)
Nonbinary/Fluid/Queer/Gender Queer	1 (3.4%)	2 (6.7%)	3 (5.1%)
Prefer not to answer	0 (0%)	1 (3.3%)	1 (1.69%)
Transgender (n=57)			
Yes	1 (3.4%)	2 (6.7%)	3 (5.1%)
Prefer not to answer	2 (6.9%)	0 (0%)	2 (3.4%)
Sexual Orientation (n=55) [1, 2]			
Asexual / Aromantic	1 (3.4%)	1 (3.3%)	2 (3.4%)
Bisexual	0 (0%)	2 (6.7%)	2 (3.4%)
Lesbian	1 (3.4%)	0 (0%)	1 (1.7%)
Pansexual	0 (0%)	2 (6.7%)	2 (3.4%)
Queer	1 (3.4%)	0 (0%)	1 (1.7%)
Questioning/ Unsure	1 (3.4%)	0 (0%)	1 (1.7%)
Straight/ Heterosexual	22 (75.9%)	20 (66.7%)	42 (71.2%)
Prefer not to answer	1 (3.4%)	5 (16.7%)	6 (10.2%)
Racial Identity (n=57) [1, 3]			
Black/African American	3 (10.3%)	3 (10.0%)	6 (10.2%)
East Asian	1 (3.4%)	4 (13.3%)	5 (8.5%)
Middle Eastern/North African (White)	1 (3.4%)	1 (3.3%)	2 (3.4%)
Latine/Hispanic (White)	3 (10.3%)	3 (10.0%)	6 (10.2%)
South/South-East Asian	4 (13.8%)	0 (0%)	4 (6.8%)
White	16 (55.2%)	18 (60.0%)	34 (57.6%)
Self-Identify [4]	1 (3.4%)	1 (3.3%)	2 (3.4%)
Prefer not to answer	0 (0%)	1 (3.3%)	1 (1.7%)
Country of Residence (n=57)			
Argentina	1 (3.4%)	0 (0%)	1 (1.7%)
Benin	0 (0%)	1 (3.3%)	1 (1.7%)
Canada	5 (17.2%)	5 (16.7%)	10 (16.9%)
Germany	0 (0%)	1 (3.3%)	1 (1.7%)
India	0 (0%)	1 (3.3%)	1 (1.7%)
Kenya	1 (3.4%)	0 (0%)	1 (1.7%)
Lithuania	1 (3.4%)	0 (0%)	1 (1.7%)
Republic of the Congo	1 (3.4%)	0 (0%)	1 (1.7%)
Sweden	0 (0%)	1 (3.3%)	1 (1.7%)
United Kingdom	0 (0%)	2 (6.7%)	2 (3.4%)
United States	19 (65.5%)	18 (60.0%)	37 (62.7%)
Education (n=57) [5]			
2-Year College	1 (3.4%)	1 (3.3%)	2 (3.4%)
4-Year College	16 (55.2%)	14 (46.7%)	30 (50.8%)

Table 1 (continued)

Demographic	Statistic		
	Control ($n=29$)	Workshop ($n=30$)	Overall ($N=59$)
Graduate Degree	7 (24.1%)	10 (33.3%)	17 (28.8%)
Some College	3 (10.3%)	5 (16.7%)	8 (13.6%)

[1] Participants could check all that apply

[2] Identities not endorsed: fluid, gay

[3] Identities not endorsed: Alaskan Native/ Native American/ Indigenous; Middle Eastern/ North African (non-white), Latine/ Hispanic (non-white), Pacific Islander/ Native Hawaiian, Multiracial

[4] Examples include: Scottish, Afro-caribbean

[5] Degrees not endorsed: grade school, some high school, high school diploma

awareness of internal experiences and engagement during interpersonal situations (Craske 2017). We deliberately chose content to directly target all three key processes of cognitive-behavioral approaches: (1) increasing awareness of internal experiences, (2) reducing biased thinking and increasing constructive and compassionate thinking, and (3) reducing avoidance. For example, to increase awareness of internal experiences, the workshop began with psychological education on the prevalence, symptoms, and function of code review anxiety, followed by a self-monitoring and functional analysis exercise of anxiety symptoms. Participants also learned and practiced the "TIPP (changing body temperature, intense exercise, paced breathing, and progressive muscle relaxation)" relaxation techniques.

To reduce biased thinking and increase constructive and compassionate thinking, participants then learned and practiced identifying and restructuring negative automatic thoughts - negatively biased thoughts about ourselves and a situation's outcomes that involuntarily occur. For example, a developer with the thought "if it doesn't work as expected, they'll think I'm stupid" might practice restructuring the thought into "things don't work as expected all the time. It doesn't mean someone is stupid," by asking themselves questions like "Can I read minds? Do I *know* that they'll think I'm stupid? Does something not working mean that someone is stupid? Am I putting unrealistic expectations on myself that I wouldn't put on others?"

Finally, to reduce avoidance, participants learned adaptations of DBT's "DEAR MAN" and "GIVE FAST" skills (Linehan 2014), which teach methods of effective and active engagement during interpersonal situations, such as gently providing and asking for validation and specific feedback. Throughout the workshop, examples were given for applying these methods during both the processes of giving as well as receiving code reviews (See Table 2).

3.4 Outcome measures

Subjective Units of Distress Scale (SUDS). The SUDS (Wolpe 1990) is a single item rating of state anxiety and distress. Participants are asked to rate their anxiety and distress about a situation of choice. To measure state code review anxiety, we asked participants to rate their anxiety and distress when thinking about giving or receiving a code review using a 0 (totally relaxed) to 10 (worst anxiety ever experienced) Likert-type scale. The SUDS is an empirically validated measure that has been shown to have good convergent, discriminant,

Table 2 Firmographic characteristics

Firmographic	Statistic		
	Control (n=29)	Workshop (n=30)	Overall (N=59)
Years Code (n=59)	15.0 (16.8)	16.1 (17.4)	15.6 (17.0)
Engineering Area [1, 2] (n=59)			
Backend	12 (41.4%)	17 (56.7%)	29 (49.2%)
Frontend	6 (20.7%)	7 (23.3%)	13 (22.0%)
Full Stack	12 (41.4%)	16 (53.3%)	28 (47.5%)
Database Admin	4 (13.8%)	3 (10.0%)	7 (11.9%)
System Admin	4 (13.8%)	2 (6.7%)	6 (10.2%)
DevOps	8 (27.6%)	6 (20.0%)	14 (23.7%)
Site Reliability	5 (17.2%)	0 (0%)	5 (8.5%)
Other	5 (17.2%)	4 (13.3%)	9 (15.3%)
Industry (n=59)			
Education	6 (20.7%)	5 (16.7%)	11 (18.6%)
Energy	1 (3.4%)	0 (0%)	1 (1.7%)
Financial Services	1 (3.4%)	2 (6.7%)	3 (5.1%)
Government	2 (6.9%)	1 (3.3%)	3 (5.1%)
Healthcare & Pharmaceuticals	2 (6.9%)	3 (10.0%)	5 (8.5%)
Insurance	1 (3.4%)	1 (3.3%)	2 (3.4%)
Non-profit	2 (6.9%)	1 (3.3%)	3 (5.1%)
Retail/Consumer/e-Commerce	2 (6.9%)	1 (3.3%)	3 (5.1%)
Technology	12 (41.4%)	13 (43.3%)	25 (42.4%)
Industrials & Manufacturing	0 (0%)	1 (3.3%)	1 (1.7%)
Other	0 (0%)	2 (6.7%)	2 (3.4%)
Organization Size [2] (n=59)			
5–9	0 (0%)	1 (3.3%)	1 (1.7%)
10–19	3 (10.3%)	2 (6.7%)	5 (8.5%)
20–99	6 (20.7%)	6 (20.0%)	12 (20.3%)
100–499	5 (17.2%)	7 (23.3%)	12 (20.3%)
500-1,999	8 (27.6%)	5 (16.7%)	13 (22.0%)
2,000–4,999	2 (6.9%)	4 (13.3%)	6 (10.2%)
5,000–9,999	2 (6.9%)	1 (3.3%)	3 (5.1%)
10,000 +	3 (10.3%)	4 (13.3%)	7 (11.9%)

[1]Participants could check all that apply

[2]Firmographics not endorsed: mobile engineering area; 1–4 org size

concurrent, and predictive validity (Kim et al. 2008). Higher scores indicate greater code review anxiety. Participants completed the SUDS at baseline and follow-up.

Probability Bias Rating (PB). The PB is a single item rating of probability bias, adapted from the empirically validated Social Cost and Probability Questionnaire (SCPQ; Foa et al. 1996). To measure probability bias in code reviews, we asked participants to rate the likelihood of "something negative (e.g. feeling incompetent, being embarrassed, making a mistake)" occurring during a code review using a 0 (not at all likely) to 8 (extremely likely)

Likert-type scale. Higher scores indicate higher levels of probability bias. Participants completed the PB at baseline and follow-up.

Cost Bias Rating (CB). The CB is a single item rating of cost bias, adapted from the empirically validated Social Cost and Probability Questionnaire (SCPQ; Foa et al. 1996). To measure cost bias, we asked participants to rate how "bad" it would be "for something negative (e.g. feeling incompetent, being embarrassed, making a mistake) to happen during a code review" using a 0 (not at all) to 8 (extremely) Likert-type scale. Higher scores indicate higher levels of cost bias. Participants completed the CB at baseline and follow-up.

Anxiety Self Efficacy (ASE). The ASE (Lee et al. 2022) is a two-item self-report measure, assessing respondents' perceived ability to cope with and manage code review anxiety. Participants utilize a 1 (not at all) to 9 (very) Likert response scale. The items are adapted from the Self-Efficacy for Social Situation Scale (Gaudiano and Herbert 2003) and modified to be about code review anxiety coping and management skills (e.g.: "Is it possible for you to cope with and manage stress and anxiety well, despite any weaknesses you might have in anxiety coping and management skills?") rather than about social skills (e.g.: "Is it possible for you to perform well in social situations, despite any weaknesses you might have in social skills?"). Higher scores indicate higher code review anxiety self-efficacy. The measure has been shown to have good internal consistency (Lee et al. 2022). Similarly, the ASE had good internal consistency at both baseline ($\alpha = 0.80$) and follow-up ($\alpha = 0.78$) in our sample.

State Self Compassion Scale (S-SCS). The S-SCS is a six item self report measure (Neff et al. 2021) assessing respondents' state self-compassion using a 1 (not at all true for me) to 5 (very true for me) Likert scale. The S-SCS has been shown to have good reliability ($\alpha = 0.86$; Neff et al. 2021). Higher scores indicate greater self-compassion. Similarly, the S-SCS had good internal consistency at both baseline ($\alpha = 0.84$) and follow-up ($\alpha = 0.83$) in our sample.

Code Review Avoidance (CRA). The CRA is a two-item measure assessing the extent to which respondents behaviorally and experientially avoided code reviews over the past seven days. Respondents answered using a 0 (Never) to 4 (Always) Likert-type scale. The CRA was adapted from the empirically validated Severity Measure for Specific Phobia - Adult (SMSP; Craske et al. 2013) to be about code reviews and abbreviated to only include the two items loading strongest onto the avoidance factor of the SMSP. The CRA had acceptable internal consistency in our sample ($\alpha = 0.74$). Higher scores indicate greater avoidance. Participants completed the CRA at baseline.

3.5 Workshop reactions survey measures

Perceived Helpfulness Rating (PH). The PH (Lee et al. 2022) is a single item rating assessing the perceived helpfulness of the workshop using a 1 (Not at all) to 5 (Extremely) Likert scale, that has been shown to have good face validity. We used this item as a manipulation check for effective workshop facilitation across conditions. Higher scores indicate a greater perceived helpfulness. Participants completed the PH after completing the workshop.

Behavioral Action Rating (BA). The BA (Lee et al. 2022) is a single item rating assessing the likelihood of practicing the skills taught in the workshop. The item is adapted from the general behavioral action rating, which has been shown to have good face and predictive validity (Lee and Hayes-Skelton 2020). Higher scores indicate a higher likelihood of skills

practice and application. We used this item as a manipulation check for effective workshop facilitation across conditions. Participants completed the BA after completing the workshop.

3.6 Procedures

After providing informed consent, all eligible participants completed the baseline measures (SUDS, PB, CB, ASE, S-SCS, CRA, and demographics/firmographics). Upon completion of baseline measures, participants were allowed to select a time slot during which they would attend the workshop. To control for natural changes in anxiety over time (maturation) and time elapsed from baseline across time slots, participants were only able to sign up for time slots 4–7 days in advance. Time slots were then stratified by time and randomly selected to either be in the workshop or waitlist control conditions. This allowed us to control for any group differences at baseline and the impact of time zones and time slots on study outcomes.

Participants completed the follow-up measures 4–7 days after completing the baseline measures. Participants in the workshop condition completed the follow-up measures (SUDS, PB, CB, ASE, and S-SCS) *after* completing the workshop intervention. In contrast, participants in the waitlist control condition completed the follow-up measures *before* being offered the workshop intervention. Participants across conditions completed the workshop reactions survey (PH, BA) after completing the workshop intervention.

Since the purpose of this study was to assess the impact of our workshop intervention on code review anxiety, probability bias, cost bias, anxiety self-efficacy, and self-compassion, we asked participants to complete the SUDS, PB, CB, ASE, and S-SCS at both baseline and follow-up. Since the workshop reactions survey measures (PH and BA) asked participants about their reactions to the workshop itself, participants only completed these measures at follow-up. Because the CRA assesses code review avoidance over the last week, we did not examine avoidance as an outcome variable, as it would be impossible for participants to change their code review avoidance immediately after completing the workshop, without having had the opportunity to engage in a real-world code review. As such, participants only completed the CRA at baseline. Finally, to reduce participant burden at follow-up, demographics and firmographics were only collected once, at baseline. Participants completed the follow-up measures, the workshop, and the workshop reactions survey on the same day. Figure 2 summarizes the study design.

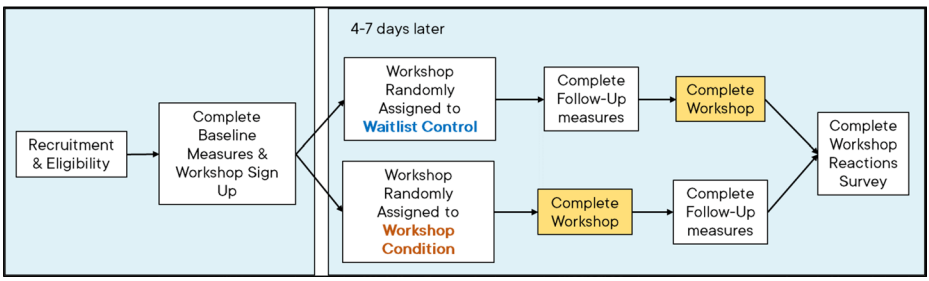

Fig. 2 Study design

3.7 Statistical analyses

All data were analyzed using *R* and a variety of *R* packages developed for statistical analysis (Fletcher 2022; Friendly et al. 2024; Wickham et al. 2019; R Core Team 2023; Revelle 2024; Rizopoulos 2022). We used Soper's sample size calculator (2024), to confirm that we were sufficiently powered, assuming a moderate effect size for anxiety reduction. To test the normality of our variables, we obtained skew and kurtosis values of all variables. All variables were within normal ranges. To identify whether any demographic and firmographic variables covaried with code review anxiety, we also conducted a series of ANOVAs and correlations between the SUDS and our demographic and firmographic variables. There were no significant effects ($ps = 0.27 - 0.74$); given this, we did not include any demographic or firmographic covariates in our model of code review anxiety. Zero order correlations indicated that all variables in our model were significantly correlated in the expected directions, for example negative correlations between ASE to PB, CB, and SUDS (see Table 3).

We created our model of code review anxiety by conducting a regression-based path analysis using the baseline scores of all participants in our sample. Path analyses allow us to examine the effect of an independent variable on a dependent variable, while accounting for the relationships between all of the independent variables in the model. A path analysis extends regression by modeling complex relationships between variables, including the direct and indirect effects of all independent variables, using linked regression equations. These equations are solved simultaneously to estimate the standardized and unstandardized path coefficients, which represent the strength and direction of the relationships between variables in the model. Standardized coefficients within a model can be compared to identify the strongest contributors to a dependent variable (Tabachnick and Fidell 2021). Variance inflation factor (VIF) checks were calculated at each step (*VIF*s = 1.34–1.89) and did not identify multicollinearity as a concern in our model.

A series of ANOVAs and chi-square analyses indicated baseline equivalence between conditions on demographic characteristics ($ps = 0.10 - 1$), firmographic characteristics ($ps = 0.06 - 1$), and baseline measures ($ps = 0.22 - 0.86$). Additionally, our workshop survey reaction measures, the PH and BA did not significantly differ by condition, serving as a manipulation check on effective workshop facilitation across conditions. We next assessed the outcomes of our workshop intervention using a MANCOVA, with the baseline scores of all outcome variables entered as covariates. We used a MANCOVA due to its ability to control for regression to the mean when assessing change over time in experimental designs.

Along with assessing statistical significance, we computed the reliable change index (RCI; Jacobson and Truax 1991) to examine clinically significant change. Clinical signifi-

Table 3 Zero-order correlations between code review anxiety model variables

Variable	PB	CB	ASE	SUDS	CRA
PB	1.00	0.50	-0.33	0.46	0.51
CB	—	1.00	−0.56	0.58	0.47
ASE	—	—	1.00	−0.58	−0.40
SUDS	—	—	—	1.00	0.60
CRA	—	—	—	—	1.00

Note. Scores reflect measures taken at baseline across all participants in our sample. PB = Probability Bias, CB = Cost Bias, ASE = Anxiety Self-Efficacy score, SUDS = Subjective Units of Distress Scale rating, CRA = Code Review Avoidance. All $ps < 0.001$

cance is a supplemental and additive approach to statistical significance that calculates the statistically reliable change of a measure over time, allowing practitioners to have further confidence that the amount of expected and reliable change from a clinical intervention is warranted given the potential costs of the intervention. Clinical significance is thus frequently used to identify the proportion of intervention "responders" who experience a large magnitude of change that noticeably impacts their daily lives (Ranganathan et al. 2015). Observed changes may be statistically significant, but not reach the higher bar of clinical significance. The RCI is computed by subtracting the pretreatment score from the posttreatment score, and dividing the sum by the standard error of measurement difference. If the absolute value of the RCI is greater than 1.96 (denotes 95% confidence), then the change in scores is considered to be statistically reliable and clinically significant (Jacobson and Truax 1991).

4 Results

4.1 Model of code review anxiety

To test H1 and H2, we conducted a regression-based path analysis. This allowed us to test each path of the model while accounting for all other variables in the model. The results indicated that, of the paths leading to the SUDS, the paths from the CB and the ASE were significant, while the path from the PB was not. Of the paths leading to the CRA, the paths from the SUDS and the PB were significant, while the paths from the CB and the ASE were not (See Fig. 3; Table 4). When comparing the standardized regression coefficients (β), our results also indicate that, of the cognitive mechanisms, anxiety self-efficacy was the strongest contributor to code review anxiety. Additionally, of the paths leading to avoidance, code review anxiety was the strongest contributor to avoidance.

4.2 Code review anxiety workshop intervention outcomes

To test H3, we conducted a MANCOVA. We entered condition as the independent variable (workshop versus waitlist control group), and the SUDS, PB, CB, ASE, and SCS-SF scores at follow-up as the dependent variables. To adjust for baseline scores, we entered the SUDS, PB, CB, ASE, and SCS-SF scores at baseline as covariates. The results indicated that, controlling for baseline measures, the combined dependent variables were significantly different between conditions with a large effect size ($F(5, 48)=6.42$, $p<.001$, $\eta^2 = 0.36$). Univariate analyses indicated that, controlling for baseline measures, participants in the workshop condition had significantly higher follow-up ASE scores with a large effect size, significantly higher follow-up S-SCS scores with a medium effect size, and significantly lower follow-up SUDS ratings with a medium effect size. In contrast, there were no significant differences between conditions on the PB or the CB (see Fig. 4; Table 5).

4.3 Clinically significant change

To further assess the magnitude of change in the SUDS, ASE, and S-SCS, we also examined clinically significant change in our sample using the RCI (Jacobson and Truax 1991). We

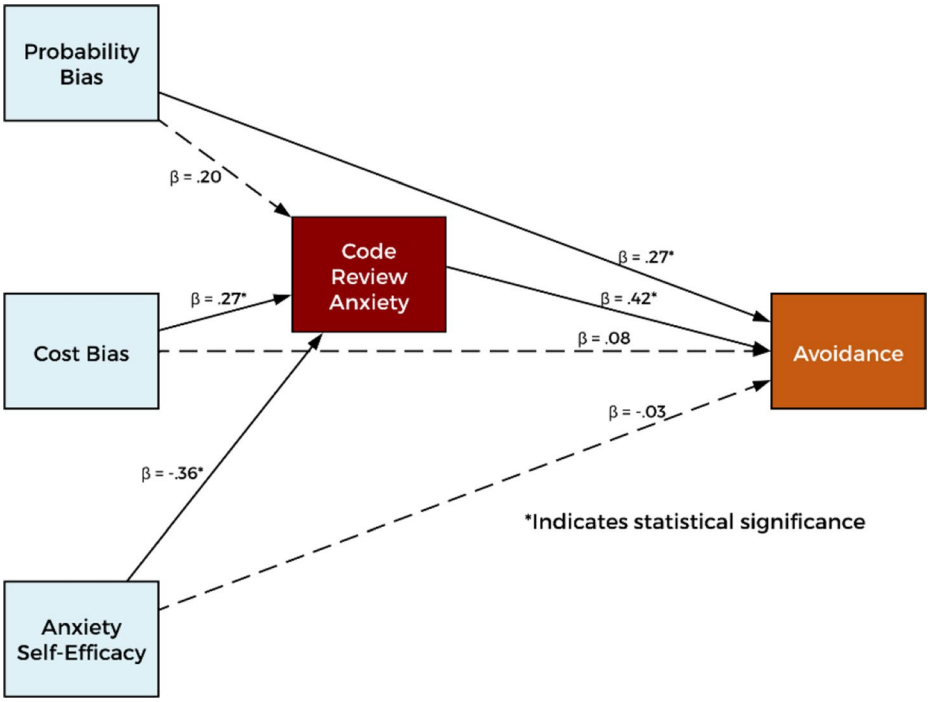

Fig. 3 Path analysis of model of code review anxiety. Standardized regression coefficients represented. Bold arrows indicate statistically significant paths; dashed arrows indicate nonsignificant paths

Table 4 Path analysis model of code review anxiety

Antecedent	Consequent					
	SUDS $R^2=0.43$, F (3, 55) = 15.61, $p<.001$			CRA $R^2=0.39$, F (4, 54) = 10.41, $p<.001$		
	b	β	95% CI	b	β	95% CI
PB	0.20	0.20	[-0.03, 0.43]	0.29*	0.27*	[0.02, 0.56]
CB	0.28*	0.27*	[0.01, 0.56]	0.09	0.08	[-0.23, 0.41]
ASE	-0.23**	−0.36**	[-0.38, -0.08]	-0.02	−0.03	[-0.21, 0.16]
SUDS	—	—	—	0.46**	0.42**	[0.15, 0.77]

Note. Scores reflect measures taken at baseline across all participants in our sample. PB = Probability Bias, CB = Cost Bias, ASE = Anxiety Self-Efficacy score, SUDS = Subjective Units of Distress Scale rating, CRA = Code Review Avoidance. * $p<.05$; ** $p<.01$

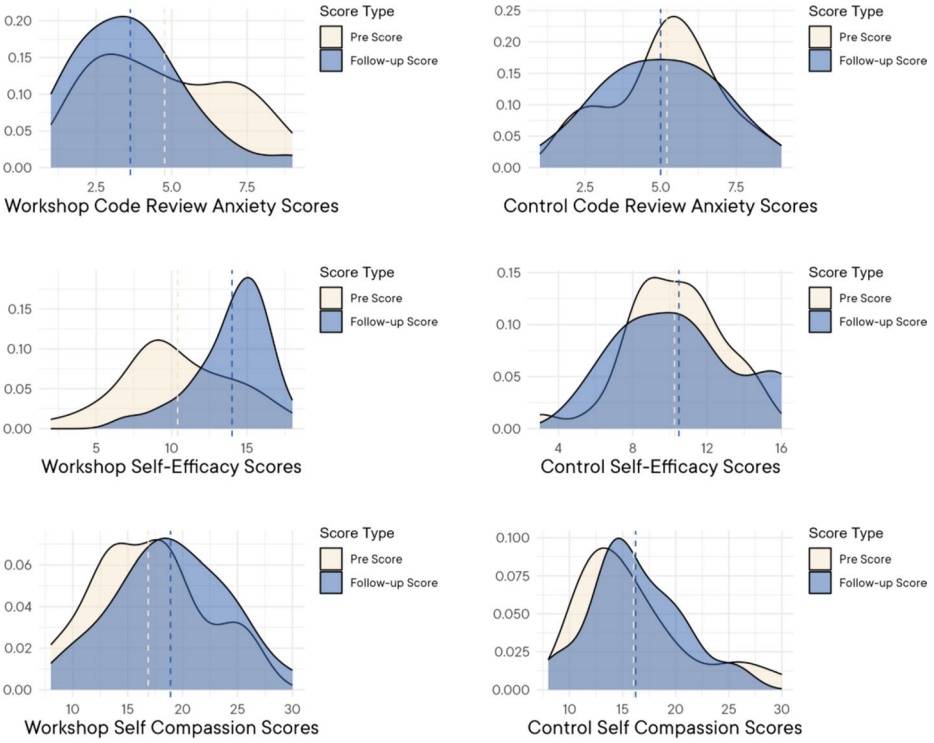

Fig. 4 Density plots of code review anxiety, self-efficacy, and self-compassion scores by condition at pre- and follow-up

Table 5 Outcome variables by condition

Variable	Baseline - Mean (SD)			Follow-Up - Mean (SD)			F (1, 58)	η^2
	Control (n=29)	Workshop (n=30)	Overall (N=59)	Control (n=29)	Workshop (n=30)	Overall (N=59)		
PB	4.45 (1.88)	3.80 (2.14)	4.12 (2.03)	3.97 (1.90)	3.60 (1.90)	3.78 (1.89)	0.73	0
CB	4.14 (1.60)	3.80 (2.25)	3.97 (1.95)	3.83 (1.87)	3.03 (1.75)	3.42 (1.84)	4.02	0.04
SCS	16.0 (5.26)	16.9 (5.05)	16.5 (5.13)	16.2 (4.32)	18.9 (5.03)	17.6 (4.85)	9.32*	0.10
ASE	10.2 (2.61)	10.4 (3.68)	10.3 (3.18)	10.5 (3.26)	14.0 (2.44)	12.3 (3.35)	30.80***	0.34
SUDS	5.21 (1.82)	4.77 (2.21)	4.98 (2.02)	5.00 (1.98)	3.63 (1.85)	4.31 (2.02)	11.32**	0.12

Note. PB=Probability Bias, CB=Cost Bias, ASE=Anxiety Self-Efficacy, SUDS=Subjective Units of Distress Scale, SCS=Self-Compassion Scale score. * $p<.05$; ** $p<.01$; *** $p<.001$

measured clinically significant change using the RCI. The findings indicated that a higher percentage of participants in the workshop condition experienced clinically significant change across all three measures (see Table 6).

5 Discussion

Healthy and efficient code review practices are central to modern technology development (Kudrjavets and Rastogi 2024). To obtain the technical and sociocognitive benefits of code reviews, developers must commit to actively engaging with, rather than avoiding, the code review process. Given that previous research implicates anxiety as a primary mechanism of avoidance and engagement, the present study aimed to (1) develop a model of code review anxiety and (2) adapt scientific best practices in clinical research to a software development context to first administer, then assess the effectiveness of a novel evidence-based code review anxiety workshop intervention.

5.1 Model of Code Review anxiety

Our path analysis found that developers were most likely to procrastinate, avoid, or disengage (e.g. "rubber stamping," skimming, or not reading requests or feedback) during code reviews when they experienced higher code review anxiety and believed that something negative would happen (probability bias; see Fig. 5). These findings partially support H2 ("lower anxiety self-efficacy, higher cost bias, higher probability bias, and higher code review anxiety would be associated with greater avoidance during code reviews"). Our path analysis also indicated that code review anxiety was the strongest contributor to avoidance. These findings are consistent with other models of anxiety, which highlight anxiety as the strongest predictor of avoidance, due to avoidance's role in temporarily reducing anxiety by removing or delaying the anxiety-provoking stimulus (Hofmann 2007; Rapee and Heimberg 1997). For example, in the case of code review anxiety, a developer who experiences a fear of judgment or negative evaluation about an impending code review can delay and temporarily remove their code review anxiety by avoiding or delaying the code review.

Notably, because avoidance can lead to negative long-term cognitive, emotional, social, and professional outcomes (Heimberg and Becker 2002), previous research has examined the effect of anxiety interventions on avoidance and found that reducing anxiety is one of the most effective ways to reduce avoidance behaviors and increase engagement with feared situations (Barlow 2002; Hofmann 2007; Lee et al. 2022; Rapee and Heimberg 1997). As such, it is likely that one of the most effective ways to reduce code review avoidance and

Table 6 Percent of participants achieving clinically significant change

Measure	*n* (%)	
	Control ($n=29$)	Workshop ($n=30$)
SUDS	17.24%	43.33%
ASE	6.9%	60%
SCS	3.45%	20%

Note. Anxiety Self-Efficacy score, SUDS = Subjective Units of Distress Scale rating (code review anxiety rating), SCS = Self-Compassion Scale score

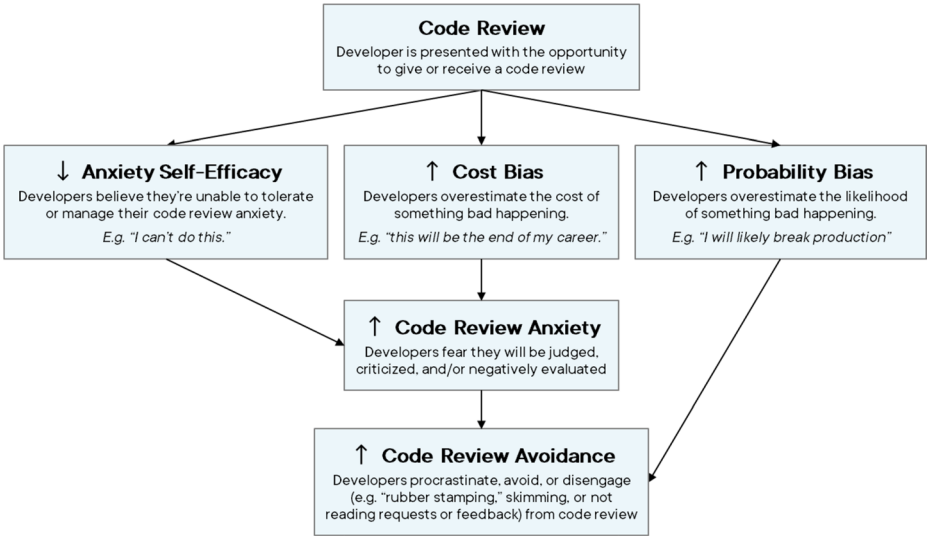

Fig. 5 Conceptual model of code review anxiety

increase code review engagement is to reduce code review anxiety. For example, a developer who learns to reduce their code review anxiety will not only reduce their code review avoidance, but also increase their code review engagement, leading to increased learning outcomes and peer support, a central aspect of large-scale professional software development. Further, the developers' increased code review engagement can subsequently enable colleagues, teams, and organizations to access the documented benefits of effective code reviews, including effective knowledge transfer (Bacchelli and Bird 2013; Bosu and Carver 2013; Cunha et al. 2021), trust (Bosu and Carver 2013), collaboration (Bacchelli and Bird 2013; Wurzel Gonçalves et al. 2023), code quality (Bacchelli and Bird 2013; Shull and Seaman 2008), and security (Bacchelli and Bird 2013; Shull and Seaman 2008).

Our path analysis also indicated that developers were more likely to experience code review anxiety when they overestimated the cost of a negative outcome (cost bias) and believed that they were unable to tolerate or manage their anxiety during a code review (anxiety self-efficacy; see Fig. 5), partially supporting H1 ("lower anxiety self-efficacy, higher cost bias, and higher probability bias would be associated with higher code review anxiety"). Low anxiety self-efficacy was the strongest contributor to code review anxiety in this model, suggesting that an effective means for reducing code review anxiety would be to increase developers' perceived ability to manage their anxiety. For example, a developer who experiences code review anxiety may experience the greatest reduction in code review anxiety if they can increase their confidence in their ability to manage their anxiety by learning new strategies for managing their anxiety and practicing existing ones. Taken together, the findings are consistent with previous research identifying anxiety self-efficacy as a predictor of anxiety reduction (Hofmann 2005, 2007; Lee et al. 2022).

Notably, our demographic and firmographic variables were not significantly associated with code review anxiety. In particular, code review anxiety was not significantly higher in individuals identifying as women, nonbinary, gender neutral, or gender queer (36.21% of our sample; $p = .57$). Code review anxiety was also not significantly associated with coding

experience ($M = 15.6$ years, $SD = 17.0$ years, range = 1–65 years; $p = .72$). Our nonsignificant findings here counter industry myths about code review anxiety being a "junior developer issue," as well as overgeneralizations about anxiety being more common in women (Remes et al. 2016).

Instead, the findings suggest that *any* developer can experience code review anxiety, an important point that can be used to normalize and openly discuss code review anxiety on software teams. Thus, understanding code review anxiety is relevant to *any* software team or organization, and investing in reducing it may be a powerful way to benefit *all* developers in a variety of circumstances. While seemingly surprising, these findings echo previous research on social and performance anxiety, indicating that its prevalence is consistent across genders, and that it impacts individuals regardless of their experiences and abilities (McLean et al. 2011; Norton and Hope 2001). Nevertheless, software research has also found that many developers systematically experience more punitive code reviews based on others' biased evaluations of personal attributes such as age, race, ethnicity and gender (Nadri et al. 2021; Murphy-Hill et al. 2022). These findings suggest that while it is important to see the common and shared experience of code review anxiety across different developers, more research is needed to understand how organizations and software teams can ensure their code review processes are equitably implemented.

5.2 Workshop outcomes

Our analysis of workshop outcomes indicated that, controlling for baseline scores, participants in the workshop condition experienced significant reduction in code review anxiety and significant increase in their self-efficacy to tolerate or manage their anxiety. Our findings here partially support H3, and suggest that a brief, clinically-informed code review anxiety workshop may be an effective way to reduce code review anxiety, and that such an intervention primarily works by targeting developers' self-efficacy to tolerate and manage their anxiety. Additionally, our RCIs provided evidence that this intervention was robust, with more developers in the workshop condition experiencing clinically significant changes in code review anxiety and anxiety self-efficacy. Although some participants in the control condition experienced clinically significant change, the proportion of participants experiencing clinically significant change is within expected limits, when considering maturation in a waitlist control design (Pusswald et al. 2019). This intervention efficacy was further supported by medium to large effect sizes in the observed workshop changes.

Overall, this evidence suggests that the effects from the workshop may not only improve developers' ability to face immediate code reviews but may also initiate large-scale positive changes that result in developers feeling noticeably better during code reviews. This is consistent with previous research indicating that single-session cognitive-behavioral workshop interventions can be highly effective in reducing stress and anxiety, and that increased self-efficacy is a key predictor of treatment outcomes (Goldin et al. 2012; Lee et al. 2022; Tahmassian and Jalali Moghadam 2011). It is also consistent with behavioral science theory for how brief targeted interventions which help people craft an adaptive narrative around their psychological experiences can create recursive, reinforcing cycles of positive change for workplaces (Brockner and Sherman 2019).

Interestingly, the results also indicated that, controlling for baseline scores, developers in the workshop condition experienced significantly more self-compassion about their code

review anxiety and any challenges they may face during code reviews. The RCI also indicated that developers in the workshop condition were more likely to experience clinically significant increases in self-compassion. Although increasing self-compassion was not an explicit target of the workshop, self-compassion has been shown to be a common mechanism of change across therapeutic interventions for anxiety, including advanced cognitive restructuring (Baer 2010; Mayasari et al. 2022). As such, these findings partially support H3 and provide evidence that the cognitive restructuring component of our workshop was effective in reducing code review anxiety, primarily by noticeably increasing developers' compassion towards their thoughts, experiences, and abilities. Beyond code reviews, our findings also suggest that interventions that help developers cultivate self-compassion and self-efficacy in the face of technological challenges, may prove a valuable route for software teams and researchers in understanding how developers navigate daily challenges in their workplace and surface actionable recommendations to help developers and their organizations. For example, self-compassion interventions have been shown to mitigate anxiety in a variety of settings (Neff et al. 2021), and software teams and researchers could take advantage of published self-compassion scales to amplify developers' experiences of software development processes and promote investment in healthy organizational practice.

Contrary to H3, the results indicated that, controlling for baseline scores, participants in the workshop condition did not experience significantly lower probability and cost biases - two common targets for cognitive restructuring. Our findings here likely reflect our choice to prioritize self-compassion and anxiety self-efficacy over the probability and cost biases during cognitive restructuring. These choices were made to create an adaptive intervention applicable to a wide range of software team contexts and to increase the cultural sensitivity of our intervention. For example, unlike panic or specific phobias, where feared negative outcomes are highly unlikely to occur (e.g. dying or having a heart attack), feared social or performance anxiety-related outcomes *can* occur and are often frequently experienced (e.g. making a mistake). Because of this, targeting a negative automatic thought like "I will make a mistake" by decreasing probability bias ("I probably won't make a mistake") can be contextually unhelpful, whereas targeting the thought by increasing self-compassion ("it's okay if I make a mistake because mistakes are a part of learning and everyone makes them") can be more broadly functional.

Similarly, because the negative valence of feared social or performance outcomes can be objectively damaging for individuals holding minoritized identities, targeting negative automatic thoughts such as "People will say something sexist about it" by decreasing cost bias ("it won't be that bad") can create avoidance and be invalidating, whereas targeting self-compassion and self-efficacy ("It's them being sexist, it doesn't mean there's something actually wrong with me. It's okay to feel hurt and I can do something to take care of myself.") can be validating and empowering (Iwamasa and Hays 2019; Martinez et al. 2020). Self-efficacy and positive self-concepts have also been put forward as key supportive mechanisms by the stereotype inoculation model (Dasgupta 2011), which proposes that contact with successful ingroup peers and ingroup experts can create a protective buffer for minoritized individuals in high achievement contexts. For software developers in our workshop, encouraging self-compassion in the face of difficulty with code reviews may help developers to contextualize these experiences as a shared human experience in software work, rather than evidence of their unique failure to be a skilled developer ("all developers make mistakes, and we invest in helping each other with reviews because this

is how we improve together"). Software developers, managers, and organizations could similarly reduce code review anxiety by encouraging self-compassion amongst themselves and each other – for example, by reminding one another that coding is a complex task that is impossible to do perfectly, that mistakes are normal and expected, and that there is always something new to learn in software development. Developers could further encourage self-efficacy, by openly discussing code review anxiety management strategies such as the strategies outlined in our workshop materials (publicly available at https://osf.io/dq4fs/) to promote new learning about anxiety management and increase developers' confidence in abilities to manage their code review anxiety.

5.3 Limitations and future directions

Our study findings should be considered in the context of several limitations and potential threats to validity (Vazire et al. 2022). First, our code review anxiety model is based on an observational study with cross-sectional data. Additionally, although the SUDS and S-SCS were empirically validated measures, the CB, PB, ASE, and CRA were adapted from empirically validated measures, which can threaten the construct validity of our study. This is common in research on emerging health related topics and is frequently used to build preliminary evidence to better understand a new area (Ambuehl and Inauen 2022). Although previous experimental research gives us confidence in interpreting this evidence to make recommendations for practitioners based on both known empirical data about anxiety and our new model for code review anxiety reduction, future longitudinal study designs could validate our adapted measures and contribute evidence to make stronger conclusions about the causality or directionality of effects. For instance, while this study contributes evidence about changing developers' immediately reported experience of code review anxiety, repeating these measures over a longer duration could test whether these effects persist over time and in the context of developers' daily work, adding to the ecological and external validity of the intervention evidence.

Relatedly, in order to reach developers in real workplaces who were actively experiencing code reviews and who were comfortable participating in the study topic, our observational study relied on participant opt-in and used convenience sampling. While we ensured that our participant pool represented a diverse range of developer identities and workplace contexts, and widely advertised our study to a broad range of channels, it is likely that the developers who voluntarily enrolled in this study were particularly interested in research and in taking action on their experience of anxiety, representing a potential selection bias and threat to external validity. Additionally, although we collected data on participants' coding experience, we did not collect data on participants' code review experience. This threatens the internal validity of our study, as code review experience may be a confound that impacts developers' code review anxiety and avoidance. Future experimental study designs could incorporate random sampling, include additional firmographics such as code review experience, and include larger sample sizes of firmographic and demographic characteristics in order to learn more about how developers experience code review anxiety and avoidance in different circumstances.

Second, our workshop intervention study design was unable to utilize a double-blind procedure, which can impact internal validity. As such, the workshop facilitator was aware if workshops were in the control or workshop conditions. While all personal efforts were made

to facilitate equivalently, a concern may be that the facilitator's awareness of the conditions led them to facilitate the workshops inequitably without realizing it (for instance, being more persuasive and engaged during the workshop condition). However, we included the workshop reactions survey as a validity check in the study design and were therefore able to validate that participants perceived both conditions as equally helpful and were just as likely to report planned behavioral action in the control conditions, suggesting that facilitation was equally engaging. Nevertheless, future research could achieve stronger confidence in controlling for any unmeasured effects of facilitator knowledge by testing the efficacy of a different intervention administered in a double-blind procedure.

Third, all workshops were facilitated by the first author, a clinical psychologist with experience providing cognitive-behavioral interventions to clinical populations. This particular characteristic of the facilitator represents a depth of training and expertise in presenting interventions, along with expertise in the subject matter of anxiety. Both conditions in this study were facilitated by a clinical psychologist, therefore balancing the presence of an expert facilitator across all workshop experiences. However, our study therefore does not have direct evidence to compare between expert and non-expert facilitators, and it is possible that these findings may be attenuated in workshops facilitated by those without clinical training, thus reducing our work's external validity by limiting our ability to generalize these results to interventions led by facilitators without this background. In developing clinical interventions, it is typical for interventions to first be developed in the context of a live session with an expert facilitator, and once initial evidence has been gathered on the efficacy of the intervention content in this context, for research to seek to scale these approaches to more accessible and lightweight solutions such as self-guided asynchronous interventions (Lee et al. 2022). One route for future research may be to develop and assess the impact of more scalable solutions that allow software facilitators to benefit from clinical expertise, such as a "train the trainer" or asynchronous delivery program. Future research can add to the understanding of the external validity of these findings by testing whether similar results are obtained in this different format. Future study designs could also incorporate a design which contrasts expert and non-expert facilitators, to directly test for characteristics of facilitators that may enhance the delivery of code review anxiety interventions. For instance, it is possible that the impact of an expert facilitator differs from non-expert facilitators in the credibility that participants give to the facilitator as an instructor. Future research could incorporate a measure of credibility and trust in the facilitator in order to learn more about the factors which are important to developers in code review anxiety interventions.

6 Conclusion

Code review anxiety is a widely documented experience for software developers that is characterized by a fear of judgment, criticism, and negative evaluation. Due to code review anxiety's similarities to social anxiety, we examined code review anxiety from a clinical perspective to (1) develop a model of code review anxiety and (2) develop, then assess the effectiveness of a novel evidence-based code review anxiety workshop intervention using a randomized waitlist control trial.

Our research shows that when developers experience code review anxiety, they are likely to engage in a range of code review avoidance behaviors – from completely ignoring

reviews, to "rubber stamping" or skimming them, to procrastinating in opening and reviewing pull requests. To reduce developers' avoidance in a human-centered way, our research highlights the importance of mitigating code review anxiety by targeting developers' cost bias or anxiety self-efficacy. Notably, our research indicates that code review anxiety does not depend on developers' cultural identities or their coding experience, making code review anxiety relevant to *any* developer, and thus *any* software team or organization.

Finally, our research provides evidence that a single-session cognitive-behavioral workshop intervention can effectively reduce code review anxiety by significantly increasing anxiety self-efficacy and self-compassion. As this is a notably cost-effective protocol relative to the value and impact of code review activities throughout a developer's career, this finding is an optimistic and important signal for the compounding benefit of empirically-justified interventions to create a more human-centered and healthier developer experience within technology companies. While further research and development, such as a double-blind randomized control trial that measures developers' outcomes during real world code reviews, is necessary to maximize its outcomes and generalizability, the findings provide an uplifting initial model and intervention for reducing code review anxiety.

Acknowledgements The authors would like to thank our participants for their willingness to engage with and support one another during the workshops, as well as Kristen Foster-Marks and Marina Wilding for their early feedback on workshop materials.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Carol Lee. Data visualization was performed by Catherine Hicks. The first draft of the manuscript was written by Carol Lee and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding No funding was received for conducting this study.

Data availability The data that support the findings of this study are not openly available due to reasons of sensitivity and are available from the corresponding author upon reasonable request.

Declarations

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

References

Ackerman AF, Fowler PJ, Ebenau RG (1984) Software inspections and the industrial production of software. In: Proceedings of a symposium on software validation: inspection-testing-verification-alternatives, p 13–40. https://doi.org/10.5555/3541.3543

Ackerman AF, Buchwald LS, Lewski FH (1989) Software inspections: an effective verification process. IEEE Softw 6(3):31–36. https://doi.org/10.1109/52.28121

Alden LE, Bieling PJ (1998) Interpersonal consequences of the pursuit of safety. Behav Res Ther 36(1):53–65. https://doi.org/10.1016/s0005-7967(97)00072-7

Ambuehl B, Inauen J (2022) Contextualized measurement scale adaptation: a 4-Step tutorial for health psychology research. Int J Environ Res Public Health 19(19):12775. https://doi.org/10.3390/ijerph191912775

Bacchelli A, Bird C (2013) Expectations, outcomes, and challenges of modern code review. In: 2013 35th international conference on software engineering (ICSE), p. 712–721. https://doi.org/10.1109/ICSE.2013.6606617

Baer RA (2010) Self-compassion as a mechanism of change in mindfulness- and acceptance-based treatments. In Baer RA (ed) Assessing mindfulness and acceptance processes in clients: illuminating the theory and practice of change. Context Press/New Harbinger Publications, pp 135–153

Baltes S, Ralph P (2022) Sampling in software engineering research: a critical review and guidelines. Empr Softw Eng 27(4):10072–10078. https://doi.org/10.1007/s10664-021-10072-8

Bandura A (1977) Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 84(2):191. https://doi.org/10.1037/0033-295X.84.2.191

Barlow DH (2002) Anxiety and its disorders. Guilford Press, New York

Baysal O, Kononenko O, Holmes R, Godfrey MW (2016) The influence of non-technical factors on code review. Empr Softw Eng 21(2):932–959. https://doi.org/10.1007/s10664-015-9366-8

Bertuzzi V, Fratini G, Tarquino C, Cannistrà F, Granese V, Giusti EM, Castelnuovo G, Pietrabissa G (2021) Single-session therapy by appointment for the treatment of anxiety disorders in youth and adults: a systematic review of the literature. Front Psychol 12:721382. https://doi.org/10.3389/fpsyg.2021.721382

Bird C, Ford D, Zimmermann T, Forsgren N, Kalliamvakou E, Lowdermilk T, Gazit I (2022) Taking flight with copilot: early insights and opportunities of AI-powered pair-programming tools. Queue 20(6):35–57. https://queue.acm.org/detail.cfm?id=3582083

Bosu A, Carver JC (2013) Impact of peer code review on peer impression formation: a survey. In: 2013 ACM / IEEE international symposium on empirical software engineering and measurement, p 133–142. https://doi.org/10.1109/ESEM.2013.23

Brockner J, Sherman DK (2019) Wise interventions in organizations. Res Org Behav 39:100125

Brown LA, Forman EM, Herbert JD, Hoffman KL, Yuen EK, Goetter EM (2011) A randomized controlled trial of acceptance-based behavior therapy and cognitive therapy for test anxiety: a pilot study. Behav Mod 35:31–53. https://doi.org/10.1177/0145445510390930

Calafell JM (2019) How I learned to stop worrying and love the code review. https://medium.com/@joshuamichaelcalafell/how-i-learned-to-stop-worrying-and-love-the-code-review-fa87926d6f24

Call CC, Eckstrand KL, Kasaparek SW, Boness CL, Blatt L, Jamal-Orozco N, Novacek DM, Foti D, Scholars for Elevating Equity and Diversity (SEED) (2023) An ethics and social justice approach to collecting and using demographic data for psychological researchers. Perspect Psychol Sci 18(5):979–995. https://doi.org/10.1177/17456916221137350

Clark DM, Wells AA (1995) A cognitive model of social phobia. In: Heimberg RG, Liebowitz MR, Hope DA, Schneier FR (eds) Social phobia: diagnosis, assessment, and treatment. Guilford Press, New York, pp 69–93

Cohen J, Teleki S, Brown E (2013) Best kept secrets of peer code review. SmartBear Software, Beverly

R Core Team (2023) R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Cougle JR, Wilver NL, Day TN, Summers BJ, Okey SA, Carlton CN (2020) Interpretation bias modification versus progressive muscle relaxation for social anxiety disorder: a web-based controlled trial. Behav Ther 51(1):99–112. https://doi.org/10.1016/j.beth.2019.05.009

Craske M (2017) Cognitive-behavioral therapy. American Psychological Association, Washington DC

Craske M, Wittchen U, Bogels S, Stein M, Andrews G, Lebeu R (2013) Severity measure for specific phobia-adult. https://www.psychiatry.org/File%20Library/Psychiatrists/Practice/DSM/APA_DSM5_Severity-Measure-For-Specific-Phobia-Adult.pdf

Cukrowicz KC, Joiner TE (2007) Computer-based intervention for anxious and depressive symptoms in a non-clinical population. Cognit Ther Res 31:677–693. https://doi.org/10.1007/s10608-006-9094-x

Cunha A, Conte T, Gadelha B (2021) Code Review is just reviewing code? A qualitative study with practitioners in industry. In: Proceedings of the XXXV Brazilian Symposium on Software Engineering (SBES '21), p 269–274. https://doi.org/10.1145/3474624.3477063

Danitz SB, Orsillo SM (2014) The mindful way through the semester: an investigation of the effectiveness of an acceptance-based behavioral therapy program on psychological wellness in first-year students. Behav Mod 38:549–566. https://doi.org/10.1177/0145445513520218

Darmuki A, Andayani A, Nurkamto J, Saddhono K (2017) Evaluating information processing-based learning cooperative model on speaking skill course. JLTR 8:44–51. https://doi.org/10.17507/jltr.0801.06
Dasgupta N (2011) Ingroup experts and peers as social vaccines who inoculate the self-concept: the stereotype inoculation model. Psychol Inq 22(4):231–246
David D, Cristea I, Hofmann SG (2018) Why cognitive behavioral therapy is the current gold standard of psychotherapy. Front Psychiatry 9:4. https://doi.org/10.3389/fpsyt.2018.00004
Eustis EH, Williston SK, Morgan LP, Graham JR, Hayes-Skelton SA, Roemer L (2017) Development, acceptability, and effectiveness of an acceptance-based behavioral stress/anxiety management workshop for university students. Cogn Behav Pract 24:174–186. https://doi.org/10.1016/j.cbpra.2016.03.011
Fletcher TD (2022) QuantPsyc: Quantitative psychology tools. https://cran.r-project.org/package=QuantPsyc
Foa EB, Franklin ME, Perry KJ, Herbert JD (1996) Cognitive biases in generalized social phobia. J Abnorm Psychol 105:433–439. https://doi.org/10.1037/0021-843X.105.3.433
Friendly M, Fox J, Monette G, Chalmers P, Murdoch D (2024) heplots: visualizing hypothesis tests in multivariate linear models. https://cran.r-project.org/package=heplots
Gaudiano BA, Herbert JD (2003) Preliminary psychometric evaluation of a new self-efficacy scale and its relationship to treatment outcome in social anxiety disorder. Cognit Ther Res 27(5):537–555. https://doi.org/10.1023/A:1026355004548
Goldin PR, Ziv M, Jazaieri H, Werner K, Kraemer H, Heimberg RG, Gross JJ (2012) Cognitive reappraisal self-efficacy mediates the effects of individual cognitive-behavioral therapy for social anxiety disorder. J Consult Clin Psychol 80:1034–1040. https://doi.org/10.1037/a0028555
Hayes-Skelton SA, Lee CS (2020) Decentering in mindfulness and cognitive restructuring: an experimental study of a potential common mechanism. Behav Mod 44(6):817–840. https://doi.org/10.1177/2F0145445519850744
Hazzan O, Har-Shai G (2014) Teaching and learning computer science soft skills using soft skills: the students' perspective. Proc 45th ACM Tech Symp Comput Sci Educ 567–572. https://doi.org/10.1145/2538862.2538885
Heimberg RG, Becker RE (2002) Cognitive-behavioral group therapy for social phobia: Basic mechanisms and clinical strategies. The Guilford Press, New York
Hindo CS, González-Prendes AA (2011) One-session exposure treatment for social anxiety with specific fear of public speaking. RSWP 21(5):528–538. https://doi.org/10.1177/1049731510393984
Hofmann SG (2005) Perception of control over anxiety mediates the relation between catastrophic thinking and social anxiety in social phobia. Behav Res Ther 43:885–895. https://doi.org/10.1016/j.brat.2004.07.002
Hofmann SG (2007) Cognitive factors that maintain social anxiety disorder: a comprehensive model and its treatment implications. Cognit Behav Ther 36(4):193–209. https://doi.org/10.1080/16506070701421313
Iwamasa GY, Hays PA (2019) Culturally responsive cognitive behavior therapy: Practice and supervision (2nd edition). American Psychological Association, Washington D.C
Jacobson NS, Truax P (1991) Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 59(1):12–19. https://doi.org/10.1037/0022-006X.59.1.124
Kim D, Bae H, Park YC (2008) Validity of the subjective units of disturbance scale in EMDR. J EMDR Pract Res 2(1):57–62. https://doi.org/10.1891/1933-3196.2.1.57
Komarraju M, Nadler, D (2013) Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter? Learn Individ Differ 25:67–72. https://doi.org/10.1016/j.lindif.2013.01.005
Kudrjavets G, Rastogi A (2024) Does code review speed matter for practitioners? Empir Softw Eng 29(7). https://doi.org/10.1007/s10664-023-10401-z
Leary MR, Kowalski RM (1995) The self-presentational model of social phobia. In: Heimberg RG, Liebowitz MR, Hope DA, Schneier FR (eds) Social phobia: diagnosis, assessment, and treatment. Guilford Press, New York, pp 94–112
Lee CS, Hayes-Skelton SA (2020) Finding personal meaning as a predictor of behavioral action over and above social anxiety. J Theor Soc Psychol 5:26–34. https://doi.org/10.1002/jts5.81
Lee CS, Bowman M, Wu JL (2022) Preliminary outcomes of an online asynchronous stress and anxiety management workshop for college students. Trends Psychiatry Psychother 45:e20210448. https://doi.org/10.47626/2237-6089-2021-0448
Lent RW, Brown SD, Larkin KC (1984) Relation of self-efficacy expectations to academic achievement and persistence. J Counsel Psychol 31(3):356. https://doi.org/10.1037/0022-0167.31.3.356
Li Z, Lu S, Guo D, Duan N, Jannu S, Jenks G, Majumder D, Green J, Svyatkovskiy A, Fu S, Sundaresan N (2022) Automating code review activities by large-scale pre-training. Proc 30th ACM Joint Eur Softw Eng Conf Symp Found Softw Eng (ESEC/FSE 2022) 1035–1047. https://doi.org/10.1145/3540250.3549081
Linehan MM (2014) DBT training manual. The Guilford Press, New York

Martinez JH, Eustis EH, Arbid N, Graham-LoPresti JR, Roemer L (2020) Experiential avoidance moderates the relation between racial discrimination and negative mental health outcomes. J Am Coll Health 70(2):461–468. https://doi.org/10.1080/07448481.2020.1754221
Mayasari S, Mujiyati, Adiputra S (2022) Cognitive restructuring techniques in developing student self-compassion. JPSP 6(3):2564–2575
McLean CP, Asnaani A, Litz BT, Hofmann SG (2011) Gender differences in anxiety disorders: prevalence, course of illness, comorbidity and burden of illness. J Psychiatr Res 45(8):1027–1035. https://doi.org/10.1016/j.jpsychires.2011.03.006
Moss R (2018) Handling code review feedback. https://www.raquelmoss.com/handing-code-review-feedback/
Muhomba M, Chugani CD, Uliaszek AA, Kannan D (2017) Distress tolerance skills for college students: a pilot investigation of a brief DBT group skills training program. J Coll Stud 31:247–256. https://doi.org/10.1080/87568225.2017.1294469
Murphy-Hill E, Jaspan C, Egelman C, Cheng L (2022) The pushback effects of race, ethnicity, gender, and age in code review. CACM 65(3):52–57. https://doi.org/10.1145/3474097
Nadri R, Rodríguez-Pérez G, Nagappan M (2021) On the relationship between the developer's perceptible race and ethnicity and the evaluation of contributions in oss. IEEE Trans Softw Eng 48(8):2955–2968
Nazligul MD, Yilmaz M, Gulec U, Gozcu MA, O'Connor RV, Clarke P (2017) Overcoming public speaking anxiety of software engineers using virtual reality exposure therapy. European Conf Software Process Improvement. https://doi.org/10.1007/978-3-319-64218-5_15
Neff KD, Germer CK (2013) A pilot study and randomized controlled trial of the Mindful Self-Compassion program. J Clin Psychol 69(1):28–44. https://doi.org/10.1002/jclp.21923
Neff KD, Tóth-Király I, Knox MC, Kuchar A, Davidson O (2021) The development and validation of the state self-compassion scale (long- and short form). Mindfulness 12(1):121–140. https://doi.org/10.1007/s12671-020-01505-4
Norton PJ, Hope DA (2001) Kernels of truth or distorted perceptions: self and observer ratings of social anxiety and performance. Behav Ther 32(4):765–786. https://doi.org/10.1016/S0005-7894(01)80020-4
Panepinto AR, Uschold CC, Olandese M, Linn BK (2015) Beyond borderline personality disorder: dialectical behavior therapy in a college counseling center. J Coll Stud 29:211–226. https://doi.org/10.1080/87568225.2015.1045782
Peters L, Moreno AM (2015) Educating software engineering managers-revisited what software project managers need to know today. ICSE '15: Proc 37th Int Conf Softw Eng 2:353–359
Piper M (2023) Overcoming anxiety in code reviews. https://www.viget.com/articles/overcoming-anxiety-in-code-reviews/
Pusswald G, Wiesbauer P, Pirker W, Novak K, Foki T, Lehrner J (2019) Depression, quality of life, activities of daily living, and subjective memory after deep brain stimulation in Parkinson disease—A reliable change index analysis. Int J Geriatr Psychiatry 34(11):1698–1705. https://doi.org/10.1002/gps.5184
Ranganathan P, Pramesh CS, Buyse M (2015) Common pitfalls in statistical analysis: clinical versus statistical significance. Perspect Clin Res 6(3):169–170. https://doi.org/10.4103/2229-3485.159943
Rapee RM, Heimberg RG (1997) A cognitive-behavioral model of anxiety in social phobia. Behav Res Ther 35:741–756. https://doi.org/10.1016/s0005-7967(97)00022-3
Remes O, Brayne C, van der Linde R, Lafortune L (2016) A systematic review of reviews on the prevalence of anxiety disorders in adult populations. Brain Behav 6(7):e00497. https://doi.org/10.1002/brb3.497
Revelle W (2024) psych: Procedures for psychological, psychometric, and personality research. https://cran.r-project.org/package=psych
Rizopoulos D (2022) ltm: Latent trait models under IRT. https://cran.r-project.org/package=ltm
Rizvi SL, Steffel LM (2014) A pilot study of 2 brief forms of dialectical behavior therapy skills training for emotion dysregulation in college students. J Am Coll Health 62:434–439. https://doi.org/10.1080/07448481.2014.907298
Roberts SO, Bareket-Shavit C, Dollins FA, Goldie PD, Mortenson E (2020) Racial inequality in psychological research: Trends of the past and recommendations for the future. Perspect Psychol Sci 15(6):1295–1309. https://doi.org/10.1177/1745691620927709
Schwarzer R, Luszczynska A (2023) Self efficacy. In: Ruch W, Bakker AB, Tay L, Gander F (eds) Handbook of positive psychology assessment. Hogrefe Publishing, Newburyport, MA, pp 207–217
Shull F, Seaman C (2008) Inspecting the history of inspections: an example of evidence-based technology diffusion. IEEE Softw 25(1):88–90. https://doi.org/10.1109/MS.2008.7
Söderberg E, Church L, Börstler J, NiehorsteR DC, Rydenfält C (2022) Understanding the experience of code review: misalignments, attention, and units of analysis. In: EASE '22: Proceedings, pp. 170–179. https://doi.org/10.1145/3530019.3530037
Soper D (2024) Sample size calculators. https://www.danielsoper.com/statcalc/category.aspx?id=19
Tabachnick BG, Fidell LS (2021) Using multivariate statistics. Pearson, Boston, MA

Tahmassian K, Jalali Moghadam N (2011) Relationship between self-efficacy and symptoms of anxiety, depression, worry and social avoidance in a normal sample of students. IJPBS 5(2):91–98. https://pubmed.ncbi.nlm.nih.gov/24644452/

Trower P, Gilbert P (1989) New theoretical conceptions of social anxiety and social phobia. Clin Psychol Rev 9(1):19–35. https://doi.org/10.1016/0272-7358(89)90044-5

Uliaszek AA, Rashid T, Williams GE, Gulamani T (2016) Group therapy for university students: a randomized control trial of dialectical behavior therapy and positive psychotherapy. Behav Res Ther 77:78–85. https://doi.org/10.1016/j.brat.2015.12.003

Vazire S, Schiavone SR, Bottesini JG (2022) Credibility beyond replicability: improving the four validities in psychological science. Curr Dir Psychol Sci 31(2):162–168. https://doi.org/10.1177/09637214211067779

Vitasari P, Wahab MNA, Herawan T, Othman A, Sinnadurai SK (2011) A pilot study of pre-post anxiety treatment to improve academic performance for engineering students. Procedia-Social Behav Sci 15(1):3826–3830. https://doi.org/10.1016/j.sbspro.2011.04.380

Ward-Ciesielski EF, Jones CB, Wielgus MD, Wilks CR, Linehan MM (2016) Single-session dialectical behavior therapy skills training versus relaxation training for non-treatment engaged suicidal adults: a randomized controlled trial. BMC PsychoL 4:13. https://doi.org/10.1186/s40359-016-0117-4

Wickham et al (2019) Welcome to the tidyverse. JOSS 4(43):1686. https://doi.org/10.21105/joss.01686

Wolpe J (1990) The practice of behavior therapy (4th edition). Pergamon Press, New York

Wurzel Gonçalves P, Calikli G, Serebrenik A, Bacchelli A (2023) Competencies for code review. Proc ACM Hum Comput Interact 7(38):1–33. https://doi.org/10.1145/3579471

Zessin U, Dickhäuser O, Garbade S (2015) The relationship between self-compassion and well-being: a meta-analysis. Appl Psychol Health Well Being 7(3):340–364. https://doi.org/10.1111/aphw.12051