Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Feasibility of Longitudinally Measuring Psychosocial Outcomes in a Nursing Student Doula Program: A Single-Arm Pilot Study

Rita Ewaz

rewaz@ucsd.edu

University of California San Diego

Annika Huilin Li

University of California San Diego

Anneka Kleine

University of California San Diego Medical Center

Maggie Conway

University of California San Diego Medical Center

April Ricotta

University of California San Diego Medical Center

Research Article

Keywords: feasibility study, doula, nursing students

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11246216/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations:
No competing interests reported.

Tables and Figures are available in the Supplementary Files section.

Feasibility of Longitudinally Measuring Psychosocial Outcomes in a Nursing Student Doula Program: A Single-Arm Pilot Study

Rita Ewaz[1*], Annika Huilin Li[1], Anneka Kleine[2], Maggie Conway[2], April Ricotta[2]

[1] Department of Education Studies, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

[2] Women and Infants Services, UC San Diego Health, 9300 Campus Point Drive, La Jolla, CA 92037, USA

***Corresponding author:** Rita Ewaz, Department of Education Studies, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA. Email: rewaz@ucsd.edu

Author email addresses

Rita Ewaz: rewaz@ucsd.edu

Annika Huilin Li: hul029@ucsd.edu

Anneka Kleine: ankleine@health.ucsd.edu

Maggie Conway: m2conway@health.ucsd.edu

April Ricotta: aricotta@health.ucsd.edu

ORCID: Rita Ewaz, 0009-0006-5350-1775

Abstract

Background: Doula training has been associated with positive professional and psychosocial outcomes among nursing students, including increased confidence, preparedness, and perceived professional growth. However, whether these outcomes can be reliably measured over time in this population remains untested. Establishing the feasibility of the longitudinal design is an important step toward definitively understanding the program's effect on psychosocial outcomes across multiple timepoints.

Methods: This single-arm pilot assessed the implementation and research feasibility of a nursing student doula training program delivered with a concurrent longitudinal research component. Twenty-two nursing students were followed across an academic year to assess the delivery and completion of the intervention (a 16-hour training, three mentored labor-support shifts, and ongoing supervised volunteer practice), as well as recruitment, retention, data completeness, and repeated administration of psychosocial measures. Four self-reported outcomes: general self-efficacy, nurse self-concept, psychological sense of program membership, and clinical activity exposure, were assessed at three points during the academic year following the didactic training. The first two points fell within the mentored-shift phase in a clinical setting. Feasibility indicators were recruitment, retention, and data completeness, and per-wave internal consistency of the new measure. Trajectories were analyzed using linear mixed models treating time as categorical, with nonparametric sensitivity checks.

Results: Recruitment and retention were 100% (22) across all three waves (66/66 scale-level observations). The four outcomes traced three distinct patterns: nurse self-concept rose substantially by the final wave ($b = 1.51$, 95% CI [1.18, 1.83], $p < .001$), psychological sense of membership peaked at the mentored-shift phase before returning to baseline (rise $b = 0.73$, 95%

CI [0.42, 1.05], $p < .001$; decline $b = 0.69$, 95% CI [0.38, 1.00], $p < .001$), and general self-efficacy and clinical activity exposure remained stable. The lack of a pre-training baseline prevented estimation of change directly attributable to the program.

Conclusions: Longitudinal evaluation of the psychosocial outcomes of implementing a doula training program for nursing students was feasible across both domains, with full recruitment and retention and acceptable instrument performance over time. Findings support a larger controlled evaluation incorporating pre-training baseline measurement, refined instrumentation, and a comparison condition.

Keywords: feasibility study, doula, nursing students

Key messages regarding feasibility

1) What uncertainties existed regarding the feasibility?

- It was uncertain whether nursing students could be recruited into a volunteer doula training program with a concurrent research component. If they could be retained across a nine-month academic year given competing clinical and academic demands, and whether validated psychosocial instruments (measuring self-efficacy, professional self-concept, and sense of membership) could be repeatedly administered in this training context with acceptable completion and reliability. It was also uncertain whether the program itself could be delivered as designed. Specifically, it was unclear whether students would complete the didactic training and mentored shifts, and whether opportunities to practice doula care would remain available throughout the academic year.

2) What are the key feasibility findings?

- Recruitment yielded a full cohort of 22 participants, retention was 100% across all three measurement timepoints and repeated administration of the instruments was achieved with high completion rates and acceptable internal consistency. Results detected meaningful change trajectories, including sustained growth in professional self-concept and a transient peak in sense of membership. The full cohort completed the 16-hour didactic training and clinical activity exposure remained stable across the year, indicating sustained access to hands-on opportunities.

3) What are the implications of the feasibility findings for the design of the main study?

- A definitive study is feasible with design modifications: a pre-training baseline measurement is needed to estimate effect magnitude; a comparison or waitlist-control condition would allow training effects to be distinguished from maturation across the

academic year; and refinement of the study-developed clinical activity exposure measure, including a clearly specified recall window, would improve its sensitivity to change. Because the program was delivered as a continuing education course, whether it can be embedded within a required nursing curriculum remains an open question for the main study.

Background

Doulas are non-clinical support professionals trained to provide physical, emotional, and educational support before, during, and after labor and birth (Chaudhary et al., 2025). They have increasingly been recognized as valuable members of the labor and delivery care team, though their consistent integration remains a gap that individual institutions are still working to address (Perkins & Isaacs, 2025). Continuous labor support, including support provided by doulas, has been associated with fewer negative appraisal of birth experience, higher rates of spontaneous vaginal delivery, and lower cesarean rates (Bohren et al., 2017). Among patients who have had doula support, particularly those from diverse or underserved backgrounds, promising results of improved maternal wellbeing, and postpartum mental health have been found (Falconi et al., 2022). Doulas may serve as a bridge to address gaps in healthcare for underserved populations (Marudo et al., 2023); this may include cultural and linguistic congruence, long-term partnership and understanding of needs, and a consistent birthing companion, addressing common gaps in comprehensive medical support.

There is a gap in understanding how providing doula support contributes to the long-term development and growth, both professionally and personally, of nursing students, and how this affects their practice. Training nursing students to provide continuous labor support has been an avenue for expanding access to doula care (Boire & Lucas, 2022). Published examples remain concentrated in a small number of programs, and have been offered as elective or extracurricular options rather than as required coursework (O'Brien & Hotelling, 2018; Jordan et al., 2008). Nursing student doula training may support the development of self-confidence and professional identity; however, these outcomes have been examined only to a limited extent with the rigor or longitudinal design needed to establish whether related constructs, such as self-efficacy and

sense of belonging, can be reliably measured in this population (Hinde & Lanning, 2021; Jordan et al., 2008). Existing studies of doula-related training for nursing students have examined narrower, shorter-term outcomes, such as preparedness and self-confidence (Hinde & Lanning, 2021) and professional identity (Jordan et al., 2008), rather than the longitudinal, validated measurement needed to characterize long-term development. Published evaluations of doula training programs for nursing students, such as Jordan et al. (2008), have generally relied on client satisfaction data and student testimonials. This gap in quantitative, longitudinal data on how doula education supports nursing students' growth and development is the focus of the current study.

Self-Concept

Self-concept refers to an individual's perception of themselves, formed through experience within their environment and shaped by interactions with and evaluations from others (Shavelson et al., 1976). It is a multidimensional psychosocial construct, organized hierarchically from a general sense of self into distinct domains, including academic, personal, and physical self-concepts (Marsh & Craven, 2006; Shavelson et al., 1976).

Nursing self-concept, however, refers to one's subjective experience as a nurse, encompassing the ability to communicate and lead effectively, one's attitudes and beliefs about the profession, and one's overall nursing knowledge (Miao et al., 2024; Angel, Craven, & Denson, 2012). Overall, a healthy nursing self-concept contributes to shared decision-making and valuing colleagues, while positively challenging individuals to grow professionally (Cowin et al., 2008). In doula training specifically, students learn to respond emotionally to patients and draw on their lived experiences to provide support. Measures of nursing self-concept capture this

development across domains such as caring, knowledge, staff relations, and leadership, as assessed by the Nurses' Self-Concept Instrument (NSCI; Angel, Craven, & Denson, 2012).

Self-Efficacy

Self-efficacy, defined as an individual's beliefs about their capacity to organize and execute actions necessary to achieve desired outcomes (Bandura, 1997), is an important construct for understanding how and why educational and training programs may be associated with participant outcomes. Self-efficacy provides insight into the mechanisms through which individuals develop confidence in their abilities to perform tasks, communicate knowledge, and apply newly acquired skills in practice (Bandura, 1977).

Research has consistently demonstrated the importance of self-efficacy in educational settings. In a cross-sectional study of 944 first- and second-year nursing students at four vocational colleges in Guangdong, China, self-efficacy mediated the relationship between professional identity and learning engagement. This suggests that students who perceive themselves as capable engage more meaningfully in the learning process (Liao et al., 2024).

In the context of our doula training program, we were interested in examining self-efficacy as a broader indicator of participant growth and development. Examining self-efficacy alongside self-concept measures provides a more holistic understanding of participants' experiences and development. Together, these constructs help us examine participants' acquisition of knowledge and skills and how the program shapes their perceptions of themselves as capable practitioners.

Psychological Sense of Membership

To better understand the student experience in this program, we examined the psychological sense of membership (PSOM) (Goodenow, 1993). PSOM is described as one's

sense of belonging or psychological membership within a school or classroom, the extent to which students feel included, accepted, and supported by both peers and educators (Goodenow, 1993). Although participants are already enrolled in nursing school, we sought to examine whether the doula program enhances their sense of belonging within the nursing profession and whether program participation is associated with changes in this construct. For example, Levett-Jones and Lathlean (2008) found that nursing students' sense of belonging during clinical placements were associated with greater confidence, motivation to learn, and willingness to engage with the healthcare team.

Clinical Activity Exposure

Lastly, we assessed the frequency in which students had the opportunity to perform core doula care activities during clinical shifts. This includes continuous labor support, support in non-pharmacological pain management, cesarean birth support, emotional support, postpartum and early breastfeeding assistance, and patient advocacy. Direct engagement in clinical activities is central to how doulas learn, adapt, and grow in the profession. Mastery experience in authentic practice settings is the primary source of efficacy beliefs (Bandura, 1977), and learning in such settings proceeds through cycles of concrete experience and reflection (Kolb, 1984). Access to those settings, in turn, depends on students feeling included and supported by the clinical team (Levett-Jones & Lathlean, 2008). Assessing the range of clinical opportunities provided and performed, serves the feasibility aims of this study, as it characterizes whether the program reliably delivers the hands-on experiences that make up the core component of the intervention. Because no existing instrument captures the specific activities of doula care within a nursing student volunteer program, we developed a brief six-item measure for this study. Internal consistency was evaluated at each wave as a feasibility outcome.

The current study addressed the following research questions: (1) Is it feasible to deliver the doula training program as designed and to recruit and retain nursing students in a doula training program with a concurrent longitudinal research component and to repeatedly administer validated psychosocial instruments across an academic year? and (2) What are the preliminary trajectories of general self-efficacy, professional self-concept, psychological sense of membership, and clinical activity exposure among nursing students completing doula training?

Methods

Study Design

This was a single-arm, prospective pilot study describing outcome trajectories among nursing students who completed a doula training program. The study had two aims: (1) to assess the feasibility of delivering the training and conducting repeated-measures follow-up in this population, including recruitment, retention, data completeness, and the measurement performance of a newly developed scale; and (2) to describe the shape of four self-report outcome trajectories across a nine-month observation window following the didactic training.

Feasibility was examined in two domains. First, implementation feasibility concerned delivery of the program itself: uptake, completion of the didactic training and the mentored labor-assist shifts, and sustained access to clinical opportunities across the academic year. Second, research feasibility concerned conduct of the longitudinal study: recruitment, retention, data completeness, and the measurement performance of the instruments, including the study-developed scale. The two domains are reported separately because the evidence supporting each rests on different indicators.

The observation window is described as the period following the didactic training rather than as a post-training window. The first two measurement waves (November and December)

occurred while participants were still completing their mentored labor-assist shifts; only the final wave (May) occurred after the conclusion of the mentored phase.

All measurements occurred after the didactic training was delivered; therefore, the study was designed and analyzed as a descriptive trajectory study; no pre-training baseline was available, and the design does not support causal inference regarding training effects. Reporting follows the CONSORT 2010 statement extension to randomized pilot and feasibility trials (Eldridge et al., 2016), applied here to a single-arm design.

Participants and setting

Participants were 22 nursing students (N = 22) enrolled in the inaugural cohort of a Nursing Student Doula Program offered through a teaching hospital in California who completed the training program in August 2025. Participants were recruited in collaboration with the university's extended studies division, which administered the doula program as a continuing education course. Recruitment materials, including a program flyer, were distributed to nursing schools across Southern California. Interested applicants completed an online application that collected demographic information and short-answer responses about their background, language abilities, and motivation for joining the program. A four-person review committee screened applications for basic eligibility, and eligible applicants completed interviews in which reviewers assessed communication skills, attitudes toward patient-centered care, interest in doula work, and readiness for hands-on training. Thirty-six individuals applied, thirty were selected for an interview, and the top twenty-two applicants were selected and enrolled in the program.

Enrollment in the research component was voluntary and separate from enrollment in the course. Students were informed that their decision to participate or decline would have no effect on their standing in the course, access to training opportunities, or relationship with the

university. Inclusion criteria for the study were being at least 18 years of age, current enrollment in a nursing program or graduation from one within the past year, and enrollment in the doula course. No additional exclusion criteria were applied. All 22 trainees enrolled in the study completed all three measurement waves (100% retention; 66 of 66 possible observations).

As a pilot study, no formal power calculation was performed; the sample size was determined by the capacity of the training cohort. A sample of 22 with complete repeated measures is consistent with published recommendations for pilot studies aimed at estimating parameters and assessing feasibility rather than testing hypotheses.

Intervention

The intervention was defined as the 16-hour didactic training, the three mentored labor-assist shifts, and the ongoing supervised volunteer practice that participants undertook across the academic year. The training program, delivered in August 2025, comprised a 16-hour didactic component covering continuous labor support, non-pharmacological pain management, cesarean birth support, emotional support, postpartum and early breastfeeding assistance, and patient advocacy, followed by three mentored labor-assist shifts in which participants provided hands-on labor support under the supervision of an experienced mentor doula. Participants then continued in volunteer practice, scheduling two 12-hour labor-support shifts per month across the remainder of the academic year.

The intervention comprises didactic training combined with supervised practice with patients, delivered through mentored shifts and subsequent supervised volunteer shifts. These shifts are a required program element, not an optional addition, and are the mechanism by which participants apply labor-support skills in real care settings.

Periodic meetings and informational webinars were offered, but not required, throughout the observation window. These are treated as supplemental activities rather than as components of the intervention. Because attendance was optional and variable, exposure to them cannot be expressed as a dose, and adherence to the intervention could not be characterized if they were included in its definition.

Because supplemental activity exposure accumulated concurrently with the follow-up measurement, the observed trajectories reflect the combined influence of the initial training and supervised practice, ongoing supplemental activities, and professional maturation; these sources cannot be disentangled in this design.

Measures

Four self-report outcomes were assessed at each wave. Scale scores were computed as the mean of constituent items, with negatively worded items reverse-coded prior to scoring.

General self-efficacy (SE). General self-efficacy was measured with the General Self-Efficacy Scale (GSE) (Schwarzer & Jerusalem, 1995) a widely used 10-item measure of broad beliefs in one's capacity to cope with difficult demands (sample item: "I can usually handle whatever comes my way"). Items are rated from 1 (not at all true) to 4 (exactly true). The GSE has extensive validation evidence across populations, with internal consistency typically in the 0.76 - 0.90 range.

Nurse self-concept (SC). Professional self-concept was measured with the 14-item Nurses' Self-Concept Instrument (NSCI) (Angel et al., 2012), comprising four dimensions: caring (3 items), knowledge (4 items), staff relations (3 items), and leadership (4 items). Items are rated on an 8-point scale from 1 (definitely false) to 8 (definitely true). Item wording was adapted to the labor-support context (e.g., "doula knowledge" in place of "nursing knowledge");

because the instrument was adapted, published psychometric properties may not fully generalize to this version, which is acknowledged as a limitation. The total mean score was analyzed.

Psychological sense of program membership (PM). Sense of belonging within the program was measured with the 18-item Psychological Sense of School Membership scale (PSSM) (Goodenow, 1993), with "school" adapted to refer to the doula training program. Items are rated on a 5-point scale from 1 (not at all true) to 5 (completely true). Five negatively worded items were reverse-scored. The PSSM has demonstrated good internal consistency ($\alpha \approx 0.77$-0.88) in prior work.

Clinical activity exposure (CAE). A six-item scale developed for this study assessed clinical activity exposure. CAE assessed how often participants performed core labor-support activities during their shifts: providing continuous support during active labor, assisting with non-pharmacological pain management, supporting a patient during cesarean birth, providing antepartum emotional support, assisting with postpartum care and initial breastfeeding, and advocating for a patient. Items were rated on a Likert-scale ranging from 1 (Never) to 5 (Always) and averaged; higher scores indicated more frequent exposure. Because the scale had no prior validation evidence, internal consistency was evaluated at each wave as a feasibility outcome; Cronbach's alpha was ≥ 0.77 at all three waves (0.83, 0.83, and 0.75), meeting the conventional $\geq$ 0.70 threshold and remaining stable across waves.

Demographic and engagement measures. The survey also collected demographic characteristics (age, generation status, language, marital status, parenting status, program year, and prior labor and delivery exposure) and program engagement indicators (number of shifts completed, patients supported in labor, births attended, doula meetings attended, and webinars or

workshops attended). Engagement variables were measured after the didactic training and were therefore not included as model covariates (see Statistical analysis).

Data collection procedures

Data were collected at three time points following the didactic training: November 2025, December 2025, and May 2026, with approximately 3, 4, and 9 months after the training. The November and December waves fell within the mentored-shift phase of the program, and the May wave followed it. Surveys were self-administered electronically via Qualtrics using individualized links, which enabled longitudinal tracking of participants across waves. The uneven spacing (one month between waves 1 and 2; five months between waves 2 and 3) was accounted for analytically by treating time as a categorical factor rather than a linear trend.

Feasibility outcomes

Consistent with pilot study objectives, feasibility indicators were prespecified in two domains. Implementation feasibility indicators were (a) uptake (proportion of trainees offered a place who enrolled in the program), (b) completion of the 16-hour didactic training and the mentored labor-assist shifts, and (c) sustained access to clinical opportunities across the observation window, indexed by the stability of clinical activity exposure and by program engagement indicators. Research feasibility indicators were (a) recruitment (proportion of eligible trainees enrolling in the research component), (b) retention and data completeness across three waves, and (c) measurement performance of the study-developed clinical activity exposure scale (per-wave internal consistency). No formal progression criteria were prespecified.

Statistical analysis

All analyses were conducted in R (version 4.6.0) using the lme4, lmerTest, emmeans, pbkrtest, and psych packages.

Primary models. Each outcome was analyzed with a linear mixed model including a random intercept for participant and time and a three-level categorical fixed effect for time (November, December, May). Categorical coding of time was chosen a priori to accommodate non-monotonic trajectory shapes and the uneven spacing of waves. Models were estimated by restricted maximum likelihood, with Kenward-Roger degrees-of-freedom approximation for small-sample inference. Estimated marginal means with 95% confidence intervals were computed for each wave, and the three pairwise contrasts were tested with Tukey adjustment for multiple comparisons. The intraclass correlation coefficient (proportion of variance attributable to stable between-person differences) was derived from each model's variance components.

Effect reporting. Effects are reported as model-estimated mean differences on each scale's original metric with 95% confidence intervals, rather than as standardized effect sizes. In repeated-measures designs, standardized effects computed against the residual standard deviation are inflated because within-person correlation removes between-person variance from the denominator; raw-scale contrasts with confidence intervals were therefore judged the more interpretable and less misleading summary.

Covariates. Models were unadjusted. This decision reflected four considerations: the single-arm design precludes confounding by indication; stable participant characteristics are absorbed by the random intercept; the sample size (n = 22) does not support additional fixed effects without overfitting; and candidate engagement variables (e.g., supplemental activity attendance) were measured post-baseline and would function as mediators rather than confounders.

Sensitivity analyses. Each mixed-model result was triangulated against non-parametric alternatives: the Friedman test as an omnibus check and Wilcoxon signed-rank tests for pairwise

comparisons. Where residual diagnostics suggested non-constant variance across waves, a heteroscedastic model allowing wave-specific residual variances was fit as an additional check. Model assumptions were evaluated via residual and quantile-quantile plots, and random-effects structures were checked for singular fits.

Missing data. Scale-level data were complete: all 22 participants provided scores for all four outcomes at all three waves (66 of 66 observations per outcome). At the item level, a single CAE item at the May wave was missing for four participants (18/22 responding; item-level completeness 99.7%); scale scores for these cases were computed as the mean of available items. No other items were missing, and no imputation was required.

Ethical considerations

This study was determined to be exempt by the University of California, San Diego Institutional Review Board (protocol #813766). Participants provided oral informed consent prior to data collection and were informed that participation was voluntary and would not affect their standing in the program. Risk was minimal and related primarily to loss of confidentiality. Participants could skip any item or withdraw at any point, and no incentives were offered.

Results

Participant characteristics

Twenty-two nursing students participated (mean age 26.4 years, SD 7.0). Half (n = 11, 50.0%) identified as first- or second-generation immigrants and half as third generation or above. Most were in the advanced years of their nursing program (n = 17, 77.3%), and most reported prior experiential exposure to labor and delivery (n = 16, 72.7%); three (13.6%) reported only didactic or simulated exposure and three (13.6%) reported no prior exposure. Program engagement over the study period was high: participants completed a median of 15 labor-support

shifts (range 9-21; M = 14.9, SD = 2.7) across the nine-month program. Most attended one to two doula meetings (n = 18, 81.8%). Full characteristics are presented in Table 1.

Feasibility outcomes

Implementation feasibility. All 22 students offered a place in the inaugural cohort enrolled in the program and completed the 16-hour didactic training delivered in August 2025. Program engagement across the observation window is reported above and in Table 1. Clinical activity exposure was stable across all three waves (see Outcome trajectories), indicating that opportunities to perform core labor-support activities remained available across the full academic year rather than only at the program's outset.

Research feasibility. All 22 students enrolled in the program's inaugural cohort were invited to participate, and all 22 enrolled (recruitment, 100%). Retention was likewise 100%: all participants completed surveys at all three waves, yielding 66 of 66 possible scale-level observations for each outcome. Item-level completeness was 99.7%: a single CAE item at the May wave was missing for four participants (18/22 responding), and scale scores for these cases were computed as the mean of available items. All other items were complete at all waves. The study-developed CAE scale showed acceptable internal consistency at every wave (Cronbach's $\alpha = 0.83$, 95% CI [0.68, 0.92], in both November and December; $\alpha = 0.75$, 95% CI [0.54, 0.88], in May), meeting the conventional ≥ 0.70 criterion at each administration. No formal progression criteria were prespecified; however, the observed recruitment (100%), retention across three waves (100%), and per-wave reliability of the study-developed scale (all $\alpha \geq 0.75$) would meet thresholds commonly applied in pilot research and support proceeding to a larger controlled evaluation.

Outcome trajectories

The four outcomes showed three distinct trajectory patterns over the nine-month observation window following the didactic training: a sustained rise (nurse self-concept), a transient peak (program membership), and stability (general self-efficacy and clinical activity exposure). Estimated marginal means, omnibus tests, and intraclass correlation coefficients are shown in Table 2; pairwise contrasts with nonparametric sensitivity results are shown in Table 3.

Nurse self-concept (SC). Self-concept differed across waves, $F(2, 42) = 69.76$, $p < .001$. Estimated means rose from 5.78 (95% CI 5.56-6.00) in November to 6.10 (5.87-6.32) in December and 7.29 (7.06-7.51) in May on the 8-point scale. The November-December difference was small and did not reach the adjusted significance threshold (-0.32, 95% CI [-0.65, 0.01], $p = 0.058$), whereas the increases from November to May (-1.51, 95% CI [-1.83, -1.18], $p < .001$) and December to May (-1.19, 95% CI [-1.52 to -0.86], $p < .001$) were large relative to the scale range. Nonparametric results were concordant for the larger contrasts (Friedman $\chi^2(2) = 32.50$, $p < .001$; both $p < .001$) and reached significance for November-December ($V = 58$, $p = .037$), a minor divergence from the mixed-model contrast consistent with its borderline magnitude. Because between-wave variability increased over time (observed SDs 0.26, 0.64, and 0.59), a heteroscedastic model allowing wave-specific residual variances was fit and improved on the homoscedastic model (likelihood ratio $\chi^2(2) = 32.65$, $p < .001$); pairwise conclusions were unchanged. The ICC was 0.27, indicating that most variability in self-concept was within persons over time.

Psychological sense of program membership (PM). Membership followed an inverted-U pattern, $F(2, 42) = 20.73$, $p < .001$. Estimated means rose from 3.75 (95% CI [3.56, 3.94]) in November to 4.49 (4.30-4.68) in December, then returned to 3.80 (3.61-3.99) in May on the 5-point scale. Both the November-to-December increase (-0.73, 95% CI [-1.05, -0.42], $p < .001$)

and the December-to-May decline (0.69, 95% CI [0.38, 1.00], $p < .001$) were significant, while November and May did not differ (-0.05, 95% CI [-0.36, 0.27], $p = .933$). Nonparametric results were fully concordant (Friedman $\chi^2(2) = 22.46$, $p < .001$; Wilcoxon November-December $p < .001$, December-May $p = .001$, November-May $p = .665$). The ICC was 0.11.

General self-efficacy (SE). Self-efficacy was stable, $F(2, 42) = 0.29$, $p = .747$, with estimated means of 3.33 (95% CI 3.19-3.47), 3.31 (3.17-3.46), and 3.37 (3.23-3.52) on the 4-point scale; no pairwise contrast exceeded 0.06 points and none was significant. The Friedman test was likewise nonsignificant ($\chi^2(2) = 3.91$, $p = .142$). One nonparametric pairwise result diverged: the November-May Wilcoxon signed-rank test was significant ($V = 0$, $p = .004$), reflecting that among participants whose scores changed, all shifted in the same upward direction. The corresponding model-estimated difference, however, was only 0.04 points on the 4-point scale (November -May = -0.04, 95% CI [-0.23, 0.15]), indicating directional consistency without practical magnitude; scores near the scale ceiling at all waves (means ≥ 3.31 of a possible 4) constrained observable change.

Clinical activity exposure (CAE). Exposure to core labor-support activities did not change across waves, $F(2, 42) = 0.01$, $p = .992$, with estimated means of 3.33 (95% CI [2.95-3.72]), 3.33 (2.95-3.72), and 3.34 (2.95-3.73); all parametric and nonparametric contrasts were null (Friedman $\chi^2(2) = 0.12$, $p = .943$). The stability was near-absolute at the individual level: every participant reported identical CAE responses in November and December, and scores were largely unchanged by May (median within-person SD across the three waves = 0.10 scale points). This pattern is plausible given the measurement schedule and observed engagement levels: only one month separated the first two waves, and most participants had completed 15 or fewer shifts over the full study period so consecutive reports referenced substantially the same

accumulated clinical experience. Consistent with this interpretation, the ICC was 0.96, indicating that nearly all variance in exposure reflected stable between-person differences, features of each student's placement and shift context, rather than within-person change over time.

Discussion

This pilot yielded encouraging findings for both aims. Feasibility was demonstrated in both domains. The program was delivered as designed: the full cohort completed the didactic training and clinical opportunities remained available across the academic year. The research component performed on every indicator: a full cohort was recruited and retained, and all four instruments, including the study-developed CAE scale, performed acceptably under repeated administration. On the descriptive aim, the four outcomes traced three distinct patterns: a sustained rise in professional self-concept, a transient peak in psychological sense of membership, and stability in general self-efficacy and clinical activity exposure. Together, these four measures offer a comprehensive understanding of program design and student growth. Our results highlight that growth during training is not linear and some domains change while others remain stable.

The sustained growth in self-concept was not unexpected. Doula work involves spending prolonged periods with patients, often during active labor, one of the most vulnerable experiences of a patient's life. Supporting patients physically, advocating for them, and providing sustained emotional care appears to operate bidirectionally: students actively support the patient while simultaneously experiencing a profound sense of reward and purpose from having done so. Consistent with self-concept theory, perceptions of oneself are formed through direct experience and the evaluations of others (Shavelson et al., 1976). This pattern is consistent with experiential accounts of professional learning, in which development is understood to occur through

engagement in authentic practice settings (Kolb, 1984; Levett-Jones & Lathlean, 2008). Whether the relational intensity of doula care specifically contributes to professional self-concept development cannot be established in this design, but it is a plausible mechanism worth testing in a controlled evaluation. We note, however, that in this design the rise in self-concept cannot be attributed to the program itself: ongoing supplemental activities, accumulating clinical experience in students' nursing programs, and professional maturation across the academic year all operate in the same direction.

Second, psychological sense of membership showed a notable transient pattern: scores peaked significantly at T2 (December) before returning to baseline levels by May. During the early months, students were completing their mentored shifts and were in frequent contact with fellow trainees, experienced doulas, and mentors; this sustained interaction during a period of active learning may have supported their sense of belonging within the program. The subsequent decline is consistent with the program's transition toward independence. After December, mentored shifts conclude and students begin practicing on their own, scheduling shifts according to their availability and choosing whether to attend optional supplemental activities. As students become more autonomous in managing patients independently the structured community contact that likely fostered belonging naturally declines. This is complemented by the work of Levett-Jones and Lathlean (2008) where they found belongingness among nursing students was closely tied to the quality and continuity of contact with clinical staff during placements. Together, these findings suggest that community reinforcement during the independent practice phase may help sustain students' sense of belonging.

Third, self-efficacy remained stable across all three waves, which was also not unexpected. Our participants were all current nursing students, and as the inaugural cohort they

were selected through a rigorous application and interview process before being accepted into the doula training program. The stable high baseline scores are consistent with this background: most students likely entered the program with an already high sense of general self-efficacy. This leaves relatively little room for measurable growth which may indicate a possible ceiling effect likely reflecting the competitive selection process. In addition, generalized self-efficacy is conceptualized as a relatively stable, trait-like belief (Schwarzer & Jerusalem, 1995), one that develops gradually through the accumulation of mastery experiences rather than shifting in response to discrete training events (Bandura, 1997). A nine-month observation window may therefore provide neither sufficient time nor the sustained structure needed to shift a broad, generalized sense of efficacy, even as it shapes more specific developing domains such as professional self-concept.

Lastly, clinical activity exposure remained stable across all three waves. We interpret this as a positive implementation finding: students consistently had opportunities to perform the core activities of doula care including, supporting cesarean deliveries, providing emotional support, and advocating for patients, throughout the program rather than only at its outset. Given the unpredictable nature of labor and delivery, this stability suggests that the program reliably generated hands-on clinical opportunities across the full observation period, supporting the claim that the intervention was implemented as designed. This is the principal evidence for implementation feasibility across the academic year. Consistent access to authentic clinical practice is central to experiential learning, through which professional competence and identity are developed (Kolb, 1984; Levett-Jones & Lathlean, 2008).

Limitations

This study has several limitations. First, no pre-training measurement was collected. Without a baseline, the observed trajectories describe development after training rather than change attributable to it. Therefore, the magnitude of any training effect cannot be estimated. Relatedly, the single-arm design provides no comparison condition, so the observed patterns cannot be separated from professional maturation, concurrent nursing-program experiences, or other influences unfolding across the academic year.

Second, the sample was small (n = 22) and comprised a competitively selected inaugural cohort from a single program at a single site. Models were also unadjusted: there were too few observations to support covariate adjustment, and the homogeneity of the cohort left little between-person variation for covariates to explain. Stable participant characteristics are absorbed by the random intercept, and a single-arm design carries no confounding by indication, so the estimates should be read as describing trajectories within this cohort rather than as adjusted effects.

Third, there are measurement limitations. Few validated instruments exist for assessing psychosocial development among nursing students engaged in doula care. This constrained our measurement options. The GSE assesses general rather than nursing- or doula-specific self-efficacy. They are broad and trait-like efficacy beliefs that may be insensitive to program-specific gains, and this may be an explanation for why we observed the uniformly high ceiling-effect scores in this sample. A domain-specific efficacy measure, such as one assessing labor-support skills, may be better suited to future evaluations. The NSCI and PSSM questions were adapted to the doula training context; therefore, the psychometric properties of the original validated measures may not fully generalize to these adapted versions. Lastly, the CAE scale was

developed for this study and requires independent validation in larger samples. The CAE items also asked about the frequency of activities over an unspecified recall window, which, combined with the short interval between the first two waves and the modest number of shifts completed, likely contributed to the identical November and December responses. Finally, all outcomes were self-reported and therefore subject to social desirability and shared-method variance. Additionally, the November and December waves coincided with the mentored-shift phase of the program, so occasion effects on the membership peak cannot be excluded.

Finally, the program was delivered as a continuing education course through the university's extended studies division rather than as a component of a required nursing curriculum. Whether the doula training can be embedded within an undergraduate nursing curriculum, and whether uptake, completion, and access to clinical opportunities would be comparable under that arrangement, remains untested by this study.

Implications

This pilot provides the design parameters for the program's continued evaluation. Three lessons directly shape the next iteration. First, the absence of a pre-training baseline was the principal limitation of the present design; baseline measurement is essential for any future evaluation seeking to attribute change to the program. Second, measurement performance in this cohort identified specific instrument refinements: clarifying the recall window of the clinical activity exposure items, whose responses were insensitive to the short interval between the first two waves, and considering domain-specific efficacy measures better matched to the training content. Third, the transient peak in program membership suggests that belonging may require active maintenance between program events, informing the design of ongoing supplemental support. A second cohort (n = 25) has since begun the program under a protocol refined on these

lessons, incorporating revised survey items, extended follow-up to examine whether favorable post-training profiles persist, and strengthened continuing support. Together with the recruitment (100%) and retention (100%) observed here, these findings suggest the nursing student doula model is feasible both to deliver and to evaluate longitudinally; whether the favorable profiles observed across the year are attributable to the program requires a controlled evaluation with baseline measurement, for which this study establishes the groundwork.

Conclusions

In this single-arm pilot study, it proved feasible to deliver the doula training program as designed. It also proved feasible to recruit a full cohort of nursing students into the concurrent research component, retain every participant across three measurement waves spanning nine months, and repeatedly administer psychosocial instruments. This includes a newly developed measure of clinical activity exposure, with a near-perfect completion and acceptable reliability. The outcome trajectories were interpretable and distinct: professional self-concept rose substantially, sense of program membership peaked during the mentored phase and declined with independence, and general self-efficacy and clinical activity exposure remained stable. These findings support proceeding to a larger controlled evaluation with pre-training baseline measurement, and they offer concrete design parameters, on instrumentation, measurement scheduling, and community support, for that study and for nursing programs considering adoption of the doula training model.

Declarations

Ethics approval and consent to participate: This study was determined to be exempt by the University of California, San Diego Institutional Review Board (protocol #813766; determination date 12/22/2026). Participants were adult nursing students enrolled in the UC San Diego Extended Studies doula program. Participants provided oral informed consent prior to data collection. Participation was voluntary and had no effect on standing in the program.

Consent for publication: Not applicable.

Availability of data and materials: The datasets generated and analysed during the current study are not publicly available to protect participant confidentiality given the small cohort but are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' contributions: R.E. conceptualized and designed the study, collected the data, conducted the analysis, interpreted the results, and wrote the main manuscript text. A.H.L. contributed to the analysis, results, writing, editing, and review of the manuscript. A.K. collected data, reviewed, and edited the manuscript. M.C. collected data, reviewed, and edited the manuscript. A.R. conceptualized and designed the study, collected data, supervised the project, edited and reviewed the manuscript. All authors reviewed and approved the final manuscript.

Acknowledgements: The authors thank Ala B. Garza, DHA, MSN, RN, NEA-BC, Senior Director of Nursing, Neonatal, Newborn, and Obstetrical Inpatient Services, UC San Diego Health, for her support of the doula program and the nursing student initiative. We also thank

Shirish Balachandra, MD, Staff Physician, Innercare Brawley, Diplomate, ABFM, HIV Expert, AAHIVM, for his editorial review of the manuscript. Finally, we thank the nursing student doulas who participated in this study.

608 **References**

609 Angel, E., Craven, R., & Denson, N. (2012). The Nurses' Self-Concept Instrument (NSCI):
610 Assessment of psychometric properties for Australian domestic and international student
611 nurses. *International Journal of Nursing Studies, 49*(7), 880–886.
612 Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change.
613 *Psychological Review, 84*(2), 191–215.
614 Bandura A. Self-efficacy: the exercise of control. New York: W. H. Freeman and Company;
615 1997.
616 Boire, C., & Lucas, L. (2022). Birth companions: Teaching nursing students to become doulas.
617 The Journal of Perinatal Education, 31(1), 14–20. https://doi.org/10.1891/J-PE-D-20-
618 00064
619 Bohren, M. A., Hofmeyr, G. J., Sakala, C., Fukuzawa, R. K., & Cuthbert, A. (2017). Continuous
620 support for women during childbirth. *Cochrane Database of Systematic Reviews*, (7).
621 Chaudhary, P., Rech, J.P., Kumar, G. et al. (2025). Role of doulas across the pregnancy care
622 continuum on maternal and child health: a scoping review. *npj Women's Health, 3*, 63.
623 https://doi.org/10.1038/s44294-025-00109-4
624 Cowin LS, Johnson M, Craven RG, Marsh HW. Causal modeling of self-concept, job
625 satisfaction, and retention of nurses. Int J Nurs Stud. 2008 Oct;45(10):1449-59. doi:
626 10.1016/j.ijnurstu.2007.10.009.
627 Eldridge, S. M., Chan, C. L., Campbell, M. J., Bond, C. M., Hopewell, S., Thabane, L., &
628 Lancaster, G. A. (2016). CONSORT 2010 statement: extension to randomised pilot and
629 feasibility trials. *BMJ, 355*, i5239. https://doi.org/10.1136/bmj.i5239

630 Falconi AM, Bromfield SG, Tang T, Malloy D, Blanco D, Disciglio RS, Chi RW. Doula care
631 across the maternity care continuum and impact on maternal health: Evaluation of doula
632 programs across three states using propensity score matching. EClinicalMedicine. 2022
633 Jul 1;50:101531. doi: 10.1016/j.eclinm.2022.101531.
634 Goodenow, C. (1993). The psychological sense of school membership among adolescents: Scale
635 development and educational correlates. *Psychology in the Schools, 30*(1), 79–90.
636 Hinde TA, Lanning RK. Utilizing Doulas to Improve Nursing Student Preparation and Self-
637 confidence in Providing Labor and Birth Support. Nurse Educ. 2021 May-Jun
638 01;46(3):136-137. doi: 10.1097/NNE.0000000000000899.
639 Jordan, E. T., Van Zandt, S. E., & Wright, E. (2008). Doula care: Nursing students gain
640 additional skills to define their professional practice. Journal of Professional Nursing,
641 24(2), 118–121. https://doi.org/10.1016/j.profnurs.2007.06.018
642 Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and*
643 *development*. Englewood Cliffs, NJ: Prentice Hall.
644 Levett-Jones, T., & Lathlean, J. (2008). Belongingness: A prerequisite for nursing students'
645 clinical learning. *Nurse Education in Practice, 8*(2), 103–111.
646 Liao M, Xie Z, Ou Q, Yang L, Zou L. Self-efficacy mediates the effect of professional identity
647 on learning engagement for nursing students in higher vocational colleges: A cross-
648 sectional study. Nurse Education Today. 2024 Aug 1;139:106225.
649 Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from
650 a multidimensional perspective. *Perspectives on Psychological Science, 1*(2), 133–163.
651 Marudo C, Nicotra C, Fletcher M, Lanning R, Nelson M, Hancock C. Bridging Health
652 Disparities and Improving Reproductive Outcomes With Health Center-Affiliated Doula

Programs. Obstet Gynecol. 2023 Oct 1;142(4):886-892. doi: 10.1097/AOG.0000000000005337.

Miao, C., Liu, C., Zhou, Y., Zou, X., Song, L., Chung, J. W. Y., Tan, W., Li, X., & Li, D. (2024). Nurses' perspectives on professional self-concept and its influencing factors: A qualitative study. BMC Nursing, 23(1), Article 237. https://doi.org/10.1186/s12912-024-01834-y

O'Brien, E., & Hotelling, B. A. (2018). Amazing things happen when student nurses are given birth doula training. The Journal of Perinatal Education, 27(2), 98–103. https://doi.org/10.1891/1058-1243.27.2.98

Perkins, H. D., & Isaacs, C. (2025). Strategies for promoting doula inclusivity in the labor and delivery setting. *Maternal and Child Health Journal, 29*(3), 287–293. https://doi.org/10.1007/s10995-025-04051-4

Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright, & M. Johnston, *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor, UK: NFER-NELSON.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*(3), 407–441.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- QuantManuscriptTablesandFigures.docx