Innovative Higher Education (2021) 47:155–174
https://doi.org/10.1007/s10755-021-09574-6

Doctoral writing workshops: A pre-registered, randomized controlled trial

Barbara W. Sarnecka[1] · Paulina N. Silva[1] · Jeff Coon[1] · Darby C. Vickers[2] · Rena B. Goldstein[3] · Jeffrey N. Rouder[1]

Accepted: 8 July 2021 / Published online: 19 January 2022

Abstract

Doctoral students were randomly assigned to a five-week (30-h) faculty-led writing workshop intervention, either preceded by a five-week (waiting list) control phase or followed by a five-week maintenance phase. In the workshop, students wrote together, received instruction in genres of academic writing (literature reviews, scientific articles, funding proposals, and presentations), and exchanged feedback on drafts. As a result of the workshop students enjoyed writing more, found writing easier, and gained confidence in themselves as academic writers. They felt able to write productively in shorter blocks of time, and they engaged in more short-term, medium-term, and long-term planning of their research. The intervention also caused participants to pause more frequently for reflection or positive thinking and to generate more new writing. Effects were maintained in a peer-led writing maintenance group for at least five weeks after the intervention ended. This is the first randomized controlled trial of a doctoral-level writing intervention to date and has the potential to support doctoral training in academic and scientific writing across the Social Sciences, Education, and the Humanities.

Keywords Mentorship · Dissertation · Doctoral education · Graduate education · Intervention · Writing

> *I simply wanted to extend a heartfelt THANK YOU for the workshop experience. I came in feeling really burnt out, overwhelmed, and kind of gritting my teeth when it came to the dissertation. I hadn't read anything and was feeling very resistant. In short, writing the dissertation sounded like a miserable and impossible process. Five weeks later, and I'm in the best place*

✉ Barbara W. Sarnecka
sarnecka@uci.edu

[1] Department of Cognitive Sciences, University of California, Irvine, CA 92697, USA

[2] Department of Philosophy, University of San Diego, San Diego, CA 92110, USA

[3] Department of Philosophy, University of California, Irvine, CA 92697, USA

I've been since entering graduate school. My writing habits still need work, but I'm moving in the right direction (wrote two pages yesterday!). I'm reading, researching, journaling, and genuinely excited about my project. The workshop provided me with the support and accountability I needed. Honestly, I'm scared to imagine where I'd be this summer without the workshop. At the risk of sounding dramatic, I think the workshop may have saved my dissertation. It has been an incredible gift. Thank you for advocating on behalf of our graduate students and creating this space.

-Jacob B., fourth-year doctoral student and writing workshop participant (Summer 2019)

Becoming an academic writer is among the most challenging aspects of doctoral training, but doctoral programs typically provide little instruction in scholarly writing (Aitchison et al., 2012). Most PhD students were successful undergraduate writers, but doctoral programs require them to write in new genres and follow disciplinary conventions that are often implicit (Maher et al., 2014). While advisors lament students' poor writing skills, lack of familiarity with academic conventions, and unwillingness to revise, the single greatest challenge to successfully completing a dissertation according to faculty is students' inability to plan, implement, and write up large-scale, independent projects (D'Andrea, 2002). Many students hold harmful misconceptions about writing, such as the belief that they can only write in large blocks of uninterrupted time (Silvia, 2007). However, as faculty know, such stretches of time are rare in an academic schedule. Without support, students may or may not discover on their own how to accomplish their research in smaller, more realistically manageable chunks of time.

The lack of instruction in writing and research planning also adds unnecessarily to students' stress. Writing dysfunction is widespread in academia (Stolzenberg et al., 2019), with feelings of frustration and paralysis often exacerbated by external factors such as the increasingly competitive academic job market and reduced time to degree (Aitchison et al., 2012). We take a very broad view of what constitutes "writing", because we take the position that writing is thinking, and that thinking aided by any form of external symbolic representation is part of writing. Thus, "writing" in this study includes designing experiments, developing computer code, creating figures and preparing presentations, in addition to outlining, drafting, and revising research reports and proposals.

Traditional models conceptualize doctoral training as an apprenticeship, with the primary agent of socialization into a discipline being the student's faculty advisor (Austin, 2002; Gardner & Mendoza, 2012; Weidman & Stein, 2003). But unlike the professions where apprenticeship evolved (e.g., tailoring, cabinetry, blacksmithing), the processes involved in writing are not readily observable. An apprenticeship in writing requires advisors and students to write together in a way that allows the student to observe the advisor's writing process from beginning to end– something that happens all too rarely (Collins et al., 1991). As one longitudinal study noted, "Particularly noteworthy and a cause for concern is the...

minimal feedback and mentoring from faculty... Although some students had faculty mentors who guided them carefully through the process, most did not" (Austin, 2002, p. 104).

Although the traditional apprenticeship model has clear shortcomings, there is no widely accepted alternative to take its place. The present study tested the effectiveness of an intervention that combines several forms of writing support, and that can be integrated into existing doctoral programs. Some of these forms of support, such as providing dedicated graduate student writing rooms, summer dissertation-writing fellowships, and dissertation-writing retreats or boot camps are already considered best practices by the Council of Graduate Schools (Sowell et al., 2010). To date however, most of these practices have not been rigorously tested. In one case, a study of summer bridge programs for beginning graduate students found no evidence that such programs benefited students in the long term (Feldon et al., 2017).

Doctoral education needs evidence-based practices. The present study tested a writing workshop intervention that combined several practices for which at least some prior evidence existed. Although effects in this study were measured after only five weeks, the workshop is ideally run as an ongoing support program for doctoral students, supplementing the mentorship of their primary faculty advisors. Key elements of the model are briefly reviewed below; complete workshop materials are available on the Open Science Framework at https://osf.io/ftuhp/.

Quiet Writing Time Each workshop meeting started with a period of 30 min during which students simply wrote quietly in each other's presence (Rockquemore, 2010). Evidence for the effectiveness of this practice has been shown by programs such as the Scholar's Retreat at the University of Colorado at Denver (Smallwood, 2004), and the Dissertation House at the University of Maryland Baltimore County (Carter-Veale et al., 2016). Although the Scholar's Retreat and Dissertation House programs were aimed at students who had already reached the dissertation stage, the present intervention was open to students in any year of a doctoral program.

Planning and Accountability The intervention also taught students how to plan their research and writing. Research from outside academia shows that people who make deliberate career plans with specific, step-by-step goals go on to earn higher salaries, more promotions, and more responsibility in their jobs (Ng et al., 2005), and report feeling more satisfied and more successful than people who do not make such plans (Abele & Wiese, 2008). Within academia, long- and short-term goal-setting are central tenets of the "solution-focused counseling" approach to dissertation support (Johnson & Conyers, 2001). And for good reason: A study of 7,600 postdoctoral fellows found that those who worked with their advisors to develop a plan for their postdoctoral training were more productive, more satisfied with their jobs, and less likely to experience conflict with the advisors (Davis, 2005). The National Institutes of Health (NIH) now strongly recommends that all graduate students and postdocs make multiyear "individual development plans" (IDPs), and that advisors refer to these plans in their annual progress reports (Rockey, 2013). In the present intervention, students were guided to make research and writing plans on three time

scales: Long term (1–5 years), medium term (10–13 weeks), and short term (weekly and daily). Progress toward weekly and daily goals was recorded on a shared daily (online) writing log, for which social accountability was provided during workshop meetings.

Support for Daily Writing The intervention helped students to establish and maintain a regular, daily writing practice. As experiments by the psychologist Robert Boice (1990) demonstrated, writing little and often is a powerful technique for managing writing-related anxiety and for dispelling common misconceptions about writing (e.g., "I can't do any useful writing in less than two hours"; "I can't start writing until I know what I am going to say"; "I need to be inspired in order to write"; etc.). The intervention emphasized writing little and often as a way of overcoming writing dysfunction, and provided social support for daily writing through the shared daily writing log.

Feedback on Writing in Progress Writing skill develops through practice and feedback. According to Sadler's (1989) theory of formative assessment and instructional design, effective feedback requires that three conditions be met. First, the student must know what good writing in the target genre looks like. Second, the student must be able to see how their current writing falls short of that goal. Third, the student must have a repertoire of steps or strategies to follow in order to get from their current performance to the goal. In practice, these conditions are met when students have frequent opportunities to exchange feedback with other scholars in their discipline. As researchers Caffarella and Barnett (2000, p. 39) found in their study of 45 doctoral student writers, "Preparing and receiving critiques from professors and peers was perceived to be the most influential element in helping [students] to understand the process of scholarly writing and in producing a better written product." In traditional apprenticeship models of doctoral mentoring, feedback comes primarily from the faculty advisor and occasionally from other faculty members. Yet faculty are rarely able to provide feedback as promptly, as frequently, or in as much detail as students need. Students also benefit from feedback early in the writing process, at the outline or rough draft stage, but may feel self-conscious showing unpolished drafts to faculty. Additionally, because advisors often do not have time to give multiple rounds of feedback on one document, students may wait until they have a polished draft to share. One solution is for students to get regular, low-stakes feedback from peers (Burnett, 1999; Delyser, 2003). In the present intervention, students regularly exchanged feedback on short (1–2 pages) written drafts.

Growth Mindset The writing workshop aims to instill what developmental psychologists call a "growth mindset" about research and writing. The term was coined by researcher Carol Dweck who explains, "Individuals who believe their talents can be developed (through hard work, good strategies, and input from others) have a growth mindset. They tend to achieve more than those with a more fixed mindset (those who believe their talents are innate gifts). This is because they worry less about looking smart and they put more energy into learning" (Dweck, 2016). In

academia, fixed mindsets are associated with the underrepresentation of women and Black scholars in a field. That is, the more people believe that success in a given discipline depends on raw, innate talent, the fewer women and Black scholars that field has (Leslie et al., 2015; Meyer et al., 2015). If research skill were physical fitness, a person with a growth mindset would say that physically fit people got that way by exercising; a person with a fixed mindset would say that physically fit people are just naturally blessed with good health. The writing workshop reinforces a growth mindset by teaching participants to focus more on the process of research and writing, and less on the products.

The present paper describes a small-scale, randomized controlled trial of the effects of our writing workshop intervention. At the time of pre-registration,[1] we hypothesized that the intervention would (1) make participants write more; (2) improve the quality of participants' writing; (3) improve participants' attitudes and work habits related to writing, and (4) improve participants' subjective well-being (e.g., by alleviating anxiety related to writing dysfunction, and/or increasing social support).

Method

Participants

Over 1,000 doctoral students in the schools of Social Sciences, Social Ecology, Education, and Humanities at a large, public research university in the U.S. were invited by email to participate in the study. The email contained a link to an initial interest survey, which was returned by 87 students (59 in the summer session, 28 in the fall), of whom 35 eventually enrolled in the study. Of these, 3 dropped out after attending at least one meeting (all cited last-minute schedule changes), leaving a final sample of 32 participants. Of these, 14 described their research methods as quantitative, 12 worked in philosophy or the humanities, and 6 did qualitative research.

Women, first-generation college students, and students from underrepresented racial and ethnic groups returned the surveys at higher rates than others. For example, women made up 52.15% of doctoral students in the invited programs, but they made up 67.82% of respondents to the initial interest survey and 81.25% of participants in the study. First-generation students made up an estimated one-third of students in the programs (Gardner, 2013; Hoffer et al., 2001; Roksa et al., 2018) but made up 45.98% of those returning the initial interest survey and 53.13% of study participants. Underrepresented ethnic and racial minority students were 18.31% of students in the programs, but made up 32.18% of those returning the initial interest survey and 25.00% of those in the study.

[1] Pre-registration for the summer data collection period is posted at https://osf.io/t7r4f. Pre-registration for the fall data collection period is posted at https://osf.io/39j2q.

Procedure

The full experiment was run twice: Once in the 10-week summer academic quarter of 2019 and again in the 10-week fall quarter. The experimental design for each 10-week period is illustrated in Fig. 1. Each time, participants were divided into two experimental groups and the quarter was divided into two halves (Weeks 1–5 and Weeks 6–10). Half of the participants were randomly assigned to the Control + Intervention (*Con + Int*) group. This group stayed on a waiting list during Weeks 1–5 (Control phase) and received the intervention during Weeks 6–10 (Intervention phase). The other half of the participants were assigned to the Intervention + Maintenance (*Int + Main*) group. This group received the intervention during Weeks 1–5 (Intervention phase) and during Weeks 6–10 (Maintenance phase), were invited to attend maintenance meetings led by graduate students (the second through fifth authors). Participants completed three surveys over the course of the study: one at the beginning (before Week 1), another halfway through (after Week 5) and one at the end (after Week 10).

Participants met twice per week for 2 h, 50 min per meeting. Meetings were led by a faculty member (the first author). Most meetings began with 30 min of quiet writing time, followed by 30 min of check-in, when participants were invited to share briefly about their writing since the last meeting and to discuss with the group any challenges or obstacles they were facing. After the check-in, the leader facilitated a 45–50-min group discussion of an assigned reading from the book *The Writing Workshop* (Sarnecka, 2019). Topics included research planning, time management, writing strategies (e.g., freewriting, mind-mapping, reverse outlining, etc.), writing genres (literature reviews, journal articles, conference presentations, or funding proposals), writing style (e.g., paragraph structure, sentence structure, word choice), and strategies for maintaining physical and psychological well-being in academia (e.g., sleep, exercise, social interaction, etc.) Most meetings ended with one or two 30-min feedback forum sessions. In each session, one participant shared a brief (1–2 page) sample of their work. The class spent 20 min reading quietly and

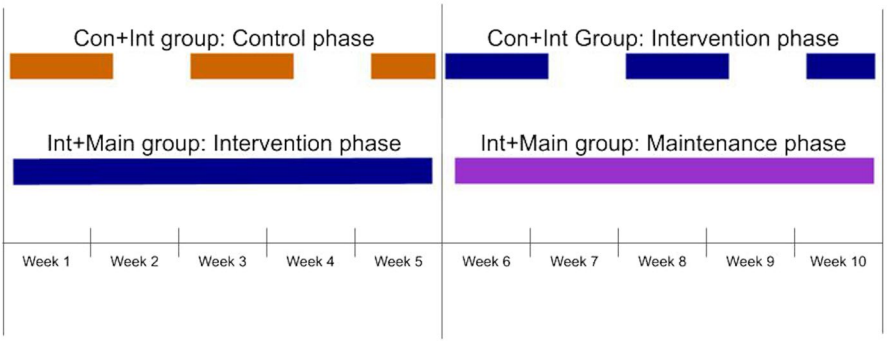

Fig. 1 Experimental design. The Con + Int group (dashed lines) served as controls in Weeks 1–5 and completed the intervention in Weeks 6–10. The Int + Main group (solid lines) had the intervention in Weeks 1–5 and attended a maintenance group in Weeks 6–10

making written suggestions and comments, followed by 10 min giving verbal feedback to the author.

In addition to attending the workshop meetings, participants were encouraged to use an online spreadsheet called the "shared daily writing log" to record writing goals, log writing, and engage in a daily positivity practice (e.g., list something you are grateful for today). Additionally, participants were invited to add any rejections they received to an online "rejection collection," which would result in a party whenever a group total of 100 rejections was reached. All surveys, data, analysis code and workshop materials (syllabus, writing log, rejection collection, and course textbook) are posted on the Open Science Framework.[2]

Results

This results section presents statistical analyses of survey data, along with narrative explanations of the findings. Throughout this section, we refer to publicly registered hypotheses from the summer (https://osf.io/t7r4f) and fall (https://osf.io/39j2q) data collection periods. Although our pre-registered analyses used Bayesian statistical methods, we present traditional frequentist analyses here on the assumption that they are more familiar to many readers. Bayesian versions of all ANOVAs are posted online for interested readers[2]. Here in the main text, we have included Bayes factors with the ANOVAs only for null or marginal results, as frequentist tests cannot quantify evidence in favor of the null hypothesis. All Bayes factors presented here are in relation to the null model. We present the results in order of the strength of evidence observed for each hypothesis.

Strong Evidence for Hypothesis 3: Improvements in Writing-Related Beliefs and Practices

Hypothesis 3 was that the intervention would "make [participants'] writing-related beliefs and practices more positive and sustainable, as measured by our novel, 19-item 'Writing Beliefs and Practices' scale." The scale was divided into three sections. Items 1–10 asked about attitudes toward writing; Item 11 asked participants to estimate the shortest block of time required for them to write productively; and Items 12–19 asked about participants' use of writing and research plans.

Items 1–10: Improved attitudes toward writing The workshop caused large improvements in participants' responses to each of ten survey items asking about attitudes toward writing. Overall, participants held more positive beliefs about writing after participating in the workshop. Responses were given on a 7-point Likert scale ranging from "strongly disagree" to "strongly agree." The intervention increased participants' agreement with the four positive statements: "I enjoy

[2] Surveys, data, analysis code, and all materials used in the workshop intervention are posted at https://osf.io/ftuhp/.

writing" (labeled *Enjoy* in Fig. 2 and Table 1); "I am good at writing" (labeled *Good*); "I know how to describe my research in a way that non-experts can understand" (labeled *Non-experts*); and "I am good at giving constructive, useful feedback on other people's writing" (labeled *Give feedback*). The intervention also caused participants to disagree more with six negative statements: "I often feel discouraged or disappointed with myself for not writing enough" (labeled *Discouraged* in Fig. 2 and Table 1); "I struggle with writer's block" (labeled *Writer's block*); "I need to be inspired in order to write" (labeled *Need inspiration*); "I can't start writing until I've figured out in my head what I'm going to say" (labeled *Can't start*); "If I know something is likely to be rejected, I don't waste time writing and submitting it" (labeled *Fear rejection*); and "Receiving feedback on my writing is unpleasant" (labeled *Dislike feedback*).

For each participant and each survey, we constructed a "writing attitudes score" from the average of that participant's responses for the ten "writing attitudes" survey items. (Items 6–10, which were negative statements, were reverse-coded for this summary

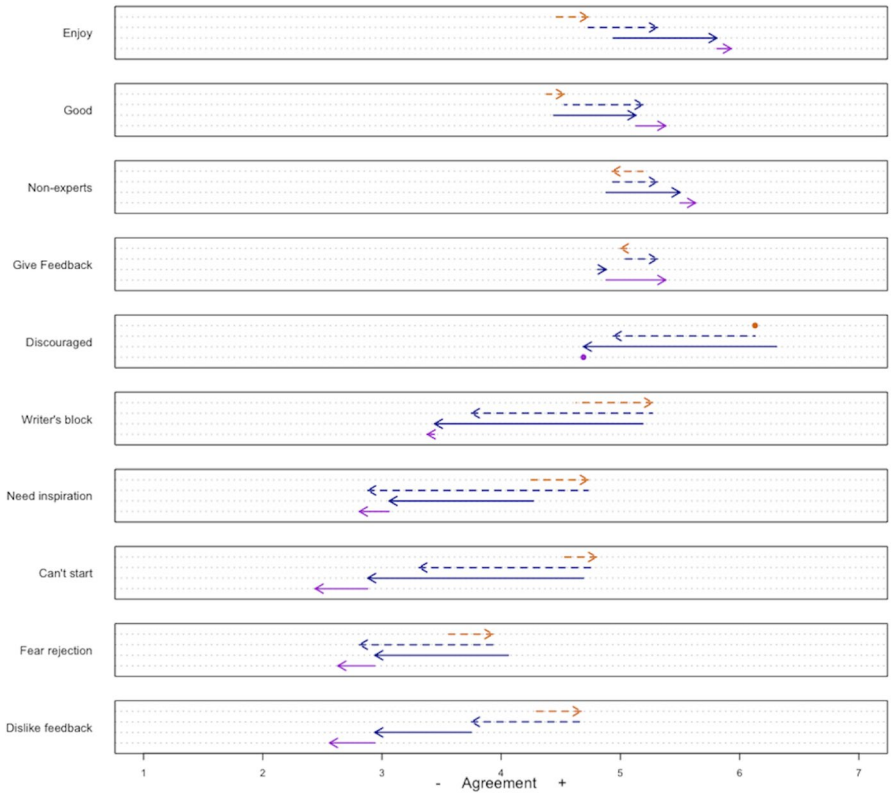

Fig. 2 Changes in attitudes toward writing. Dotted lines show changes in the Con + Int group during the control (orage) and intervention (blue) phases. Solid lines show the Int + Main group during intervention (blue) and maintenance (purple)

Table 1 Agreement/disagreement with statements about writing beliefs and practices. Values are means and (standard deviations)

	ConInt Group			IntMain Group		
	Survey 1	Survey 2	Survey 3	Survey 1	Survey 2	Survey 3
Enjoy	4.44 (1.59)	4.73 (1.28)	5.31 (1.25)	4.94 (1.18)	5.81 (.98)	5.93 (.77)
Good	4.38 (1.46)	4.53 (1.25)	5.19 (.98)	4.44 (1.37)	5.13 (1.15)	5.38 (1.09)
Non-experts	5.19 (1.47)	4.93 (1.39)	5.31 (1.35)	4.88 (1.41)	5.50 (.97)	5.63 (.81)
Give feedback	5.06 (1.24)	5.00 (1.46)	5.31 (.95)	4.81 (1.28)	4.88 (1.31)	5.38 (1.03)
Discouraged[a]	6.31 (.89)	6.13 (.99)	4.94 (1.44)	6.31 (1.01)	4.69 (1.78)	4.69 (1.58)
Writer's block[a]	4.63 (1.82)	5.27 (1.67)	3.75 (1.77)	5.19 (1.38)	3.44 (1.67)	3.38 (1.86)
Need inspiration[a]	4.25 (1.95)	4.73 (1.75)	2.88 (1.78)	4.27 (2.09)	3.06 (1.95)	2.81 (1.72)
Can't start[a]	4.50 (1.86)	4.80 (2.01)	3.31 (1.49)	4.69 (2.09)	2.88 (1.71)	2.44 (1.21)
Fear rejection[a]	3.56 (1.67)	3.93 (1.75)	2.81 (1.11)	4.06 (1.91)	2.94 (1.24)	2.63 (1.41)
Dislike feedback[a]	4.25 (2.18)	4.67 (2.16)	3.75 (1.95)	3.75 (1.65)	2.94 (1.57)	2.56 (1.41)
Scale score (average)[b]	**3.98 (1.85)**	**3.77 (1.85)**	**4.77 (1.57)**	**3.88 (1.82)**	**4.94 (1.56)**	**5.06 (1.49)**

[a]The intervention was predicted to decrease agreement with this item; lower scores indicate better outcomes

[b]To compute this scale score, negative items were reverse coded so that higher scores indicate better outcomes for all items

score; higher scores represented more positive attitudes.) In order to determine whether the intervention changed attitudes, we looked at changes in each participant's writing attitudes score across the three surveys. The changes in scores from Survey 1 to 2 and from Survey 2 to 3 for each participant represent changes in that participant's attitudes during the first and second phases of the experiment, respectively.

Because the experiences of the two groups differed (i.e., both groups received the intervention, but for the Con + Int group it was preceded by a waiting period, and for the Int + Main group it was followed by a maintenance period), we analyzed the data for each group separately. Repeated measures ANOVA of the Con + Int group analyzed the difference between the amount that writing attitudes changed in Phase 1 (the control phase) and the amount that writing attitudes changed in Phase 2 (the intervention phase). A repeated measures ANOVA of the Con + Int group's scale scores shows an effect of phase, as writing attitude scores changed more during the intervention phase than during the control phase. $F(2,28) = 17.46$, $p < 0.001$, $\eta^2 = 0.29$. A repeated measures ANOVA of the Int + Main group's scores also shows an effect of phase, as scores changed more during the intervention phase than during the maintenance phase. $F(2,30) = 29.24$, $p < 0.001$, $\eta^2 = 0.30$.

Item 11: Shorter blocks of time required for writing We also found support for the prediction that the intervention would enable students to write in shorter blocks of time. Each survey included the question, "What's the shortest block of time you can use productively for writing? In other words, what is the minimum amount of time that you need to have free in order to get any useful writing done?" During

the intervention, these estimates dropped by over 58%, from a mean of 41.67 min (SD = 32.61) to a mean of 17.35 min (SD = 21.70) across both intervention groups. Estimates continued to drop during the maintenance phase, from a mean of 22.07 min (SD = 29.32) for the Int + Main group at the end of the intervention to a mean of 13.81 min (SD = 12.36) for the same group at the end of the maintenance period. Data from the first survey of the summer data collection period for this question were lost due to experimenter error, reducing the number of observations. However, of the 21 participants for whom we had complete pre- and post-intervention phase data, 18 reported a shorter time block needed for writing at the end of the intervention than at the beginning; 3 reported no change; and no participants reported a longer time block. If we exclude the three participants who reported no change, the probability of this result or one more extreme occurring by chance is $p < 0.00001$. This is very strong evidence that the intervention enabled participants to write productively in shorter blocks of time.

Items 12–19: Greater use of plans for writing and research A third prediction for which we found strong evidence was that the workshop would increase participants' belief in the value of making explicit plans for their writing and research and would increase their use of such plans over three timescales: long term (1–5 years), medium term (10 weeks), and short term (weekly/daily). Each survey included two questions about each type of plan. The first asked how important/helpful the plan was; participants in both groups and all phases of the experiment saw planning as important/helpful, with mean judgments ranging from 70.71 to 90.67 on a 100-point scale (see Table 2 and Fig. 3).

The second question asked whether the participant had used each type of plan for their own writing during the preceding five weeks. We used a McNemar exact binomial test to evaluate whether participants' propensity to make plans changed during the different phases of the study. At the start of the study, most participants did not have plans for their writing. Many students made such plans during the intervention phase (a McNemar exact binomial test across all durations yielded a $p < 0.000001$). Some students also made plans during the control phase, but not as many (a McNemar exact binomial test across all durations yielded a $p \approx 0.02$). No significant change was observed from the intervention phase to the maintenance phase (see Table 3 and Fig. 3).

Weak Evidence for Hypothesis 4: Improvements in Subjective Well-Being

We predicted that the intervention would "improve [participants'] subjective well-being, as measured by three scales: The Flourishing Scale (Diener et al., 2010), the PHQ-4 (Lowe et al., 2010) and a novel 'Healthy Lifestyles' scale." We found positive evidence for this prediction on only one item of the "Healthy Lifestyles" scale—a question asking how often the participants paused for reflection or positive thinking. No changes were observed on the other well-being measures.

Table 2 Importance/helpfulness of writing plans for graduate students and postdocs, on a scale from 0 (very unhelpful) to 100 (very helpful). Data are means and (standard deviations)

	ConInt Group			IntMain Group		
	Survey 1	Survey 2	Survey 3	Survey 1	Survey 2	Survey 3
Long-term plans (1–5 years)	75.25 (28.68)	72.60 (26.87)	89.75 (12.51)	70.71 (27.60)	79.19 (25.57)	85.00 (16.79)
Medium-term plans (10 weeks)	81.85 (18.07)	84.27 (12.83)	87.69 (16.03)	72.71 (28.22)	87.47 (11.93)	85.69 (17.80)
Short-term (daily) plans	78.14 (19.14)	79.36 (17.59)	86.13 (18.13)	87.00 (11.84)	86.88 (12.28)	83.88 (18.06)

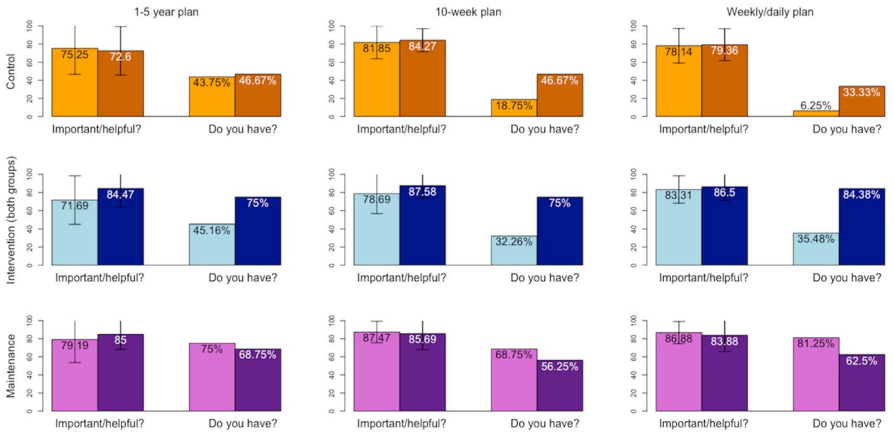

Fig. 3 Planning. Each pair of columns shows the data at the start (lighter column) and end of each 5-week phase. The first pair of columns shows judgments of importance/helpfulness; the second pair shows the proportion of students who had such a plan for their own writing. Error bars are standard deviations

Table 3 Proportion of students reporting that they had used each type of plan for their own writing and research over the preceding five weeks. Data are 'yes' responses over total responses

	ConInt Group			IntMain Group		
	Survey 1	Survey 2	Survey 3	Survey 1	Survey 2	Survey 3
Long-term plans (1–5 years)	7/16	7/15	12/16	7/16	12/16	11/16
Medium-term plans (10 weeks)	3/16	7/15	13/16	3/16	11/16	9/16
Short-term (daily) plans	1/16	5/15	14/16	6/16	13/16	10/16

More frequent pausing for reflection or positive thinking The "reflection or positive thinking" question was one of six questions asking participants how many days they had been sick or ill over the past week, how often they had gotten enough sleep, and how often they had taken time out for play or relaxation, social interaction, exercise, and reflection or positive thinking. Responses were on a four-point Likert scale, from 'Never or rarely (0–1 days)' to 'Every or almost every day (6–7 days).' The intervention affected only the "reflection or positive thinking" question. A repeated measures ANOVA showed an interaction for this item, as participants reported a significant increase in frequency of pausing for reflection/positivity during the intervention phase only, $F_{survey}(2,58)=3.97$, $p=0.02$, $\eta_p{}^2=0.12$; $F_{survey*group}(2,58)=5.80$, $p=0.01$, $\eta_p{}^2=0.17$; $F_{group}(1, 29)=0.03$, $p=0.87$.

No effect on other "Healthy Lifestyle" measures The intervention had no measurable effect on the frequency of participants' play/relaxation, social interaction, exercise, sleep, or frequency of sickness/illness. An ANOVA of Healthy Lifestyle data including all phases and groups showed a main effect of survey only, reflecting the fact that

students in all groups showed a slight decline in healthy habits during the first five weeks of each term and a slight improvement during the second five weeks, although post hoc analyses showed no significant differences between any pairs of surveys. ($F_{survey}(2, 58) = 3.32$, $p = 0.04$, $\eta_p^2 = 0.10$; $M_{S2\text{-}S1} = -0.65$ (SE = 0.44), $M_{S3\text{-}S2} = 1.23$ (SE = 0.54), $BF_{survey} = 1.17$,). There were no other main effects or interactions, ($F_{survey*group}(2,58) = 1.01$, $p = 0.37$; $F_{group}(1,29) = 1.89$, $p = 0.18$; $BF_{full\ model} = 0.28$). A repeated measures ANOVA for the exercise question ('During the past week, how many days did you exercise for at least 20 min?') showed an effect of survey only, reflecting the fact that students reported less frequent exercise halfway through each 10-week study period than they had at the beginning, and their exercise frequency remained low at the end of each period, $F_{survey}(2, 58) = 7.58$, $p = 0.001$, $\eta_p^2 = 0.21$; $F_{survey*group}(2, 58) = 0.01$, $p = 0.99$; $F_{group}(1,29) = 2.65$, $p = 0.11$; $BF_{full\ model} = 5.24$, $BF_{survey} = 33.80$.

No effect on "Flourishing" The intervention had no measurable effect on participants' scores on the Flourishing Scale (Diener et al., 2010), an 8-item measure of self-perceived well-being in areas such as relationships, self-esteem, purpose, and optimism. An ANOVA including all phases and groups showed only a marginal effect of group, as the Con + Int group reported slightly higher levels of flourishing than the Int + Main group overall, $F_{survey}(2, 58) = 2.20$, $p = 0.12$; $F_{survey*group}(2, 58) = 1.85$, $p = 0.17$; $F_{group}(1,29) = 4.05$, $p = 0.05$, $\eta_p^2 = 0.12$; $BF_{full\ model} = 0.52$.

No effect on anxiety/depression screener The intervention had no measurable effect on participants' scores on the 4-item Patient Health Questionnaire (Khubchandani et al., 2016; Kroenke et al., 2009), an ultra-brief screener for anxiety and depression symptoms. An ANOVA including all phases and groups showed a main effect of survey only, as students reported the highest levels of anxiety/depression at the beginning of each ten-week academic term, lower levels halfway through the term, and the lowest levels at the end, regardless of when they received the intervention, $F_{survey}(2,58) = 6.93$, $p = 0.002$, $\eta_p^2 = 0.19$. There were no other main effects or interactions ($F_{survey*group}(2, 58) = 0.50$, $p = 0.61$, $F_{group}(1, 29) = 0.05$, p = 0.82; $BF_{full\ model} = 2.15$; $BF_{survey} = 19.90$).

Weak Evidence for Hypothesis 1: Increases in the Amount of Participants' Writing

Hypothesis 1 was that the intervention would "increase the amount of time doctoral students and postdocs spend writing." Many data for this analysis were lost, and the remaining data yielded no evidence to support our prediction. Unexpectedly, however, we found that the intervention did cause participants to write more by another measure: It made them more likely to generate new writing, either by revising an old research statement or drafting a completely new one, even when new writing was not requested or required.

Measure 1. No effect on time spent writing At the beginning of the study, participants downloaded a free software application (RescueTime) to track how they spent

Table 4 Time logged via RescueTime app. Values are means and (standard deviations) in hours of 'very productive' (writing) plus 'productive' (research) time

	Week 1	Week 2	Week 3	Week 4	Week 5	Total[a]
ConInt group, Control phase	12.8 (8.3)	12.0 (6.7)	10.8 (4.8)	11.6 (5.1)	9.6 (7.7)	**50.44 (24.61)**
ConInt group, Intervention phase	16.10 (6.4)	14.90 (9.1)	11.51 (8.0)	12.15 (8.4)	18.45 (8.5)	**67.12 (38.6)**
IntMain group, Intervention phase	17.64 (10.3)	18.31 (7.7)	19.45 (10.7)	18.96 (7.6)	16.87 (7.1)	**77.19 (46.3)**
IntMain group, Maintenance phase	21.6 (6.9)	18.5 (8.1)	17.6 (6.9)	16.0 (11.7)	20.9 (7.5)	**92.0 (37.5)**

[a]Values in this column are the means and standard deviations of individual participants' totals across the five-week period

time on their computers. At the end of each phase of the experiment, participants reported their hours and minutes of productive and very productive time for the preceding five weeks. However, many data were lost when participants failed to restart or reinstall RescueTime (e.g., after updating their operating systems). In the end, we collected only 196/320, or 61.25%, of the planned observations.

Time spent writing did increase in the Con + Int group, from 50.44 h over five weeks during the five-week control phase to 67.12 h during the intervention phase. It increased even more in the Int + Main group, from 77.19 h during the intervention phase to 92.0 h during the maintenance phase (see Table 4). However, the preregistered, mixed-design ANOVA comparing the total time logged in Weeks 1–5 to Weeks 6–10 (intervention phase for the Con + Int group; maintenance phase for the Int + Main group) revealed only an effect of group, as the Int + Main group logged more total time across all ten weeks than the Con + Int group, $F(1,17) = 4.53$, $p = 0.048$, $\eta_p^2 = 0.21$. $BF_{session} = 1.82$. This is the general pattern that we would expect to see if the intervention helped people write more and the maintenance group helped them keep writing. However, there was no effect of survey, $F(1,17) = 3.21$, $p = 0.09$, nor did the predicted interaction appear, $F(1,17) = 0.87$, $p = 0.36$. $BF_{full\ model} = 1.34$. Thus, although time spent writing did increase from control to intervention and from intervention to maintenance, the relatively high variation between individuals and relatively low variation between groups (see Fig. 4) yielded an ANOVA with null effects.

Measure 2. Weak evidence that the intervention made people more likely to generate new writing Unexpectedly, we found that participants were slightly more likely to upload new writing about their research after the intervention and maintenance phases of the experiment than after the control phase. With each survey, participants uploaded a statement describing their research in approximately 1,000 words. They were free to either upload an old statement or write a new one. In the Con + Int group, only 3/16 (19%) of students submitted any new writing after the control phase; the remaining 13 students resubmitted the identical statement from the previous survey. After the intervention phase, 9/16 (56%) submitted new writing,

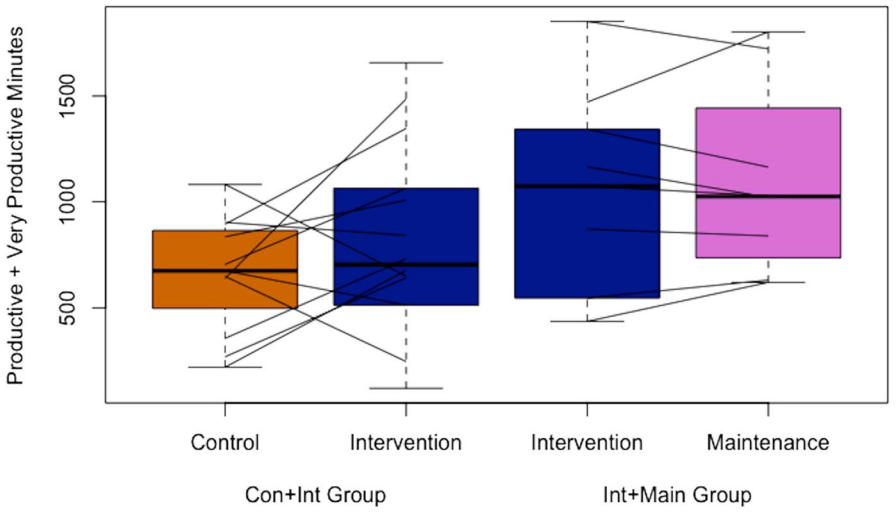

Fig. 4 Time spent writing. Average minutes per week of 'productive' (i.e., non-writing research activities) plus 'very productive' (i.e., writing) time logged by participants. Each line connects an individual's averages during the two phases of the experiment (Control and Intervention for the ConInt group; Intervention and Maintenance for the IntMain group). The thicker horizontal lines inside each box represent the medians

a marginally insignificant difference at $\alpha = 0.05$, but significant at $\alpha = 0.1$ (McNemar exact test $p = 0.07$). In the Int + Main group, 8/16 (50%) submitted new writing after the intervention phase, and 7/16 (44%) did so after the maintenance phase, which was not a significant difference (McNemar exact test, $p = 1$). In other words, we found weak evidence that the shift from control to intervention elicited new writing about research from more participants, even when such new writing was not requested or required. No such change was observed in the shift from intervention to maintenance.

No Evidence for Hypothesis 2: Improvements in the Quality of Participants' Writing

With each survey, participants uploaded a 1,000-word research statement to be scored by two independent raters. Because the intervention caused participants to generate more new writing, and because new writing received lower scores than old writing, the intervention actually caused scores for writing quality to drop slightly. Ratings for revised statements were slightly worse (by an average of -2.32 points on a 48.75-point scale), than ratings for the original statements, and ratings for brand-new statements (describing entirely different research projects) were substantially lower (by an average of -12.50 points) than the originals. Thus, a repeated-measures ANOVA revealed a main effect of survey, $F_{\text{survey}}(1.56,\ 58) = 6.33$, $p = 0.007$, $\eta_p^2 = 0.18$; $F_{\text{survey*group}}(2,58) = 0.53$, $p = 0.59$, $F_{\text{group}}(1,29) = 0.11$, $p = 0.74$; $M_{\text{S2-S1}} = -0.96$ (SE = 0.47), $M_{\text{S3-S2}} = -1.40$ (SE = 0.76); $BF_{\text{survey}} = 11.22$, $BF_{\text{full model}} = 1.78$. Post hoc

tests showed that scores at the end of each ten-week study period were lower than scores at the beginning.

Discussion

This study tested a five-week intervention during which students wrote regularly, made plans for writing and research, learned about structure and style in four academic genres (literature reviews, journal articles, funding proposals, and presentations) and exchanged feedback on writing in progress. The intervention was no more expensive to offer than a standard graduate seminar, requiring only one faculty member to lead it, with graduate student volunteers leading the maintenance groups. We pre-registered four hypotheses about the effects of this intervention, predicting that it would (1) increase the amount of time participants spent writing; (2) improve the quality of their writing; (3) make their writing-related beliefs and practices more positive and sustainable, and (4) improve their subjective well-being.

We found strong evidence on all measures for Hypothesis 3: The intervention clearly improved participants' writing-related beliefs and practices. After the intervention, participants liked writing more, felt more confident in their own writing abilities, and no longer held harmful misconceptions such as, "I need to be inspired in order to write" (Boice, 1990). The intervention taught students to write productively in shorter blocks of time—a key skill for academics, and one that graduate students often lack. The intervention also helped students with planning, which advisors identify as the single greatest obstacle to successful completion of the PhD (D'Andrea, 2002). No longitudinal studies yet exist to show how such changes affect later student success, but it seems likely that such improved attitudes and behavior should continue to benefit students even after the end of the study.

We also predicted that the intervention would improve students' subjective well-being, but we saw a change on only one measure: Students paused more frequently for reflection and positive thinking after completing the intervention. This undoubtedly reflected the practice during the intervention of having students fill out a shared daily writing log that included a daily positive reflection (e.g., something you feel grateful for today). Although we did not see a change on other well-being measures, we remain optimistic that longer-term participation in the writing workshop might yet yield improvements in well-being by decreasing isolation (Ali & Kohun, 2007) and providing sustained instrumental, informational, and emotional support to participants (Charles et al., 2021; Thoits, 2011).

Another hypothesis to receive little support was that participants would write more. We originally defined this by the hours of writing time participants logged on their computers, and by that definition the intervention was not very effective; the increases in writing time were modest, and there was wide individual variation among participants. However, when participants had a choice between uploading an old piece of writing and generating new writing, the intervention made them more likely to generate something new. On the baseline survey, students were required to upload a research statement of approximately 1,000 words. On subsequent surveys, they could either re-upload the identical statement, revise it, or write a completely

different one. Students were more likely to revise their statements or write new ones at the end of the intervention or maintenance phases of the experiment, whereas students at the end of the control phase were more likely to re-use the identical statement from before. This was an unexpected form of evidence and an unplanned measure, but it is consistent with our hypothesis that the intervention made students write more.

Finally, we hypothesized that the intervention would improve the quality of students' writing, and so we were puzzled to find that scores for writing quality actually dropped slightly. Upon examination, this turned out to be an artifact of new writing generated by the intervention, as described earlier. The research statements that many students uploaded on the baseline survey had previously been submitted as part of graduate fellowship applications. These statements were highly polished. On later surveys, students who revised their research statements or drafted new ones ended up with something less polished, which actually earned lower ratings. This finding does not deter us from thinking that the intervention does help students write better over the long term. Anecdotally, we have seen that the first change students experience in their writing practice is a reduction in writing-related anxiety, a feeling of getting "unstuck," and a new willingness to experiment with their writing. (This effect was observed in the present study, where the intervention caused participants to generate new writing even when they did not have to.) As participants continue in the workshop over time, they see improvements in quality. Thus, we interpret the null effects in this intervention only as evidence that five weeks is not long enough to see improvements in writing quality. We remain optimistic that longer intervention and measurement periods will show such improvements.

At the time of this writing, we are implementing a program of cascading graduate mentorship at our university, featuring working groups like the present study's maintenance groups. In this program, trained graduate students and postdocs will lead their peers in groups that follow the writing workshop model described here. The new program will extend the study of the writing workshop model by testing whether the intervention is effective when groups are led by graduate students rather than a faculty member; whether the program can be scaled up from one or two groups at a time to approximately 40 concurrent groups; and by measuring longer-term effects such as projects completed, program milestones achieved, workplace satisfaction, and the development of research skills.

In conclusion, the present study tested a writing workshop intervention that emphasized planning, low-stakes feedback, peer mentorship, and community. The results indicate that just five weeks of participation in the workshop produced dramatic positive changes in students' writing behavior, their attitudes about academic writing, and most importantly in their views of themselves as writers. In our eyes, these effects alone make the writing workshop worth offering. Moreover, our experience has been that improvements in writing attitudes and behavior are only the beginning. With ongoing participation in the workshop, students continue to develop as scholars, writers and mentors. Many students remain in our workshop for years, and some (including the second through fifth authors of this paper) go on to lead their own workshops to support their peers. We have described our workshop here and posted its materials online (https://osf.io/ftuhp/) in order to facilitate the

creation of as many such groups as possible. Our hope is that this relatively low-cost and easy-to-implement model can be adopted by other doctoral programs to support the research and writing success of doctoral students across a range of disciplines.

Acknowledgements The authors wish to thank the doctoral students and postdoctoral fellows who participated in the study, as well as all those who helped to develop and refine the writing workshop model over the past decade. Thanks also to all the faculty who lent their support and sent their doctoral students, especially Professors Nancy Aguilar-Roca, Jessica Borelli, Michael Gottfredson, Emily Grossman, Michael McBride, Duncan Pritchard, Stephanie Reich, Jenny Reinehard, Ramesh Srinivasan, and Joachim Vandekerckhove. Thanks to Clara Schultheiss for help with logistics, and to David Feldon for advice on methods and measures.

Authors' Contributions **Barbara Sarnecka:** Conceptualization, Methodology, Investigation, Resources, Writing–original draft, Writing–review & editing, Supervision. **Paulina Silva:** Formal analysis, Investigation, Data curation, Writing–review & editing, Visualization, Project administration. **Jeff Coon:** Validation, Formal analysis, Investigation, Visualization, Writing–review & editing. **Darby Vickers:** Methodology, Investigation, Writing–review & editing. **Rena Goldstein:** Validation, Investigation, Writing–review & editing. **Jeff Rouder:** Validation, Formal analysis, Data curation, Visualization.

Data availability Preregistrations, workshop materials, surveys, and data are available at https://osf.io/ftuhp/.

Code availability Analysis code is available at https://osf.io/ftuhp/.

Declarations

Conflicts of interest Authors declare no competing interests.

References

Abele, A. E., & Wiese, B. S. (2008). The nomological network of self-management strategies and career success. *Journal of Occupational and Organizational Psychology, 81*(4), 733–749. https://doi.org/10.1348/096317907X256726

Aitchison, C., Catterall, J., Ross, P., & Burgin, S. (2012). 'Tough love and tears': Learning doctoral writing in the sciences. *Higher Education Research & Development, 31*(4), 435–447. https://doi.org/10.1080/07294360.2011.559195

Ali, A., & Kohun, F. (2007). Dealing with social isolation to minimize doctoral attrition – A four stage framework. *International Journal of Doctoral Studies*, *2*.

Austin, A. E. (2002). Preparing the next generation of faculty: Graduate school as socialization to the academic career. *The Journal of Higher Education, 73*(1), 94–122. https://doi.org/10.1080/00221546.2002.11777132

Boice, R. (1990). Faculty resistance to writing-intensive courses. *Teaching of Psychology, 17*(1), 13–17. https://doi.org/10.1207/s15328023top1701_3

Burnett, P. C. (1999). The Supervision of doctoral dissertations using a collaborative cohort model. *Counselor Education and Supervision, 39*(1), 46–52. https://doi.org/10.1002/j.1556-6978.1999.tb01789.x

Caffarella, R. S., & Barnett, B. G. (2000). Teaching doctoral students to become scholarly writers: The importance of giving and receiving critiques. *Studies in Higher Education, 25*(1), 39–52. https://doi.org/10.1080/030750700116000

Carter-Veale, W. Y., Tull, R. G., Rutledge, J. C., & Joseph, L. N. (2016). The dissertation house model: Doctoral student experiences coping and writing in a shared knowledge community. *CBE—Life Sciences Education*, *15*(3), ar34. https://doi.org/10.1187/cbe.16-01-0081

Charles, S. T., Karnaze, M. M., & Leslie, F. M. (2021). Positive factors related to graduate student mental health. *Journal of American College Health*, 1–9. https://doi.org/10.1080/07448481.2020.1841207

Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator: The Professional Journal of the American Federation of Teachers*, *15*(3). https://www.aft.org/ae/winter1991/collins_brown_holum. Accessed 3 Sep 2020.

D'Andrea, L. M. D. (2002). Obstacles to completion of the doctoral degree in colleges of education: The professors' perspective - ProQuest. *Educational Research Quarterly, 25*(3), 42.

Davis, G. (2005). Doctors without orders. *American Scientist, 93*(3), S1.

Delyser, D. (2003). Teaching graduate students to write: A seminar for thesis and dissertation writers. *Journal of Geography in Higher Education, 27*(2), 169–181. https://doi.org/10.1080/03098260305676

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research, 97*(2), 143–156. https://doi.org/10.1007/s11205-009-9493-y

Dweck, C. S. (2016). What having a "growth mindset" actually means. *Harvard Business Review*, 4.

Feldon, D. F., Jeong, S., Peugh, J., Roksa, J., Maahs-Fladung, C., Shenoy, A., & Oliva, M. (2017). Null effects of boot camps and short-format training for PhD students in life sciences. *Proceedings of the National Academy of Sciences, 114*(37), 9854–9858. https://doi.org/10.1073/pnas.1705783114

Gardner, S. K. (2013). The challenges of first-generation doctoral students. *New Directions for Higher Education, 2013*(163), 43–54. https://doi.org/10.1002/he.20064

Gardner, S. K., & Mendoza, P. (2012). *On Becoming a scholar: Socialization and development in doctoral education*. Stylus Publishing, LLC.

Hoffer, T. B., Dugoni, B. L., Sanderson, A. R., Sederstrom, S., Ghadialy, R., & Rocque, P. (2001). *Doctorate recipients from United States universities: Summary report 2000. Survey of earned doctorates*. https://eric.ed.gov/?id=ED459639

Johnson, R. W., & Conyers, L. M. (2001). Surviving the doctoral dissertation: A solution-focused approach. *Journal of College Counseling, 4*(1), 77–77.

Khubchandani, J., Brey, R., Kotecki, J., Kleinfelder, J., & Anderson, J. (2016). The psychometric properties of PHQ-4 depression and anxiety screening scale among college students. *Archives of Psychiatric Nursing, 30*(4), 457–462.

Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Löwe, B. (2009). An ultra-brief screening scale for anxiety and depression: The PHQ–4. *Psychosomatics, 50*(6), 613–621. https://doi.org/10.1016/S0033-3182(09)70864-3

Löwe, B., Wahl, I., Rose, M., Spitzer, C., Glaesmer, H., Wingenfeld, K., Schneider, A., & Brähler, E. (2010). A 4-item measure of depression and anxiety: Validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *Journal of Affective Disorders, 122*(1–2), 86–95.

Leslie, S.-J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science, 347*(6219), 262–265. https://doi.org/10.1126/science.1261375

Maher, M. A., Feldon, D. F., Timmerman, B. E., & Chao, J. (2014). Faculty perceptions of common challenges encountered by novice doctoral writers. *Higher Education Research & Development, 33*(4), 699–711. https://doi.org/10.1080/07294360.2013.863850

Meyer, M., Cimpian, A., & Leslie, S.-J. (2015). Women are underrepresented in fields where success is believed to require brilliance. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00235

Ng, T. W. H., Eby, L. T., Sorensen, K. L., & Feldman, D. C. (2005). Predictors of objective and subjective career success: A meta-analysis. *Personnel Psychology, 58*(2), 367–408. https://doi.org/10.1111/j.1744-6570.2005.00515.x

Rockey, S. (2013). Individual development plans for NIH-supported trainees. *NIH Extramural Nexus*. https://nexus.od.nih.gov/all/2013/07/23/individual-development-plans-for-nih-supported-trainees/. Accessed 20 Jun 2019.

Rockquemore, K. A. (2010). Shut up and write. *Inside Higher Ed*. https://www.insidehighered.com/advice/2010/06/14/shut-and-write. Accessed 4 Sep 2020.

Roksa, J., Feldon, D. F., & Maher, M. (2018). First-generation students in pursuit of the PhD: Comparing socialization experiences and outcomes to continuing-generation peers. *The Journal of Higher Education, 89*(5), 728–752. https://doi.org/10.1080/00221546.2018.1435134

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119–144. https://doi.org/10.1007/BF00117714

Sarnecka, B. W. (2019). *The writing workshop: Write more, write better, be happier in academia*. Author. https://osf.io/5qcdh/. Accessed 15 Aug 2019.

Silvia, P. J. (2007). *How to write a lot: A practical guide to productive academic writing*. American Psychological Association.
Smallwood, S. (2004). A week at camp dissertation. *CHE*. https://www.chronicle.com/article/a-week-at-camp-dissertation/. Accessed 5 Aug 2020.
Sowell, R. S., Bell, N. E., & Kirby, S. N. (2010). *Ph.D. completion and attrition: Policies and practices to promote student success*. Council of Graduate Schools. https://cgsnet.org/phd-completion-and-attrition-policies-and-practices-promote-student-success-0. Accessed 15 Aug 2019.
Stolzenberg, E. B., Eagan, K., Zimmerman, H. B., Berdan Lozano, J., Cesar-Davis, N. M., Aragon, M. C., & Rios-Aguilar, C. (2019). *Undergraduate teaching faculty: The 2016–2017 HERI faculty survey*. Higher Education Research Institute, Graduate School of Education & Information Studies, University of California, Los Angeles. https://www.heri.ucla.edu/monographs/HERI-FAC2017-monograph.pdf. Accessed 23 Jun 2019.
Thoits, P. A. (2011). Mechanisms linking social ties and support to physical and mental health. *Journal of Health and Social Behavior, 52*(2), 145–161. https://doi.org/10.1177/0022146510395592
Weidman, J. C., & Stein, E. L. (2003). Socialization of doctoral students to academic norms. *Research in Higher Education, 44*(6), 16.

Barbara W. Sarnecka is Professor of Cognitive Sciences and Associate Dean of Graduate Studies and Research for Social Sciences at the University of California, Irvine.

Paulina N. Silva is a doctoral candidate in Cognitive Sciences at the University of California, Irvine.

Jeff Coon is a doctoral candidate in Cognitive Sciences at the University of California, Irvine.

Darby C. Vickers is Visiting Assistant Professor of Philosophy at the University of San Diego.

Rena B. Goldstein is a doctoral candidate in Philosophy at the University of California, Irvine.

Jeffrey N. Rouder is Professor and Falmangne Endowed Chair of Cognitive Sciences at the University of California, Irvine.