



Can Text Messaging Teach Residents?

A Randomized Controlled Trial

Hovig K. Artinian^{1,2}, Thomas G. Keens¹, Roberta Kato¹, and Todd P. Chang¹

¹Children's Hospital Los Angeles, University of Southern California, Keck School of Medicine, Los Angeles, California; and ²Helen DeVos Children's Hospital, Michigan State University College of Human Medicine, Grand Rapids, Michigan

ORCID ID: 0000-0001-9911-5008 (H.K.A.)

ABSTRACT

Background: Spaced education and the testing effect are both educational methods that increase long-term memory formation; however, these can be difficult to implement during residency training given time constraints. Text messaging is ubiquitous but has not been studied as a spaced education tool.

Objective: Assess if text messaging improves resident knowledge during an inpatient pediatric pulmonary rotation.

Methods: A prospective randomized control study with pediatric residents on a pulmonary inpatient rotation was conducted at an urban free-standing children's hospital between 2016 and 2017. The intervention arm received one daily multiple-choice text message scenario and a scripted teaching text for each response. Both arms received standard pulmonary education. Knowledge was assessed using a 23-item pretest and posttest with unique, nonrepeated items with fair reliability, following iterative revisions. Perceived value of texting was assessed using Likert scales. Paired and unpaired *t* tests compared knowledge and value scores. The difference between pretest and posttest scores (Δ) for both arms was calculated, then compared using an unpaired *t* test. Spearman's rho evaluated maturation bias. Analysis of variance evaluated year of training as a confounding factor.

Results: A total of 65 residents were randomized, with a response rate of 81%. Posttest mean scores were lower than pretest in both arms, attributed to more difficult questions randomized to the posttest. The intervention arm scored higher on the posttest ($P = 0.04$). However, the Δ mean did not show a statistically significant difference ($P = 0.6$). Text messaging was viewed as "effective" by 80% of participants in the intervention arm.

Conclusion: A scripted text messaging intervention is perceived as effective by learners but did not result in measurable increased resident knowledge.

Keywords:

resident education; text messaging; texting

(Received in original form February 26, 2020; accepted in final form May 28, 2020)

Supported by the Barbara M. Korsch Award for Research in Medical Education 2016, Children's Hospital Los Angeles. This funding source did not influence the study design, data collection, and data analysis in the writing of the manuscript, or in the decision to submit the article for publication. No corporate sponsors were involved with this project.

ATS Scholar Vol 1, Iss 3, pp 278–287, 2020

Copyright © 2020 by the American Thoracic Society

DOI: 10.34197/ats-scholar.2020-0025OC

Multiple competing demands on attendings' and residents' time make implementation of effective teaching methods difficult because of the significant time commitment required. Lectures are an efficient delivery system but risk information overload, increased knowledge decay, and reduced long-term memory formation (1). As children with medical complexity increase in prevalence and require specialized expertise to ensure their safe care, it is critical that effective teaching tools be identified that allow resident learners to improve their confidence, skill, and practice (2).

Memory performance and long-term retention benefit when items are presented repeatedly and spaced apart in time, which is known as the spacing effect (3–5). However, teaching content that is spaced requires an enormous time commitment by instructors, who must make themselves available multiple times to teach a concept. To this end, mobile phone technology, which has become ubiquitous both in daily life and during daily rounds, may provide an effective solution with which to access learners. Resident learners and medical students already frequently reach for their phone to answer questions during rounds, develop differential diagnoses, and identify optimal treatments and dosing of medications. Additionally, the current generation of learners has demonstrated decreasing interest in email and increasing interest in text messaging (6–8). Online spaced education meant to reinforce a live

course has resulted in improved self-reported global clinical behaviors (9). Our study design sought to objectively demonstrate whether discreet text messages could be an effective tool to teach novel content, in a spaced way, and result in improved knowledge retention.

Whether residents will use mobile technology to study is a complex question. Perspective pieces have encouraged development of text messaging as a teaching tool, arguing that “the prevalence, acceptance, and low cost of text messaging make it particularly inviting as a potentially high-yield learning tool in medical education.” (10) Feasibility studies have demonstrated that residents find text message-based educational materials helpful and educational as a daily study tool (11). By design, the text messages will space a larger concept into discreet episodes of learning and allow the learner to choose when to engage with the content.

The challenge for educators is to design learner-centered content that is scaffolded and delivered in a convenient and effective manner using this ubiquitous technology (12). Texting, by requiring brevity, forces the instructor to space the concept, thus reducing the risk of information overload. However, just presenting discrete facts to a learner will not result in improved memory formation; the learner must invest mental effort with a concept, in which the greater the degree of mental effort, the greater the degree of memory formation that results.

Author Contributions: H.K.A., as the principal investigator, conceptualized and designed the study, authored all the material used in the study, drafted the initial manuscript, revised it after feedback from the other authors, and approved the final manuscript as submitted. T.G.K., R.K., and T.P.C. reviewed and approved all the material, performed the initial analyses, reviewed and revised the manuscript, and approved the final manuscript as submitted.

Correspondence and requests for reprints should be addressed to Hovig K. Artinian, M.D., M.A.T., Pediatric Pulmonology and Sleep Medicine, Helen DeVos Children's Hospital, 35 Michigan Street NE, Suite 3003, Grand Rapids, MI 49503. E-mail: hovig.artinian@helendevoschildrens.org.

This is the basis for what is known as the testing effect, which is learning through testing rather than just restudying, and it has shown improved long-term memory formation (3). Using the testing effect in the form of multiple-choice questions with immediate feedback has been shown to result in improved performance on classroom examinations (13). Therefore, by writing texts in the form of test questions, ostensibly, the benefit of the testing effect combines with the benefit of the spacing effect. Also, because our design allowed for immediate feedback when a resident answered a text, there is personal gain or value for the resident in the form of teaching.

Given the existing access to mobile technology, the documented use of text messaging by residents, and the dearth of research into current delivery systems for spaced education and the testing effect, we sought to evaluate whether a complex medical topic could be effectively taught using testing methodology on a mobile device and result in increased knowledge. We hypothesized that resident learners would show high interest in learning this way, especially in the fast-paced inpatient environment, and that this would result in increased knowledge.

This work previously appeared in abstract form (14).

METHODS

Study Design and Setting

This was a single-center, prospective randomized controlled trial conducted at a pediatric residency program at a large, urban free-standing children's hospital during the 2016 to 2017 academic year. The institutional review board approved this study, and participants provided the requisite consent before enrollment.

Participants

We enrolled postgraduate year (PGY)-1, -2, and -3 trainees from a pediatric categorical training program rotating for the first time on the inpatient team serving general pediatric and pulmonary patients. Each team comprised three PGY-1 residents, led by one PGY-2 or PGY-3 resident. The pulmonary inpatient team, also termed Team 1 or 2, cares for hospitalized patients with cystic fibrosis, chronic lung disease, and chronic respiratory failure. The trainees spend the majority of the four-week rotation on day shift, alternating to night shift on a rotating basis.

Residents rotating through the inpatient team for a second time during the study period were excluded.

Study Planning Phase

Investigators conducted a six-month study planning phase using a modification of Kern and colleagues' six steps of curriculum development (15). After identifying the problem, a needs assessment was conducted of the residency to self-identify pulmonary topics of greatest interest. There were 29 respondents (30% response rate). From this needs assessment, chronic respiratory failure was chosen as the focus topic and 18 teaching objectives were developed using content experts and the American Board of Pediatrics content specifications (16). For each objective, three to four multiple-choice knowledge questions were originally authored and iteratively revised by a panel of five board-certified pediatric pulmonologists, several board-certified general pediatricians, and one board-certified pediatric emergency medicine physician. The questions were field-tested on 40 nonstudy residents to determine item discrimination and the

overall test reliability (Kuder-Richardson 20 [KR20] = 0.48). Items with the lowest item-total correlations were dropped or revised. Ultimately, 23 test items were randomly assigned for the pretest and 23 different test items were randomly assigned for the posttest, ensuring items from each objective were included in each test.

Randomization

Using a block randomization scheme, participants were randomized to either the text messaging arm (intervention) or the control arm in a 1:1 ratio (Figure 1). All rotation blocks were randomized using a random number generator (<http://www.random.org>) except for the December rotation owing to limitations with holiday scheduling of residents. All subjects within a single inpatient block were allocated to the same arm to prevent cross-contamination of data or sharing text

message interventions. Pulmonary attendings and fellows, other than the primary author (H.K.A.), were blinded to the allocation of the residents. The primary author (H.K.A.) was assigned to inpatient for 13% of the study period and documented all topics he taught to the residents during the study period.

Intervention

The intervention arm received one daily scripted multiple-choice text message scenario for three weeks (Monday through Saturday) at the resident's preferred time of day. Participants did not have input on any other aspects of the intervention. The 18 scripted text message scenarios and explanations were authored and iteratively revised by a panel of board-certified pediatric pulmonologists and general pediatricians. A scripted teaching text message was sent immediately after each

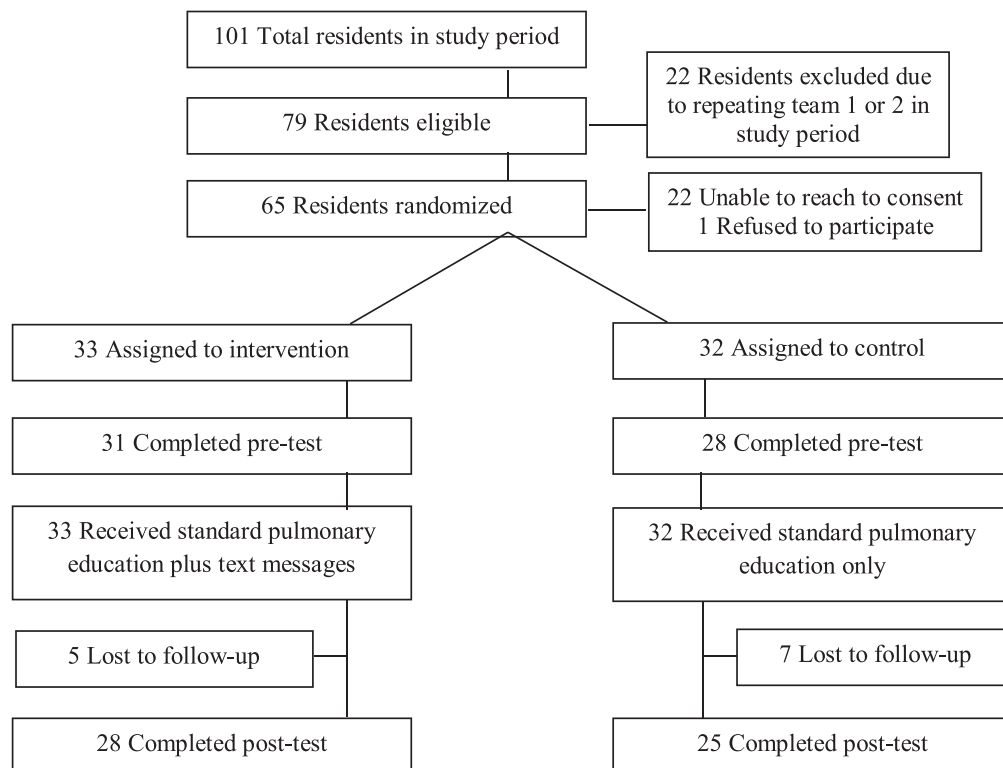


Figure 1. CONSORT (Consolidated Standards of Reporting Trials) Study flow diagram.

response. Participants could answer any text message at any time during the rotation. Participant responses were only counted if returned during their month of their inpatient rotation. All residents received the standard inpatient pulmonary education, which included one to two didactic sessions, daily family-centered rounds, and bedside teaching.

Protocol

To protect the privacy of participants, each participant was assigned an anonymous study identification number for all data collection. At the start of each block, participants completed a self-administered, 27-item questionnaire that assessed baseline knowledge ($n = 23$ items) and attitudes ($n = 4$ items) about chronic respiratory failure. Knowledge was assessed using multiple-choice questions and attitudes were assessed on a five-point Likert scale. All participants returned the pretest either before their inpatient block started or by the start of their first full week on inpatient service. At the end of each block, participants completed a self-administered 30-item questionnaire of knowledge ($n = 23$ different multiple-choice questions) and attitudes ($n = 7$ questions on a five-point Likert scale). Participants were encouraged to return the posttest within the block of their participation. Participants were incentivized to return the tests by receiving a \$25 Amazon gift card on completing both the pretest and posttest, regardless of study arm.

Primary Outcome Measure and Confounders

Improvement in knowledge about chronic respiratory failure was assessed by looking at the difference between the pretest and posttest scores, termed the delta score.

Maturation bias was tested for by using Spearman's rho correlation. Because each inpatient team is led by one PGY-2 or PGY-3, training year and its impact on scores was analyzed using a one-way analysis of variance with Bonferroni *post hoc* tests.

Analysis

Knowledge and confidence levels were compared between groups using independent and dependent samples *t* tests. Other secondary variables were analyzed using Spearman's rho correlation. Sample size was fixed based on the available resident population of approximately 60 within the eligible rotation. Preintervention sensitivity analysis revealed a large effect size (Cohen $d = 0.75$) to achieve 80% power at an α -level of 5%. All analyses were conducted using SPSS version 24 (IBM Corporation).

RESULTS

A total of 65 participants were randomized, with 32 residents in the control arm and 33 residents in the intervention arm, and 53 participants completed the study for a response rate of 81% (see Figure 1). The majority of participants were PGY-1 ($n = 37$; 56.9%), then PGY-2 ($n = 15$; 23.1%) and PGY-3 ($n = 13$; 20%) (Table 1). No maturation bias was observed ($P = 0.151$). Preintervention baseline knowledge increased linearly with year of training, and there was a statistically significant difference between the PGY-1 and PGY-3 group on the pretest mean score ($P = 0.01$); however, no statistically significant difference was observed at the end of the block in posttest score ($P = 0.07$) or delta score ($P = 0.13$) based on level of training. Among the intervention arm, the KR20 score for the text messages was calculated to be 0.58.

Table 1. Primary outcome results

	Enrolled (<i>n</i>)	Completed (<i>n</i>)	Pretest Percentage Correct		Posttest Percentage Correct		Delta Percentage Change	
			Mean	SD	Mean	SD	Mean	SD
Intervention								
PGY1	19	18	0.56	0.11	0.54	0.13	-0.02	0.12
PGY2	8	6	0.66	0.1	0.55	0.13	-0.14	0.16
PGY3	6	4	0.64	0.23	0.64	0.13	-0.11	0.15
Control								
PGY1	18	16	0.51	0.15	0.45	0.11	-0.06	0.12
PGY2	7	3	0.56	0.11	0.48	0.04	-0.03	0.1
PGY3	7	6	0.69	0.1	0.57	0.13	-0.15	0.08

Definition of abbreviations: PGY = postgraduate year; SD = standard deviation.

Primary Outcome Measure: Knowledge

There was not a statistically significant difference in delta mean score between control (-1.76) and intervention (-1.36) ($P=0.6$). Participants in the control arm had a lower pretest mean score (12.93) than those in the intervention arm (13.71); however, this was not statistically significant ($P=0.3$). There was a statistically significant difference in the posttest mean score between control (11.08) and intervention (12.75) ($P=0.04$). The lower posttest score in both groups was attributed to a lower overall reliability score on the posttest and more difficult test items randomized to the posttest (pretest KR20 during study period was 0.58; posttest KR20 during study period was 0.47). Table 1 summarizes the primary outcome results.

Secondary Outcomes

Among participants who completed the study, 83% agreed or strongly agreed that texting is an “effective” way to learn. This

feeling did not show a statistically significant difference on the posttest (79% agreed or strongly agreed) and was not different between the study arms. When asked whether participant felt “confident” managing chronic respiratory failure, no significant difference was observed between those agreeing or strongly agreeing on the pretest (18%) and posttest (32%) ($P=0.05$). Among all participants, 84% felt comfortable receiving one text message per day; none felt comfortable receiving four or more texts per day. Among those in the intervention arm, by the end of the block, 86% felt that one text per day “was just enough.” Median response rate to text messages was 14 out of 18; however, only six participants completed all 18 text messages. The majority of participants batched their responses to the text messages rather than responding daily. No correlation was observed between total texts returned (regardless of correct response) and delta score ($P=0.3$). No correlation was observed between total correct responses to the texts and delta score ($P=0.9$). No

correlation was observed between median response time to texts and delta score ($P=0.2$). Table 2 summarizes these secondary outcome results.

DISCUSSION

We investigated whether text messaging improves knowledge. Although participants perceived text messaging as effective and perceived an increase in knowledge, we did not detect an objective improvement in knowledge. These results are similar to those from a study on text messaging done in family medicine, with a similar number of participants but with a different timeline and context than our study, which focused on a discrete four-week period during inpatient rotation (17). To date, no other randomized controlled trials have evaluated texting within a similar conceptual framework in graduate medical education. However, our results are contrary to studies that have looked at other mobile technologies such as email, which have shown significant improvements in knowledge with quite significant effect sizes (6). The difference may be related to the medium used (computer vs. cellular phone) and the type of content each allows to be delivered. However, the difference may also be related to learner bias with the evaluation tool used in the

email study because an identical test was used for the pretest, posttest, and end-of-year test.

Our study did show that participants are highly motivated to learn via text messaging and they perceived it as effective. Although the completion of every text during the rotation was low ($n=6$), the median response rate by participants was robust at 14 out of 18 text questions.

This is in contrast to other studies that have shown a larger discrepancy between stated motivation and actual behavior for technology items (18). Perception can be deceiving and our study results suggest that a learner's perception of what will be effective should not necessarily drive content development decisions.

Our results demonstrated no maturation bias, consistent with previous findings that lectures are ineffective for long-term memory formation. However, spaced education and the testing effect are well-established educational methods, yet texting failed to improve knowledge. One explanation may be that previous studies have used technology to reinforce a previously taught topic but not to teach a novel topic. A drawback of our study is that no set curriculum was prescribed during rounds, therefore the texts that

Table 2. Secondary outcome results among intervention group

	Text Msgs Returned		Text Msgs Correct (%)		Response >24 h (%)		Median Response Time (h)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Intervention								
PGY1	14.3	4.5	0.47	0.18	0.41	0.31	46.2	68.6
PGY2	7.5	4.8	0.56	0.11	0.27	0.35	26	41.9
PGY3	14	6.9	0.63	0.05	0.44	0.24	16.4	19.3

Definition of abbreviations: Msgs = messages; PGY = postgraduate; SD = standard deviation.

participants received were not necessarily reinforcing a topic, rather they were teaching it anew. Future studies on texting in this conceptual framework might consider prescribing a set curriculum that is taught with standard methods, using the text messages to reinforce what is taught.

Another limitation of our study is that responses were not compelled daily. Most residents batched their responses to the texts, reducing opportunities for repetition. Additionally, the content in the text questions may have been too long, decreasing the convenience associated with brevity of texting, although reasonable response rates make this less likely; future studies should consider evaluating the correlation between text length and response rate and timing. Another potential limitation of our study was the use of pretest and posttests. Because of our fairly small, fixed sample size, the use of the pretest was important to quantify any differences in knowledge due to chance, despite the randomization (19). However, the pretest imposes an extra burden on the participants and can influence learning during the intervention (20).

Finally, the power of the mobile device may be altogether different than what we envisioned for this study. Learners are advised on rounds to read about the patients they see to improve retention through context. A wholesale reimaging of the mobile device to access a curated curriculum that reinforces learning on rounds, allows for exposure to patient experiences that may not be available at one's home institution, and allows for immediate learning gratification on any end-user chosen topic is where future study should focus. Lectures and text messages are not effective as standalone

interventions; however, a holistic educational approach that intertwines the strengths of these and other mediums in combination may prove successful and merits further study.

Conclusions

This study demonstrated that text messaging was perceived as effective but failed to improve knowledge. Future studies should ideally focus on texting as one tool within a suite of methods meant to improve context and knowledge. Additionally, consideration for how to incentivize daily engagement with the content may demonstrate a benefit in future studies. Finally, consideration should be given to whether there is a subset of learners who would potentially find greater benefit from text message-based learning.

Acknowledgment

The authors thank all the board-certified pediatric pulmonologists, pediatric hospitalists, and pediatricians who selflessly shared of their time in the development of this educational intervention. They also thank the many pediatric residents who volunteered during field-testing, both at University of California San Francisco-Fresno and at Children's Hospital Los Angeles, as well as those who participated during the study period. They also thank the Dr. Barbara M. Korsch Award for Research in Medical Education committee at Children's Hospital Los Angeles for pioneering medical education innovation.

Author disclosures are available with the text of this article at www.atsjournals.org.

REFERENCES

1. Paas F, van Gog T, Sweller J. Cognitive load theory: new conceptualizations, specifications, and integrated research perspectives. *J Educ Psychol Rev*. 2010;22:115–121.
2. Cohen E, Kuo DZ, Agrawal R, Berry JG, Bhagat SK, Simon TD, *et al*. Children with medical complexity: an emerging population for clinical and research initiatives. *Pediatrics* 2011;127:529–538.
3. Endres T, Renkl A. Mechanisms behind the testing effect: an empirical investigation of retrieval practice in meaningful learning. *Front Psychol* 2015;6:1054.
4. Logan JM, Castel AD, Haber S, Viehman EJ. Metacognition and the spacing effect: the role of repetition, feedback, and instruction on judgments of learning for massed and spaced rehearsal. *Metacognition Learn* 2012;7:175–195.
5. Carpenter SK, Cepeda NJ, Rohrer D, Kang SHK, Pashler H. Using spacing to enhance diverse forms of learning: review of recent research and implications for instruction. *Educ Psychol* 2012;24:369–378.
6. Kerfoot BP, DeWolf WC, Masser BA, Church PA, Federman DD. Spaced education improves the retention of clinical knowledge by medical students: a randomised controlled trial. *Med Educ* 2007;41:23–31.
7. Kerfoot BP, Armstrong EG, O’Sullivan PN. Interactive spaced-education to teach the physical examination: a randomized controlled trial. *J Gen Intern Med* 2008;23:973–978.
8. Lenhart, A, Ling, R, Campbell, S, Purcell, K. Teens and mobile phones: text messaging explodes as teens embrace it as the centerpiece of their communication strategies with friends. Pew Internet and American Life Project; 2010 [accessed 2017 May]. Available from: <https://files.eric.ed.gov/fulltext/ED525059.pdf>.
9. Shaw T, Long A, Chopra S, Kerfoot BP. Impact on clinical behavior of face-to-face continuing medical education blended with online spaced education: a randomized controlled trial. *J Contin Educ Health Prof* 2011;31:103–108.
10. Broom MA, Adamson GT, Draper LR. Text messaging in medical education. *Pediatrics* 2014;133:e491–e493.
11. Adamson GT, Draper LR, Broom MA. Text4Peds: feasibility of an educational text-messaging program for pediatrics residents. *J Grad Med Educ* 2014;6:746–749.
12. Herrington A, Herrington J. Authentic mobile learning in higher education. Presented at the AARE International Educational Research Conference. December 9–13, 2007, Melbourne, Australia.
13. Brosvic GM, Epstein M, Cook MJ, Dihoff RE. Efficacy of error for the correction of initially incorrect assumptions and of feedback for the affirmation of correct responding: learning in the classroom. *Psychol Rec* 2005;55:401–418.
14. Artinian H, Keens T, and Chang T. Augmenting resident pulmonary education related to management of chronic respiratory failure through implementation of a text messaging curriculum (research abstract) [abstract] [accessed 2020 Jun 24]. *Acad Pediatr* 2017;17:e20. Available from: [https://www.academicpedsjnl.net/article/S1876-2859\(17\)30229-2/abstract](https://www.academicpedsjnl.net/article/S1876-2859(17)30229-2/abstract).
15. Kern D, Thomas PA, Hughes MT. Curriculum development for medical education: a six-step approach, 2nd ed. Baltimore, MD: Johns Hopkins University Press; 2009.
16. American Board of Pediatrics. Content outline for general pediatrics maintenance of certification examination [updated 2017; accessed 2015 Feb]. Available from: https://www.abp.org/sites/abp/files/gp_contentoutline_2017.pdf.

17. Mount HR, Zakrajsek T, Huffman M, Deffenbacher B, Gallagher K, Skinker B, *et al.* Text messaging to improve resident knowledge: a randomized controlled trial. *Fam Med* 2015;47:37–42.
18. Chang TP, Pham PK, Sobolewski B, Doughty CB, Jamal N, Kwan KY, *et al.* Pediatric emergency medicine asynchronous e-learning: a multicenter randomized controlled Solomon four-group study. *Acad Emerg Med* 2014;21:912–919.
19. Cook, DA, Beckman, TJ. Reflections on experimental research in medical education. *Adv Health Sci Educ Theory Pract* 2010;15:455–464.
20. Kim E, Willson VL. Evaluating pretest effects in pre-post studies. *Educ Psychol Meas* 2010;70:744–759.