Improving physician clinical documentation quality: Evaluating two self-efficacy-based training programs

Ruthann Russo  
Stephen P. Fitzgerald  
JD Eveland  
Barry D. Fuchs  
D. Patrick Redmon

Background: Clinical documentation is critical to health care quality and cost. The generally poor quality of such documentation has been well recognized, yet medical students, residents, and physicians receive little or no training in it. When clinical documentation quality (CDQ) training for residents and/or physicians is provided, it excludes key constructs of self-efficacy: vicarious learning (e.g., peer demonstration) and mastery (i.e., practice). CDQ training that incorporates these key self-efficacy constructs is more resource intensive. If such training could be shown to be more effective at enhancing clinician performance, it would support the investment of the additional resources required by health care systems and residency training programs.

Purposes: The aim of this study was to test the impact of CDQ training on clinician self-efficacy and performance and the relative efficacy of intervention designs employing two versus all four self-efficacy constructs.

Methodology/Approach: Ninety-one internal medicine residents at a major academic medical center in the northeastern United States were assigned to one of two self-efficacy-based training groups or a control group, with CDQ and clinical documentation self-efficacy measured before and after the interventions. A structural equation model (AMOS) allowed for testing the six hypotheses in the context of the whole study, and findings were cross-validated using traditional regression.

Key words: clinical documentation, quality, self-efficacy

Ruthann Russo, PhD, MPH, is Assistant Professor, Health Information Technology, City University of New York/Borough of Manhattan Community College. E-mail: ruthannrusso@gmail.com.
Stephen P. Fitzgerald, PhD, is Director of Faculty Development, Associate Professor of Business Administration and Lead Faculty in Management, Trident University International, College of Business Administration, Cypress, California. E-mail: sfitzgerald@tuiu.edu.
JD Eveland, PhD, is Professor of Business Administration and Information Systems, Trident University International, Colleges of Business Administration and Information Systems, Cypress, California. E-mail: jeveland@tuiu.edu.
Barry D. Fuchs, MD, FCCP, is Associate Professor of Medicine, School of Medicine, University of Pennsylvania, and Medical Director, Medical Intensive Care Unit and Respiratory Care Services, Hospital of the University of Pennsylvania, Philadelphia. E-mail: barry.fuchs@uphs.upenn.edu.
D. Patrick Redmon, PhD, is Associate Professor, Graduate Program in Health Services Administration, Xavier University, Cincinnati, Ohio. E-mail: redmond@xavier.edu.

The authors have disclosed that they have no significant relationship with, or financial interest in, any commercial companies pertaining to this article.

DOI: 10.1097/HMR.0b013e31824c4c61

Copyright © 2013 Wolters Kluwer Health | Lippincott Williams & Wilkins
Accurate and complete clinical documentation in inpatient medical records serves as a communication tool for all health care personnel involved in the care of a patient and is essential for medical and legal purposes (Eichholz, Van Voorhis, Sorosky, Smith, & Sood, 2004; Rangachari, 2008; Siegler, 2010). As a result, health care systems, hospitals, residency training programs, medical schools, and patients all have an interest in ensuring physicians are practicing high-quality clinical documentation. The importance of clinical documentation quality (CDQ) and its generally poor quality have been well recognized (e.g., Cascio, Wilkens, Ain, Toulson, & Frassica, 2005; Eichholz et al., 2004; Larson, Wiggins, & Goldfarb, 2004; Novitsky, Singh, Kersher, & Griffo, 2005; Temel et al., 2009). For example, Temel et al. found that 80% of terminal cancer patients did not have code status documented in their electronic medical records. Cascio et al. found inadequate documentation in 70% of patient records and discrepancies in 62% of physicians’ notes with respect to weight, vascular lines, or medications.

According to the Centers for Disease Control (CDC) and Centers for Medicare and Medicaid Services (CMS), high-quality clinical documentation is documentation that is timely, legible, complete, clear, unambiguous, and precise (CDC, 2011; CMS, 2003). In health care, the physician’s documentation is an essential input, and accurate health information is a critical output (AHIMA Coding Products and Services Team, 2003). In the business of health care, budgeting and spending are based on documentation provided by physicians for the care they provide to patients (CMS, 2003). Moreover, poor quality documentation in a patient’s record has been linked to both excessive health care costs and poor quality of care (Institute of Medicine, 1999; Pohl, Reis, & Forjuoh, 2010; Simmons & Goldberg, 2003).

Recent legislation and administrative changes require health care providers to implement electronic health records and utilize a disease severity-based model for CMS billing (Ballentine, 2009; Forthman, Gold, Dove, & Henderson, 2010; Hripcsak, Vawdrey, Fred, & Bostwick, 2011; Mahon, Nickitas, & Nokes, 2010; Rangachari, 2008). Clinical documentation creates the groundwork for the electronic health record (Hripcsak et al., 2011), and severity-based coding and billing require physicians to document all clinically relevant diagnoses (Ballentine, 2009; Rangachari, 2008). Health care providers are compelled to improve clinical documentation practices to comply with these recent requirements.

Findings: Although both interventions increased CDQ, the training designed to include all four self-efficacy constructs had a significantly greater impact on improving CDQ. It also increased self-efficacy.

Practice Implications: CDQ may be significantly improved and sustained by (a) training physicians in clinical documentation and (b) employing all four self-efficacy constructs in such training designs.

Theoretical Framework

Peer demonstration, observation, and practice are integral to promoting self-efficacy, which has been shown to improve a variety of behavioral outcomes in clinical settings (e.g., Cabana, Rand, Slish, & Nan, 2004; Carson, Gillham, Kirk, Reddy, & Battles, 2002; Laschinger & Tresolini, 1999; Mavis, 2001; Opacic, 2003). Self-efficacy influences behavior through four primary mechanisms: mastery, vicarious experience, social persuasion, and psychological states (Bandura, 1982). Mastery entails successfully performing the focal task, vicarious experience entails observing and studying credible others demonstrating it, and social persuasion involves credible, trustworthy experts assuring trainees that they are capable of performing it (Laschinger & Tresolini, 1999). Psychological and emotional states (“psychological states”) involve enhancing individuals’ physical status and reducing stress levels (Cioffi, 1991).

Given the critical nature of CDQ, it is surprising that medical students receive little or no CDQ training and that residents and physicians receive such training inconsistently at best. When CDQ training is provided, the sponsoring organization is usually a hospital, the training is generally less than an hour, and peer demonstration and observation and/or practice opportunities are rarely, if ever, provided (Advisory Board Company, 2006; Hicks & Gentleman, 2003), thereby excluding key constructs of self-efficacy that are integral to most self-efficacy research and intervention designs: mastery and vicarious experiences.

The relative efficacy of training designs employing two versus all four constructs of self-efficacy has not been tested, nor has the impact of CDQ training on clinician self-efficacy and performance. Furthermore, research on developing physician self-efficacy has focused primarily on the constructs of mastery and/or vicarious experience (e.g., Cabana et al., 2004; Carson et al., 2002; Coffman, Shellman, & Berna, 2004; Mavis, 2001; Opacic, 2003). Because social persuasion and psychological state interventions reflect the streamlined methodology used to train physicians in CDQ by health care providers (Department of Health and Human Services, Office of the Inspector General, 1999a, 1999b, 1999c, 2005), establishing
the relative efficacy of such training is essential. These findings led to the following hypotheses.

H1a: Four-construct training is positively related to CDQ.

H1b: Two-construct training is positively related to CDQ.

H2a: Four-construct training is positively related to self-efficacy.

H2b: Two-construct training is positively related to self-efficacy.

H3a: Four-construct training positively impacts CDQ more than two-construct training does.

H3b: Four-construct training positively impacts self-efficacy more than two-construct training does.

Methods

Sample

Resident physicians in the internal medicine residency program at a 600-bed major academic medical center in the northeastern United States were chosen as the target population, because residents provide a significant amount of documentation in patient records, are likely to attend all of the educational sessions when asked by their program director, and are the attending physicians of the future; therefore, a successful intervention in resident training may have a long-term impact.

Instruments

Participants in each group completed three preintervention questionnaires: a CDQ proficiency test, a self-efficacy assessment, and a demographic questionnaire. The CDQ test measured residents’ ability to document cases consistent with the criteria for high-quality clinical documentation (CMS, 2003), with content that is timely, legible, complete, clear, unambiguous, and precise (CDC, 2005; CMS, 2003). The 27-item CDQ test was composed of six multiple-choice questions regarding basic knowledge of clinical documentation requirements and regulations and 11 case study questions requiring written documentation of patient diagnoses based on clinical information provided. Each question received a score of 0–5 based on completeness and accuracy. For example, the correct response to one question was “acute blood loss anemia”; this was scored as follows: acute = 1; blood loss = 2; anemia = 1; legibility = 1. Participants received one overall test score ranging from 0 to 100% and six subscores based on the components of clinical documentation tested by each question.

The 22 items on the clinical documentation self-efficacy instrument were adapted from previously validated instruments by pairing standard self-efficacy lead-ins (e.g., “I am confident that I can…” with specific clinical documentation criteria and practices that have been shown to be both problematic and important for ensuring high quality and cost considerations in patient care (AHIMA Coding Products and Services Team, 2003; CDC, 2005; CMS, 2003). These clinical documentation practices were also included as objectives in the training provided to the physicians (Cabana, et al., 2004; Katz, Feigenbaum, Pasternak, & Vinker, 2005); the purpose of the pretest was to control for prior knowledge going into the training that might affect posttraining scores.

Content validity of both the CDQ and self-efficacy measures was established by a panel of four physicians, a nurse, and a clinical documentation expert, each of whom independently reviewed and provided input on the two instruments. Both instruments were then piloted with a group of 22 physicians. An exploratory factor analysis on the self-efficacy measure suggested a single-factor structure as expected ($\alpha = 0.87, m = 0.73$); therefore, all items were retained.

Procedures

The study participants were assigned to one of three groups: (a) four-construct training, (b) two-construct training, or (c) control (no training). The physician instructor, with 5 years of experience in clinical documentation improvement consulting and training, was paired with the lead investigator, a certified health information professional with over 15 years of experience in designing clinical documentation improvement programs. The training programs were preplanned and scripted for content consistent with either two or four constructs of self-efficacy. The physician was coached prior to the educational intervention sessions on the self-efficacy constructs and on the importance of delivering the training script as prepared.

The same training objectives were used for both interventions: (a) demonstrate understanding of the relationship between physician documentation and its translation into ICD-9-CM coded data; (b) demonstrate an understanding that ICD-9-CM coded data are used for planning, reimbursement, quality ratings, Medicare Conditions of Participation, JCAHO Core Measures, and research; (c) provide documentation in the inpatient record that is timely, legible, complete, clear, unambiguous, and precise; and to document (d) detail and precision in the patient’s principle diagnosis, (e) all chronic coexisting secondary
diagnoses, (f) all acute coexisting secondary diagnoses, (g) the clinical significance of all abnormal diagnostic tests, and (h) the etiology or suspected etiology of symptoms.

Identical instructional materials were distributed to all study participants. The primary difference between the four-construct training program and the two-construct training program was the format for its delivery. For example, in the four-construct training, the internal medicine physician shared examples from his personal experience that showed how he became a good documenter (vicarious learning). He provided feedback to the trainees as they practiced and applied CDQ during the training (mastery). The two-construct group did not receive mastery training or hear about physician experiences in clinical documentation. All trainees were assured by both instructors that they would be able to effectively apply the skills being trained (social persuasion) and were given a free dinner (psychological state interventions). The contents of the program for each intervention can be found in Table 1.

### Statistical Analyses

The original data analysis plan called for hypothesis testing using a series of specific regression analyses. However, all variables, including the controls, were fairly highly intercorrelated. Therefore, we estimated two structural equation (latent variable) models that subsumed all of the hypothesized relationships together and allowed for the testing of each hypothesis in the context of all the others, implemented using AMOS v.18 (Arbuckle, 2008). This procedure simplified both the analysis and the presentation of findings. For confirmation, the individual hypotheses were also tested using traditional regression methods. The substantive conclusions were identical in both the

### Table 1

<table>
<thead>
<tr>
<th>Construct</th>
<th>Operationalized by</th>
</tr>
</thead>
</table>
| Vicarious experiences (included in four-construct training only) | 1. A videotape of physicians discussing documentation issues and solutions to those issues was shown.  
2. Examples of good documentation by physicians in patient records were shared.  
3. Coparticipants in the study provided an example of how they would document accurately and correctly.  
4. The physician conducting the session provided examples of how he learned to document accurately and correctly.  
5. Examples of how the information learned during the first session was applied by members of the study group were shared during the second of the two-part training session. Every member of the training group was asked to speak and share their documentation experiences over the prior week since the first training session. |
| Mastery experiences (included in four-construct training only) | 1. Power presentations were reviewed with residents, and they were given a copy of the presentation to take for review and reference purposes.  
2. Each participant had the opportunity to accurately and correctly document diagnoses in the five problem areas from 10 sample medical records.  
3. Tools were provided to each participant to use while documenting. These include a handbook, and cue cards are provided for reference. |
| Social persuasion (included in both two-construct and four-construct training) | 1. The instructor pointed out that the information that needs to be documented is “in the physician’s head”; they just needed to learn to translate it to the written word.  
2. The importance of complete and accurate documentation to the patient was discussed.  
3. The job of the “translators” of the physicians’ documentation and how they needed the physician’s complete documentation to do their job were discussed.  
4. Feedback was provided during each session based on how well the participants did in their pretest and their participation in class exercises. |
| Psychological and emotional states (included in both two-construct and four-construct training) | 1. Dinner was provided at each session.  
2. The frustrations physicians have with complying with increasing regulations in health care were discussed.  
3. The monetary impact of incomplete or inaccurate documentation was discussed.  
4. The quality of care impact of incomplete or inaccurate documentation was discussed.  
5. Time management issues were discussed. |
regression and structural equation model (SEM) procedures. The SEM employs a maximum likelihood estimation algorithm rather than the least squares algorithm used in regression analysis, and thus the precise values of the coefficients differ between the two procedures; however, with a stable model, such as that tested here, both the relative magnitudes (effect sizes) and significance of the respective coefficients are generally closely aligned between the SEM and regression procedures.

Demographic data are shown in Table 2, and descriptive statistics are shown in Table 3.

Figures 1 and 2 show the two structural models that were estimated. They were conceptually identical and differed only in the treatment of the CDQ variables. In Model 1, the simpler version, the two CDQ variables were represented by single-item indices derived from the 17 separate CDQ measurements taken at each time point.

In Model 2, the CDQ variables were represented as latent variables with the separate measurements interpreted as reflective indicators of the latent variables. The latter structure perhaps strains the interpretation of “reflective indicators” to some degree, but it does provide some insights into the issue of overall model fit, as noted below.

In both models, the two dependent variables are the posttraining test scores for self-efficacy and CDQ. The independent variables are the training conditions (two-group and four-group) operationalized as two dummy variables with the control (no training) condition as the reference category; the effects for training condition thus reflect the increment for that group above the no-training condition effect. In both models, the self-efficacy measures (pre and post) are expressed as latent variables with four reflective indicators. These are derived from the four primary multi-item sections of the clinical documentation self-efficacy instrument related to confidence in one’s ability to document: (a) patient history and diagnoses, (b) etiology based on presenting symptoms, and (c) clinical significance of abnormal tests, (d) all in a way that will meet hospital coding and Medicare requirements and the Joint Commission guidelines.

The structure of both models with regard to dependent variables (self-efficacy and CDQ), the independent variable (training condition), and control variables was identical. Pretest self-efficacy was included as a control variable in the prediction of posttraining self-efficacy, whereas pretest CDQ competence was included as a control for posttraining CDQ expertise. Two demographic variables were included as possible controls for the dependent variables: gender (operationalized as a dummy variable representing the effect of being male) and year of residency training (operationalized as separate dummy variables for second-year status and third-year status, with first-year status as the reference category).

Statistically significant ($p < .05$) paths are shown in both models as thicker black lines; paths tested but found to be nonsignificant are shown as thinner lines. The numbers in larger font size next to the black lines are the path (standardized regression) coefficients, representing the effect of one variable on another. The $R^2$ statistics (directly beneath the two posttest ovals) represent the proportion of variance in the two dependent variables accounted for by the combination of independent variables and controls (79%–80%).

The substantive findings from both models are essentially identical. Both models show large and significant positive effects for the pretests on the posttests, along with significant positive effects of both treatments on posttest CDQ (with the four group treatment effect being half again as large as the two group treatment effect); the four group treatment also affects posttest self-efficacy positively, although the two group treatment does not. In addition, being male has a small but significant positive effect on

### Table 2

<table>
<thead>
<tr>
<th>Treatment group</th>
<th>Gender</th>
<th>Program year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>Male</td>
</tr>
<tr>
<td>Frequency</td>
<td>31</td>
<td>30</td>
</tr>
<tr>
<td>Percent</td>
<td>34.1</td>
<td>33.0</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th></th>
<th>Pretest self-efficacy</th>
<th>Pretest CDQ</th>
<th>Posttest self-efficacy</th>
<th>Posttest CDQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
</tr>
<tr>
<td>Mean</td>
<td>2.74</td>
<td>0.67</td>
<td>2.83</td>
<td>0.83</td>
</tr>
<tr>
<td>$SD$</td>
<td>0.42</td>
<td>0.13</td>
<td>0.44</td>
<td>0.15</td>
</tr>
<tr>
<td>Minimum</td>
<td>1.77</td>
<td>0.34</td>
<td>1.78</td>
<td>0.33</td>
</tr>
<tr>
<td>Maximum</td>
<td>3.82</td>
<td>0.91</td>
<td>3.86</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Note. CDQ = clinical documentation quality.
posttest self-efficacy, whereas a second-year residency status has a similar effect on posttest CDQ. Table 4 summarizes the estimated coefficients and connects them to the hypotheses.

To summarize the results of the tests of the six hypotheses, see the following:

H1a: Four-construct training is positively related to CDQ

Supported: $b = .69$, $p < .05$

H1b: Two-construct training is positively related to CDQ

Supported: $b = .44$, $p < .05$

H2a: Four-construct training is positively related to self-efficacy

Supported: $b = .33$, $p < .05$

H2b: Two-construct training is positively related to self-efficacy

Not supported: $b = .14$, $p = ns$

H3a: Four-construct training positively impacts CDQ more than two-construct training

Supported: Critical ratio $= -3.75$, $p < .05$

H3b: Four-construct training positively impacts self-efficacy more than two-construct training

Not supported: Critical ratio $= 1.79$, $p = ns$
these effects into the models allows for direct interpretation of main effect coefficients without further qualification.

In both models, the latent variables (pretest and posttest self-efficacy and, in Model 2, pretest and posttest CDQ competence) appear to be reasonable combinations of their respective indicators, with path coefficients (interpretable in this case as essentially equal to factor loadings) ranging from about .5 to .7, thus supporting the argument for construct validity of both latent variables.

The issue of overall “fit” of the model to the data is always perplexing in any kind of SEM, and there is little overall consensus on either the value of specific indices or their interpretation. With dozens of possible indices to choose from, the very idea of “fit” in any absolute sense is questionable. Most indices are much more useful in assessing comparative fit of alternative models relative to the same data than in providing any absolute measure of fit comparable across studies.

Table 4 includes some fit measures commonly reported. The most commonly reported measure, the chi-square statistic and associated degrees of freedom, is significant in both models; however, this is common in models with more than about 75 cases and not necessarily an indicator of bad fit. Chi-square statistics are customarily adjusted for degrees of freedom to account for this, resulting in the commonly reported chi-square/degrees of freedom index. The value for both models (2.06, 1.53) is generally considered to be acceptable; “chi-square to degrees of freedom ratios in the range of 2 to 1 or 3 to 1 are indicative of an acceptable fit between the hypothetical model and the

---

Figure 2

Estimated structural equation model 1: CDQ variables represented as latent variables

---

CDQ = Clinical documentation quality

**Coefficients in large type indicate significance, p < .01; values omitted for indicators — all are significant**

***Coefficients in smaller type indicate non-significant values; coefficient values omitted for error terms**
sample data” (Carmines & McIver, 1981, p. 80). The most useful application of this statistic is to test the incremental effect of one model over another; in this case, Model 2 shows a modest improvement over Model 1, indicating a slightly better fit (chi-square difference = 1399.4, df = 942, p < .05).

The most commonly reported measure of overall fit of the model to the data is the root mean square error of approximation (RMSEA), but even here there are differences of opinion about interpretation. Browne and Cudeck (1993) noted that:

Practical experience has made us feel that a value of the RMSEA of about 0.05 or less would indicate a close fit of the model in relation to the degrees of freedom. This figure is based on subjective judgment. It cannot be regarded as infallible or correct, but it is more reasonable than the requirement of exact fit with the RMSEA = 0. We are also of the opinion that a value of about 0.08 or less for the RMSEA would indicate a reasonable error of approximation and would not want to employ a model with an RMSEA greater than 0.1.

The achieved RMSEAs are 0.11 for Model 1 and 0.07 for Model 2. Chen, Curran, Bollen, Kirby, and Paxton (2008), in their assessment of RMSEA cutoff criteria,

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Criterion</th>
<th>B</th>
<th>Standard Error</th>
<th>Beta</th>
<th>Model 2a Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Four-group</td>
<td>Post-CDQ</td>
<td>0.26*</td>
<td>0.02</td>
<td>0.70*</td>
<td>0.73*</td>
</tr>
<tr>
<td>Two-group</td>
<td>Post-CDQ</td>
<td>0.16*</td>
<td>0.02</td>
<td>0.44*</td>
<td>0.53*</td>
</tr>
<tr>
<td>3rd year residents</td>
<td>Post-CDQ</td>
<td>0.02</td>
<td>0.02</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>2nd year residents</td>
<td>Post-CDQ</td>
<td>0.04*</td>
<td>0.02</td>
<td>0.11*</td>
<td>0.11*</td>
</tr>
<tr>
<td>Male</td>
<td>Post-CDQ</td>
<td>−0.04</td>
<td>0.02</td>
<td>−0.09</td>
<td>−0.07</td>
</tr>
<tr>
<td>Post-SE</td>
<td>Post-CDQ</td>
<td>−0.01</td>
<td>0.03</td>
<td>−0.01</td>
<td>−0.03</td>
</tr>
<tr>
<td>Pretest CDQ</td>
<td>Post-CDQ</td>
<td>0.44</td>
<td>0.07</td>
<td>0.32*</td>
<td>0.32*</td>
</tr>
<tr>
<td>Four-Group</td>
<td>Post-SE</td>
<td>0.21</td>
<td>0.05</td>
<td>0.33*</td>
<td>0.33*</td>
</tr>
<tr>
<td>Two-Group</td>
<td>Post-SE</td>
<td>0.09</td>
<td>0.05</td>
<td>0.14*</td>
<td>0.14</td>
</tr>
<tr>
<td>3rd year residents</td>
<td>Post-SE</td>
<td>0.02</td>
<td>0.05</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>2nd year residents</td>
<td>Post-SE</td>
<td>0.09</td>
<td>0.05</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>Male</td>
<td>Post-SE</td>
<td>0.20</td>
<td>0.05</td>
<td>0.29*</td>
<td>0.29*</td>
</tr>
<tr>
<td>Pre-SE</td>
<td>Post-SE</td>
<td>0.94</td>
<td>0.16</td>
<td>0.75*</td>
<td>0.75*</td>
</tr>
<tr>
<td>Pre-SE</td>
<td>SE2overall</td>
<td>1.00</td>
<td></td>
<td>0.74*</td>
<td>0.74*</td>
</tr>
<tr>
<td>Pre-SE</td>
<td>SE3overall</td>
<td>1.21</td>
<td>0.23</td>
<td>0.62*</td>
<td>0.62*</td>
</tr>
<tr>
<td>Pre-SE</td>
<td>SE4overall</td>
<td>1.15</td>
<td>0.19</td>
<td>0.72*</td>
<td>0.72*</td>
</tr>
<tr>
<td>Pre-SE</td>
<td>SE5overall</td>
<td>1.75</td>
<td>0.38</td>
<td>0.53*</td>
<td>0.53*</td>
</tr>
<tr>
<td>Post-SE</td>
<td>PostSE2overall</td>
<td>1.00</td>
<td></td>
<td>0.79*</td>
<td>0.79*</td>
</tr>
<tr>
<td>Post-SE</td>
<td>PostSE3overall</td>
<td>1.20</td>
<td>0.18</td>
<td>0.69*</td>
<td>0.69*</td>
</tr>
<tr>
<td>Post-SE</td>
<td>PostSE4overall</td>
<td>1.41</td>
<td>0.15</td>
<td>0.91*</td>
<td>0.91*</td>
</tr>
<tr>
<td>Post-SE</td>
<td>PostSE5overall</td>
<td>1.59</td>
<td>0.26</td>
<td>0.63*</td>
<td>0.63*</td>
</tr>
</tbody>
</table>

Note. CDQ = clinical documentation quality; SE = self-efficacy.

<table>
<thead>
<tr>
<th>Model Fit</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square/df</td>
<td>173.2/84</td>
<td>1572.6/1026</td>
</tr>
<tr>
<td>Model 2 improvement</td>
<td>1399.4/942*</td>
<td></td>
</tr>
<tr>
<td>RMSEA</td>
<td>.11</td>
<td>.07</td>
</tr>
<tr>
<td>CFI</td>
<td>.81</td>
<td>.58</td>
</tr>
</tbody>
</table>

*p < .05.

*Coefficients for indicators of CDQ latent variables in model 2 are all significant; not listed individually.

aH1a: Supported.
bH1b: Supported.
cH2a: Supported.
dH2b: Not supported.
found no empirical support for the use of any universal cutoff point to determine adequate model fit and urged researchers to use human judgment in combination with fit indices to assess the model fit. Similarly, Iacobucci (2010) encouraged researchers to not take the rules of thumb regarding model fit index cutoff points too seriously, but rather to “ask good theoretical questions: Is every hypothesized link logically supported, and is there a sound, comprehensive yetparsimonious theoretical story for the entire model?” (p. 95).

In this study, in summary, the hypothesized models are certainly theoretically sound as well as generally consistent with the data. The effects are clear-cut and consistent across both models, as well as corroborated via traditional regression analysis. The overall fit indices, particularly for the more complex Model 2, are not outside the generally acceptable limits that might cause one to doubt the applicability of a model. Conceptually, then, the structural models provide considerable support for the basic contention of the study—that is, that training incorporating the four groups of factors provides significantly better results than training without these elements. In particular, the four-construct training has demonstrably greater impact on CDQ, as expected; the stronger effects for CDQ competence than for self-efficacy are not unpredictable given the greater tangibility and probably lower degree of measurement error associated with the competence construct.

One final point should be noted. Although the use of AMOS-based SEM is quite appropriate here, it could be argued that the data really more clearly match the expectations for an alternative modeling technique, partial least squares (PLS), in terms of sample size, mixture of formative and reflective indicators, and two-stage latent variables. In fact, we conducted a PLS analysis of the data. All the substantive conclusions reported here were also confirmed by the PLS analysis; however, given the very different estimation methods involved, the coefficients were somewhat different from those generated by AMOS, and we thought it better not to report them as such in the interests of avoiding confusion arising from too many differently interpretable numbers. We do plan to report the results of this comparative application of different analytical techniques in a subsequent article.

## Theory Implications

The current research contributes to theory by isolating for the first time the impact of social persuasion and psychological states from that of mastery and vicarious experience. Bandura (1982) showed that mastery contributed the most among the four constructs to developing self-efficacy. However, we were not able to identify any studies that had isolated the relative impact of social persuasion and psychological state constructs in relation to mastery and vicarious experience. The results show that although training that employs only the social persuasion and psychological state components of self-efficacy can positively impact key clinical documentation skills training outcomes, no support was found for a similarly positive impact on self-efficacy. Only the full four-construct training has a positive impact on post-training self-efficacy.

Furthermore, although the impact of self-efficacy improvement on clinical documentation outcomes has been shown in prior research, the specific impact on CDQ had not been tested. Therefore, the current study extends prior self-efficacy research to the CDQ framework for the first time.

## Practice Implications

Training designed to employ all four components of self-efficacy showed substantially greater positive impact on improved CDQ (H1) and positively impacted self-efficacy (H2). These results were robust across two different analyses (SEM and regression), even given the modest sample size. This provides considerable support for the greater resource investment that training that incorporates mastery and vicarious experience requires. Clinical documentation impacts quality of care, clinical coding, and health care costs (AHIMA Coding Products and Services Team, 2003; Institute of Medicine, 1999; Pohl et al., 2010; Simmons & Goldberg, 2003). In addition, the United States is slated to begin using the current version of diagnostic and procedural coding (ICD-10) for all health care reporting on October 1, 2013. This system is even more dependent on the detail of physician clinical documentation than the current coding system (AHIMA Coding Products and Services Team, 2003).

Given the benefits of four construct training, health care systems should adopt the four-construct approach to CDQ training employed in this research. In fact, health care organizations that understand the value of high-quality clinical documentation have begun to train their physicians and clinicians doing just that. The academic medical center that served as the test site for this research has since committed to training additional physicians on their staff using the four-construct training. Since the study was published, 12 academic medical centers, 19 acute care community hospitals, and 1 health care system comprising 27 acute care hospitals have trained their medical staffs using the four-construct methodology. Although it is still too soon to determine the long-term impact of the training, the organizations have reported more reliable coding as the result of training. In addition, over 1,800 health information professionals received training on the four-construct method of clinical documentation training. During those sessions, most attendees reported that their organizations intended to invest in further training.
The greatest benefit of the four-construct clinical documentation training will result when the method is taught during medical school. Training in clinical documentation concepts at that time will allow medical students to make errors and receive feedback prior to documenting in actual patient records, where errors could impact the quality and cost of patient care.

The study design made use of pretest and posttest instruments to determine improvement in quality. This methodology was chosen for its efficiency and ability to produce consistent results. However, the true test of CDQ is in the actual documentation practices. This study sets the stage for future research that employs actual documentation practices.

The sample was drawn from a 600-bed academic medical center, and there are 123 such centers in the United States today. However, there are over 2,000 hospitals throughout the country that employ residents. These hospitals are not part of academic medical centers like the study site. Although it is not possible to generalize the results, it is more likely that sponsorship by a large academic medical center will produce similar results to those found in this research.

Acknowledgments

The institutional review boards of Trident University International and the University of Pennsylvania both granted exempt approval for this research. A prior version of this article was nominated by the Health Care Management Division of the Academy of Management for the division’s “Best Theory to Practice Paper Award” and presented in that division’s paper track at the Academy’s annual meeting in Anaheim, California, in August 2008. Dr. Joshua Shackman and Dr. Mickey Shachar of Trident University International both provided helpful reviews of the manuscript prior to submission.

References


