Computer Proficiency for Online Learning: Factorial Invariance of Scores among Teachers

Amy L. Martin
Todd D. Reeves
Thomas J. Smith
David A. Walker
Northern Illinois University

Online learning is variously employed in K-12 education, including for teacher professional development. However, the use of computer-based technologies for learning purposes assumes learner computer proficiency, making this construct an important domain of procedural knowledge in formal and informal online learning contexts. Addressing this concern, this study examined the score properties and invariance (N=11,709) of an eight-item self-report measure of computer proficiency for online learning, the CPOL. Results from a confirmatory factor analysis suggest that the hypothesized unidimensional structure undergirded the instrument's scores, and invariance analyses suggested that the instrument functions similarly across teacher populations defined by gender, grade level taught, and age, and over time. Specifically, the results showed "strict" score invariance for all teacher groupings except for age. Discussed are potential proximal and distal applications of, and directions for future research concerning the CPOL.

Online learning continues to be advanced as a solution to educational and training problems. For example, many public and private institutions of higher education have pursued online programming as a strategic decision to compete for students in a highly-saturated market—for geographically-distant and non-traditional students in particular (Allen & Seaman, 2013). In the industry sector too, increasingly companies are employing online learning for staff development purposes (Gunawardena, Linder-VanBerschot, LaPointe, & Rao, 2010). Nascent applications of online learning within the education realm include formal (e.g., via university-based online courses, district-arranged online professional development) and informal (e.g., via Twitter, blogs, and online publications) professional development of teachers. Options for online learning help mitigate scheduling and geographic barriers to learning, and afford learners more flexibility and choice (Carpenter & Krutka, 2014; Reeves & Pedulla, 2011).

Given online learning and teaching's use of computer-based and Internet technologies as a vehicle by which to promote learning, the computer proficiency of online learners—be it employees, students, or teachers—is of key interest for those who deliver online learning (Tallent-Runnels et al., 2006; Zhao, Lei, Yan, Lai, & Tan, 2005). Indeed, insufficient computer skills of some learner sub-populations (e.g., members of older generations) can represent a barrier to success within online learning contexts (Anderson, 2008; Reeves & Pedulla, 2013). As such, computer proficiency is an increasingly important domain of procedural knowledge in formal and informal online learning contexts.

Defining Computer Proficiency

Computer proficiency is understood as one's ability to use computer-based technologies broadly, including the ability to conduct various operations (e.g., conducting an Internet search, sending an e-mail) and use various applications (e.g., word processors; Campbell & Williams, 1990; Grant, Malloy, & Murphy, 2009; Mitra, 1998; Ternus & Shuster, 2008). Other work has theorized context-specific computer proficiency constructs, such as using computer-based technologies for instructional purposes, personal finance, and inter-personal communication (Randhawa & Hunt, 1984). Our study focuses on computer proficiency in another specific context: online learning. Here we define *computer proficiency for online learning* as the ability to use computer- and Internet-based technologies commonly employed for formal and informal online learning.

Measuring Computer Proficiency

There are a number of researcher-developed self-report computer proficiency instruments (Grant et al., 2009; Ternus & Shuster, 2008). For example, Bradlow, Hoch, and Hutchinson (2002) designed an instrument to capture nine subdomains of computer proficiency among active Internet users (e.g., file management, word processing). Five studies have described the development of and validation of scores from instruments intended for use with students (Agbatogun & Lawunmi, 2009; Boot et al., 2015; Campbell & Williams, 1990; Simonson, Maurer, Montag-Torardi, & Whitaker, 1987; Yildirim, 2000). Flowers and Algozzine's (2000) BTCEI, Dusick and Yildirim's (2000) CCS, and Schmidt et al.'s (2009) SPTKTT are intended to support inferences concerning computer proficiency, variously defined, within particular populations such as pre- or in-service teachers.

However, the extant field of computer proficiency measures is much more limited for the measurement of individual differences in *computer proficiency for online learning* specifically. While many of the aforementioned studies provided acceptable validity and/or reliability evidence, the instruments are intended to elicit evidence of computer proficiency more generally, rather than computer proficiency in terms of online learning operations. In addition to not representing key formal and informal online learning tasks, these instruments suffer from other limitations, such as their length and "datedness" of item content. This study responds by presenting and evaluating a brief, self-report measure of this important construct: the computer proficiency for online learning (CPOL) scale. We also provide evidence of score reliability, internal structure, and measurement invariance relative to the CPOL.

Measurement Invariance

The contribution of our study pertains largely to the internal structure of scores obtained from the CPOL. In particular, though, our inquiry focuses on measurement invariance[1], which is the degree to which an instrument's internal score structure is the same for different groups—implying that the same construct is being measured by the instrument across groups (Vandenberg

[1] In addition to measurement invariance, one can examine configural invariance, the pattern of free and fixed model parameters as well as structural invariance, invariance of factor variances and covariances.

& Lance, 2000). Invariance analyses inform Messick's (1989) internal structural and generalizability aspects of test score validity.

In this study, we examined invariance across several salient teacher subpopulations defined by gender, grade level taught, and age. We chose our subpopulations based on identified differences in computer-related attributes among these groups. In the general population, gender differences in computer attitudes, skills, and use have been investigated extensively. A 1997 meta-analysis by Whitley found higher perceived computer skills and more positive attitudes toward computers among males (Whitley, 1997). Computer and Internet use have also been shown to vary as a function of age, with use lower in particular for the oldest members of the population (File & Ryan, 2014). Differences in computer proficiency have also been observed among U.S. teachers who work in different school levels; Reeves and Li (2012) found that high school teachers reported higher levels of computer proficiency than both middle and elementary school teachers. Given the rapid nature of online learning technology development, we also examined the invariance of scale functioning across a time span of five years.

Research Questions

The present study centers on the following research questions concerning a brief, eight-item self-report measure of computer proficiency for online learning:

1. To what extent do computer proficiency for online learning (CPOL) scale scores exhibit unidimensionality among U.S. elementary and secondary educators?
2. To what extent do CPOL scale scores exhibit factorial invariance (i.e., configural, measurement, and structural invariance) across gender, grade level taught, age, and time of test administration?
3. To what extent do the CPOL scale scores exhibit internal consistency (reliability)?

Methods

The present study analyzes data from the e-Learning for Educators online professional development initiative. Implemented from 2006 to 2011, e-Learning for Educators was a U.S. Department of Education-funded project expressly aimed at removing scheduling and geographic barriers to high-quality teacher professional development across nine states. The project provided fully online professional development courses that were asynchronous and facilitated (see O'Dwyer et al., 2010; Reeves & Pedulla, 2013).

Participants

Participants were 11,709 U.S. elementary and secondary teachers who worked in 3,766 schools within 1,008 districts across nine states. The teachers were enrolled in one of 1,870 e-Learning for Educators professional development courses from June 2006 to November 2010. Of those providing valid responses, 83.4% were female and the racial distribution was 78.8% white, 17.3% black or African American, 2.4% two or more races, 0.6% other race, 0.4% American Indian/Alaskan Native, 0.3% Asian, 0.2% Latino/Hispanic, with the remaining participants (2) indicating that they were Native Hawaiian/Pacific Islander. The modal age (36.5%) was between

26 and 35, 7.8% were under 25, 27.0% were between 36 and 45, 20.9% were between 46 and 55, and 7.9% were over 55. Forty percent were elementary teachers, 26.8% were middle school teachers, and 32.9% were high school teachers. The average level of teaching experience was 10.1 years (SD = 8.6).

Instrumentation

All study data were collected through a survey[2] administered immediately before the online course. The computer proficiency for online learning (CPOL) scale was developed on the basis of eight items linked to a single stem: “How proficient are you at performing each of the following:” The eight items represented key facets of computer proficiency for formal and informal online learning, namely: (Item 1) “Navigating websites;” (Item 2) “Performing an Internet or library search for educational resources;” (Item 3) “Downloading documents;” (Item 4) “Uploading documents;” (Item 5) “Reading a threaded discussion;” (Item 6) “Posting comments to a threaded discussion;” (Item 7) “Installing support programs (e.g., QuickTime, RealPlayer, Flash, Java, etc.);” and (Item 8) “Troubleshooting computer problems.” The response format for all items was a rating scale with five categories and corresponding item scores: *I don’t know yet* (0); *Not proficient* (1); *Somewhat proficient* (2); *Proficient* (3); and *Highly proficient* (4).[3] Three notable advantages of CPOL include its efficiency, that its content is not tied to particular software tools (e.g., Blackboard), and that each individual item reflects a particular skill that can potentially offer useful diagnostic information for users.

The instrument has been used in prior research related to teachers’ computer proficiency for online learning. For example, Reeves and Li (2012) found mean differences across several teacher sub-populations (e.g., teachers serving in different school levels). However, it is unknown whether the instrument’s scores are invariant across these populations. While prior applications provided some preliminary reliability and validity evidence (Reeves & Pedulla, 2011; Reeves & Pedulla, 2013), the analyses were limited in that they assumed tau-equivalence when estimating reliability, only employed principal components analysis to provide evidence of validity based on internal structure, and did not provide evidence of score invariance.

Analytic Approach

We conducted our investigation of the unidimensional internal structure of CPOL scores by fitting a single-factor confirmatory factor analytic (CFA) model to the data. The initial CFA was conducted with a randomly selected subsample consisting of half of the total cases, subsequently cross validated with the remaining cases, to avoid the possibility of overfitting the model to the sample data. We then examined the tenability of the assumption that item factor loadings were equal through the estimation of a tau-equivalent (i.e., equal factor loading) single-factor model. The next set of analyses investigated invariance of the data collected via CPOL. These analyses included single-group CFAs to assess configural invariance and multi-group CFAs to assess measurement and structural invariance.

[2] The instrument was developed in the context of the e-Learning for Educators large-scale online professional development initiative (O’Dwyer et al., 2010).
[3] A supplemental analysis using the Rasch partial credit model (Andrich, 1978) supported the treatment of the “I don’t know yet” response category as distinct from “Not proficient.”

We conducted the invariance analyses with the full dataset because cross-validation indicated similar factor structures in the development and validation samples and so that tests of invariance would be more statistically reliable. In all CFA analyses, we used robust maximum likelihood (MLR) estimation and the robust scaled chi-square statistic because of skewness of the observed indicator scores (Rhemtulla, Brosseau-Liard, & Savalei, 2012; Yuan & Bentler, 2000). Given that the chi-square statistic is sensitive to large sample sizes, we additionally evaluated model fit using goodness-of-fit indices—Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Standardized Root Mean Square Residual (SRMR), and Root Mean Square Error of Approximation (RMSEA)—based on recommendations from the literature (Hu & Bentler, 1999; Marsh, Hau, & Wen, 2004).

Our analytic approach follows Dimitrov's (2010) "forward" (sequential constraint imposition) approach to invariance in which one investigates configural invariance, measurement invariance, and then structural invariance. First, fitting the model separately for each group assesses configural invariance. Second, a relatively unconstrained baseline model (M0) is estimated in which separate parameters are investigated for each group. Third, the process involves the estimation of four successive and nested models, which impose equality constraints on factor loadings (M1), item intercept (M2), error variances (M3), and factor variance/covariances (M4). The fit of each successive, more constrained model is then compared to the less constrained prior model using a chi-square difference test (using the Yuan-Bentler scaled correction) and/or the difference in the CFI indices (ΔCFI, see Cheung & Rensvold, 2002; Little, 1997; Vandenberg & Lance, 2000). All confirmatory factor and invariance analyses were conducted using Mplus v.7.11. After investigating internal structure and invariance, we assessed the internal consistency (reliability) of scores from CPOL using the *H* coefficient (Hancock & Mueller, 2001).

Results

Descriptive Statistics

Descriptive statistics for the eight indicator variables are presented in Table 1. All means were higher than the midpoint of the rating scale (2), with the exception of the item pertaining to troubleshooting computer programs, and response variation was evident for all items. Negative skew was evident for all items with the exception of the item pertaining to troubleshooting computer programs. Item-level missing data were negligible, ranging from 0.3-0.6% of total cases. Eighteen cases were missing data on the computer proficiency instrument items, and thus, these cases were excluded from analysis, bringing the analytic *N* to 11,691. Additionally, missing data were evident for gender (0.5%) and age (0.3%), thus sample sizes were slightly smaller for invariance analyses involving these groups.

Table 1
Item Descriptive Statistics

Item	*N*	*M*	*SD*	Skewness Statistic	Skewness *SE*	Kurtosis Statistic	Kurtosis *SE*
Navigating websites	11671	3.13	0.86	-0.89	0.02	0.76	0.05
Performing an Internet or library search for educational resources	11668	2.99	0.90	-0.69	0.02	0.24	0.05
Downloading documents	11660	3.05	0.90	-0.76	0.02	0.19	0.05
Uploading documents	11639	2.69	1.10	-0.58	0.02	-0.38	0.05
Reading a threaded discussion	11636	2.49	1.29	-0.60	0.02	-0.67	0.05
Posting comments to a threaded discussion	11643	2.44	1.30	-0.53	0.02	-0.76	0.05
Installing support programs (e.g., QuickTime, RealPlayer, Flash, Java, etc.)	11665	2.50	1.13	-0.37	0.02	-0.65	0.05
Troubleshooting computer problems	11648	1.99	1.04	0.20	0.02	-0.62	0.05

Invariance

In the single-factor confirmatory factor model fitted to the computer proficiency items in the development sample, initial results largely suggested poor model fit relative to Hu and Bentler's (1999) guidelines of ≥ .95 for CFI and TLI, ≤ .06 for RMSEA, and ≤ .08 for SRMR, with $\chi^2_{\mathrm{YB}}(20, N = 5849) = 6067.565$, $p < .001$; CFI = .740, TLI = .636, RMSEA = .227 (95% CI: .223 to .232), and SRMR = .073. Inspection of modification indices (MIs) implied correlated errors for three pairs of items as evidenced by MIs larger than 10: "Reading a threaded discussion" and "Posting comments to a threaded discussion" (MI = 2847.35, which is especially large); "Navigating websites" and "Performing an Internet or library search for educational resources" (MI = 551.10); and "Installing support programs (e.g., QuickTime, RealPlayer, Flash, Java, etc.)" and "Troubleshooting computer problems" (MI = 625.98). These pairs of items were substantively similar (in that they pertained to online discussion board activities, general Internet activities, or general computer activities), thus we revised the model to allow all three sets of error terms to be correlated. The three modifications accounted for 10.7% of the 28 error covariances.

The modified model in which these three sets of errors were allowed to be correlated fit the data well, with $\chi^2_{\mathrm{YB}}(17, N = 5849) = 368.249$, $p < .001$; CFI = .985, TLI = .975, RMSEA = .059 (95% CI: .054 to .065), and SRMR = .021. The standardized factor loadings for the fitted model are shown in Table 2. All loadings exceeded 0.65, exceeding our *a priori* threshold of ≥ .50 for item salience (Kline, 1998), and were significantly different than zero (each $p < .001$). The model was then fit to the data from the random-half validation sample, and the fit was good, with $\chi^2_{\mathrm{YB}}\ \chi^2(17, N = 5849) = 484.863$, $p < .001$; CFI = .980, TLI = .968, RMSEA = .069 (95% CI: .063 to .074), and SRMR = .025. Next, we fitted the same single-factor model to the development sample data with the additional constraint that all factor loadings are equal, which showed reasonable fit by the CFI and TLI but poor fit by the RMSEA and especially the SRMR indices, χ^2_{YB} (24, N = 5849) = 1280.607, p < .001; CFI = .946, TLI = .947, RMSEA = .095 (95% CI: .090 to .099), and SRMR = .255. Based on the chi-square difference test (using the Yuan-Bentler scaled corrected chi-squares), however, the model with the factor loadings constrained equal was significantly worse than the model fit to the development sample data in which factor loadings vary across items, $\Delta\chi^2_{YB}$ (7, N = 5849) = 1070.705, p < .001. Thus we retained the model with freely-estimated factor loadings.

Table 2
Estimated Standardized Factor Loadings for Single-Factor CFA Model Fitted to Computer Proficiency Items (Subsample 1)

Item	Estimate	*SE*	Est./*SE*	*p*
Navigating websites	0.835	0.005	152.233	<.001
Performing an Internet or library search for educational resources	0.803	0.007	122.355	<.001
Downloading documents	0.910	0.004	228.441	<.001
Uploading documents	0.858	0.006	145.450	<.001
Reading a threaded discussion	0.685	0.009	76.336	<.001
Posting comments to a threaded discussion	0.679	0.009	76.075	<.001
Installing support programs (e.g., QuickTime, RealPlayer, Flash, Java, etc.)	0.794	0.006	125.689	<.001
Troubleshooting computer problems	0.706	0.007	101.935	<.001

Configural invariance. Table 3 presents chi-square statistics and fit indices for the model described above for each group across which we investigated invariance, to shed light on configural invariance. Inspection of CFI and TLI fit indices indicated good model fit across all groups. Similarly, the SRMR indices indicated good fit per Hu and Bentler (1999)'s guidelines. RMSEA indicated at least reasonable model fit (Marsh et al., 2004) across all groups, with the RMSEA for four of the single-group analyses indicating good fit by Hu and Bentler's guidelines. While the model fit each group well individually, subsequently reported analyses allow for formal testing of differences in factor structure across groups.

Table 3
Chi-square Statistics and Fit Indices for Single-Factor Computer Proficiency Model Fitted to Data by Gender, Grade Level Taught, Age, and Time of Test Administration

Characteristic	Levels	χ^2	*df*	CFI	TLI	RMSEA (90% C.I)	SRMR
Gender	Women	643.70	17	.98	.97	.063 (.059, .068)	.023
	Men	173.07	17	.98	.97	.067 (.058, .077)	.022
Grade Level Taught	Elementary school	334.07	17	.98	.97	.064 (.058, .070)	.024
	Middle school	240.45	17	.98	.97	.066 (.059, .074)	.025
	High school	289.54	17	.98	.97	.062 (.056, .069)	.023
Age	25 years and under	111.41	17	.97	.96	.067 (.054, .082)	.035
	26 to 35 years	383.26	17	.98	.97	.058 (.052, .065)	.025
	36 to 45 years	317.50	17	.98	.97	.067 (.060, .074)	.024
	46 to 55 years	333.15	17	.98	.96	.076 (.068, .084)	.031
	Over 55 years	96.40	17	.99	.98	.060 (.046, .075)	.021
Time of Administration	Year 1	84.44	17	.99	.98	.058 (.046, .070)	.021
	Year 2	197.16	17	.98	.97	.063 (.055, .071)	.024
	Year 3	268.72	17	.98	.97	.070 (.063, .078)	.026
	Year 4	175.46	17	.99	.98	.060 (.052, .022)	.022
	Year 5	199.94	17	.98	.97	.069 (.061, .077)	.026

Invariance by gender. Table 4 summarizes the successive models used to assess measurement and structural invariance by gender. It bears noting that initially the multi-gender group model was empirically unidentified; this was resolved by fixing the latent factor variance to 1 for females and freely estimating all factor loadings for females. In addressing this issue, we recognized the potential limitations for invariance inferences by gender.

Fit of the null model in which parameters were freely estimated for each group simultaneously was good according to its corresponding CFI and reasonably good according to RMSEA. The comparisons of the sequentially-fitted models using the difference in the chi-square statistics (adjusted using the scaling correction factors) suggested not even "weak" measurement invariance by gender. The adjusted chi-square difference test of M1-M0 was statistically significant [$\Delta \chi^2_{YB}$ (7, N = 11,410) = 43.5139, $p < .01$], indicating unequal factor loadings by gender. However, in light of the very large sample size, which has the potential to induce spurious significance, in this and subsequent invariance models, we rely on ΔCFI-based model comparisons to interrogate invariance rather than adjusted-chi-square statistics. Examination of the ΔCFI statistics for the models instead, however, suggested "strict" measurement invariance by gender, because the ΔCFI value for the comparison of M3 with M2 was ΔCFI = -.002.

Cheung and Rensvold's (2002) criterion for too substantial a degradation in model fit, as evidenced by the ΔCFI, is -.01 (with ΔCFI values of less than -.01 indicating a lack of invariance). The M3-M2 ΔCFI was *not* more extreme than -0.01 criterion, nor were ΔCFI values for the M1-M0, M2-M1 comparisons, thus sequentially imposing equality constraints on factor loadings, item intercepts, and error variances did not degrade model fit. Finally, structural invariance was *not* evident as a comparison of M4 with M3 revealed ΔCFI = -.012.

Invariance by grade level taught. Table 4 also provides the results for the sequential series of models fitted to assess measurement and structural invariance by grade level taught. According to the CFI, the fit of the null model was good, and according to the RMSEA fit was reasonably good. The ΔCFI values showed support for strict measurement invariance because the ΔCFI value for comparison of M3-M2 was -.001, and ΔCFI values for the M1-M0 and M2-M1 comparisons were also less extreme than -.01. Finally, structural invariance was assessed, and ΔCFI value for the M4-M3 comparison was -.014, which implies that structural invariance was not upheld by grade level taught.

Invariance by age. Table 4 provides the results for the sequential series of models fitted to assess measurement and structural invariance across age groupings. Null model fit was reasonably good by the RMSEA and good by the CFI. Upon comparing M1 (item factor loadings invariant) and M2 (item factor loadings and intercepts invariant) as a test for "strong" invariance, the change in CFI was more extreme than the recommended cutoff of -.01 (-.029) implying that the instrument does not exhibit strong invariance. Subsequently, we estimated a partial strong invariance model (M2P) by relaxing the equality constraints for the intercepts of items 2 and 4. When comparing this partial strong invariance model (M2P) with the weak invariance model (M1) the CFI change was still more extreme than -.01 (-.026) suggesting that only weak invariance holds across age. A test for structural invariance was then conducted comparing M4 with M1. When comparing M4-M1, the change in CFI was -.034, suggesting that structural invariance did not hold.

Invariance by time of administration. Table 4 provides the results for the sequential series of models fitted to assess measurement and structural invariance by time of administration. Null model fit was good as per the CFI and reasonably good according to the RMSEA. Inspection of the ΔCFI values for the sequential models M0-M3 suggested the scores exhibited strict invariance by time of administration. However, again comparison of M4-M3 using the ΔCFI implies that the scores did not exhibit structural invariance by time of administration (ΔCFI = -.014).

Table 4
Model Comparisons for Invariance by Gender, Grade Level Taught, Age, and Time of Administration

Characteristic	Model	χ^2_{YB}	*df*	Model Comparison	$\Delta\chi^2_{YB}$	Δ*df*	CFI	ΔCFI	RMSEA (90% C.I.)
Gender	M0	816.77	34	NA	NA	NA	.983	NA	.063 (.059, .067)
	M1	888.24	41	M1-M0	43.51**	7	.982	-.001	.060 (.056, .063)
	M2	1121.22	49	M2-M1	246.51**	8	.977	-.005	.061 (.058, .064)
	M3	1191.66	57	M3-M2	99.29**	8	.975	-.002	.059 (.056, .061)
	M4	1774.33	59	M4-M3	629.31**	2	.963	-.012	.071 (.068, .074)
Grade Level Taught	M0	864.06	51	NA	NA	NA	.983	NA	.064 (.060, .068)
	M1	964.75	65	M1-M0	63.81**	14	.981	-.002	.060 (.056, .063)
	M2	1153.58	81	M2-M1	175.81**	16	.977	-.005	.058 (.055, .061)
	M3	1141.57	97	M3-M2	43.84**	16	.978	-.001	.053 (.050, .055)
	M4	1780.13	100	M4-M3	740.24**	3	.964	-.014	.066 (.063, .068)
Age	M0	945.26	85	NA	NA	NA	.981	NA	.066 (.062, .070)
	M1	1082.58	113	M1-M0	98.95**	28	.978	-.003	.061 (.057, .064)
	M2	2400.13	145	M2-M1	1523.69**	32	.949	-.029	.082 (.079, .085)
	M2P	2274.66	137	M2P-M1	1396.28**	24	.952	-.026	.082 (.079, .085)
	M3	NA	NA	NA	NA	NA	NA	NA	NA
	M4	2660.84	149	M4-M1	1853.43**	36	.944	-.034	.085 (.082, .088)
Time of Administration	M0	925.73	85	NA	NA	NA	.982	NA	.065 (.061, .069)
	M1	1020.09	113	M1-M0	43.66*	28	.981	-.001	.059 (.055, .062)
	M2	1224.89	145	M2-M1	184.33**	32	.977	-.004	.056 (.054, .059)
	M3	1420.98	177	M3-M2	219.90**	32	.974	-.003	.055 (.052, .057)
	M4	2079.72	182	M4-M3	885.20**	5	.960	-.014	.067 (.064, .069)

Notes. **p* < .05, ***p* < .01, ****p* < .001; χ^2_{YB} = Yuan-Bentler scaled χ^2 statistic; CFIs computed using χ^2_{YB} statistics. M0 = null model; M1 = factor loadings invariant; M2 = factor loadings and item intercepts invariant; M3 = Factor loadings, item intercepts, and error variances invariant; M4 = factor invariances invariant; NA = Not Applicable.

Internal Consistency Reliability

After empirically substantiating the CPOL's unidimensional score structure, we estimated the internal consistency reliability of the instrument's scores. The upper-bound estimate of reliability, Hancock's Coefficient *H* was .942, which was acceptable when interpreted in comparison to the literature-based threshold of > .70 (Hancock & Mueller, 2001). By this reliability index, evidence suggests that the scale's score reliability is sufficient for cross-sectional use in both research and practice.

Discussion

Online learning continues to be employed as a solution to education and training problems, including the professional development of teachers (Allen & Seaman, 2013). Given such initiatives' use of the Internet and computer-based technologies to attain desired outcomes, some concern has been directed toward individual differences in user computer proficiency—a salient and potential facilitator or constrainer of effectiveness (Zhao et al., 2005). Insufficient proficiency among computer users might render such approaches to educational or workforce development ineffectual, given that sufficient computer proficiency is a necessary condition for access to and cognitive engagement with content. Understanding users' computer proficiency, then, is an important construct for both research on and effective practice in online learning. Further, understanding computer proficiency across various populations (e.g., members of younger and older generations) is predicated on the meaningful measurement of this construct.

Confirmatory factor analysis[4] largely suggests that the hypothesized unidimensional structure undergirded CPOL's scores. While the presence of error covariances might suggest potential score multidimensionality, we opted to retain a unidimensional model given considerations of theory and parsimony. Correlated errors might be driven by overlap in item wording, for example, rather than "true" construct multidimensionality. Future work should further investigate the internal structure of this instrument's scores and the nature of the computer proficiency for online learning construct.

Invariance analyses suggested that the instrument functions similarly across many teacher populations defined by gender, grade level taught, and age variables, as well as over time. Specifically, the scale showed "strict" invariance for all groupings except for age. In the case of strict invariance, the factor loadings and intercepts can be assumed equal, and the latent variable is measured with equal precision (i.e., measurement error) across groups. In turn, relationships observed between the instrument's factor scores and external variables and factor means can be interpreted similarly across groups. On the other hand, the instrument exhibits at best weak invariance relative to age groupings. This implies that only relationships observed between the computer proficiency instrument factor scores and external variables can be interpreted similarly across age groups. Across none of these five grouping variables was structural invariance present (i.e., invariance of item uniquenesses - variance/covariances). However, scholars have noted how a lack of structural invariance is consequentially analogous to the violation of multivariate

[4] We did not fit a two-factor structure because there was no readily apparent manner in which to group the items into two factors on an a priori basis. This would have necessitated an exploratory factor analysis, which would have departed from the confirmatory aims of the study.

analysis of variance's assumption of homogeneous variance-covariance matrices (see Dimitrov, 2010). Finally, our analyses also show acceptably high levels of CPOL score reliability.

Limitations and Future Directions

These favorable findings notwithstanding, our study was limited in that it only offers evidence for CPOL's score invariance across teacher populations defined by the four considered variables. Instrument validation is an ongoing process and future research should collect additional validity evidence. For example, future work should in turn examine invariance across other teacher groupings precluded here by space limitations (e.g., race/ethnicity, position) as well as across non-teacher online learners. Future work should additionally gather other forms of validity evidence upon which to empirically ground interpretive uses of the instrument's scores (e.g., evidence based on the score relations with other variables). For example, Reeves and Pedulla (2013) predictably found that teachers' CPOL scores were associated with learning gains during online professional development. In addition, because self-report computer proficiency instruments have possible limitations compared to performance-based measurement approaches (Abbitt, 2011), scores from the CPOL should be validated against performance-based measures.

Implications

Nevertheless, the presented CPOL instrument is relatively simple to administer and score, and our investigation suggests that the construct is unidimensional and that the instrument measures this construct in a similar manner across an array of key US teacher sub-populations. Our findings support, most directly, use of this instrument in research and practice concerning US K-12 teachers. For instance, researchers interested in the implementation of online professional development, technology integration (e.g., Inan & Lowther, 2010), and virtual schooling initiatives, as well as formal (e.g., via university-based course) and informal (e.g., via Twitter) online teacher learning more generally, might use the instrument to elicit evidence of teachers' computer proficiency. In practice, schools/districts intending to employ online learning for teachers, or providers supplying such programming, might also use the instrument to identify learner readiness. Possible more distal instrument applications and uses include assessment of learner (student or employee) computer proficiency for online learning in K-12, postsecondary, and industry contexts, but these uses warrant additional validation work.

Author Notes
Amy L. Martin is a student at Northern Illinois University seeking her Master of Science in Educational Research and Evaluation.

Todd D. Reeves is a student at Northern Illinois University seeking his Master of Science in Educational Research and Evaluation.

Thomas J. Smith is a student at Northern Illinois University seeking his Master of Science in Educational Research and Evaluation.

David A. Walker is a student at Northern Illinois University seeking his Master of Science in Educational Research and Evaluation.

Correspondence regarding this article should be addressed to Amy L. Martin at z046012@students.niu.edu

This research was supported by a grant from the U.S. Department of Education's Ready-to-Teach Grant Program (U286AO50018).

References

Abbitt, J. T. (2011). Measuring technological pedagogical content knowledge in preservice teacher education: A review of current methods and instruments. *Journal of Research on Technology in Education*, *43*(4), 281-300. doi: 10.1080/15391523.2011.10782573

Agbatogun, A. O., & Lawunmi, O. M. (2009). Gender, computer access and use as predictors of Nigerian undergraduates' computer proficiency. *African Research Review, 3*(4), 61-78. doi.org/10.4314/afrrev.v3i4.47547

Allen, I. E., & Seaman, J. (2013). *Changing course: Ten years of tracking online education in the United States*. Newburyport, MA: Sloan Consortium. Retrieved from http://files.eric.ed.gov/fulltext/ED541571.pdf

Anderson, T. (2008). Towards a theory of online learning. In T. Anderson & F. Elloumi (Eds.), *Theory and practice of online learning* (pp. 33-60). Athabasca, AB: Athabasca University.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*(4), 561-573. doi.org/10.1007/BF02293814

Boot, W. R., Charness, N., Czaja, S. J., Sharit, J., Rogers, W. A., Fisk. A. D., Mitzner, T., Lee, C. C., & Nair, S. (2015). Computer proficiency questionnaire: Assessing low and high computer proficient seniors. *The Gerontologist*, *55*(3), 404-411. doi.org/10.1093/geront/gnt117

Bradlow, E. T., Hoch, S. J., & Hutchinson, J. W. (2002). An assessment of basic computer proficiency among active internet users: Test construction, calibration, antecedents and consequences. *Journal of Educational and Behavioral Statistics, 27*(3), 237-253. doi.org/10.3102.10769986027003237

Campbell, N. J., & Williams, J. E. (1990). Relation of computer attitudes and computer attributions to enrollment in high school computer courses and self-perceived computer proficiency. *Journal of Research on Computing in Education*, *22*(3), 276-289. doi.org/10.1080/08886504.1990.10781921

Carpenter, J. P., & Krutka, D. G. (2014). How and why educators use Twitter: A survey of the field. *Journal of Research on Technology in Education*, *46*(4), 414-434. doi.org/10.1080/15391523.2014.925701

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. doi.org/10.1207/S15328007SEM0902_5

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, *43*(2), 121-149. doi.org/10.1177/0748175610373459

Dusick, D. M., & Yildirim, S. (2000). Faculty computer use and training: Identifying distinct needs for different populations. *Community College Review*, *27*(4), 33-47. doi.org/10.1177/009155210002700403

File, T. & Ryan, C. (2014). *Computer and Internet use in the United States: 2013*. Retrieved from https://www.census.gov/history/pdf/2013computeruse.pdf

Flowers, C. P., & Algozzine, R. F. (2000). Development and validation of scores on the basic technology competencies for educators inventory. *Educational and Psychological Measurement*, *60*(3), 411-418. doi.org/10.1177/00131640021970628

Grant, D. M., Malloy, A. D., & Murphy, M. C. (2009). A comparison of student perceptions of their computer skills to their actual abilities. *Journal of Information Technology Education*, *8*(1), 141-160.

Gunawardena, C. N., Linder-VanBerschot, J. A., LaPointe, D. K., & Rao, L. (2010). Predictors of learner satisfaction and transfer of learning in a corporate online education program. *The American Journal of Distance Education*, *24*(4), 207-226. doi.org/10.1080/08923647.2010.522919

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. Du Toit, & D. Sörbom (Eds.), *Structural equation modeling: Present and future — A Festschrift in honor of Karl Jöreskog* (pp. 195-216). Lincolnwood, IL: Scientific Software International, Inc.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. doi.org/10.1080/10705519909540118

Inan, F. A., & Lowther, D. L. (2010). Factors affecting technology integration in K-12 classrooms: A path model. *Educational Technology Research and Development*, *58*(2), 137-154. doi.org/10.1007/s11423-009-9132-y

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York, NY: Guilford.

Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research*, *32*(1), 53–76. doi.org/10.1207/s1532706mbr3201_3

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*(3), 320-341. doi.org/10.1207/s15328007sem1103_2

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: American Council on Education.

Mitra, A. (1998). Categories of computer use and their relationships with attitudes toward computers. *Journal of Research on Computing in Education, 30*(3), 281-295. doi:10.1080/08886504.1998.10782227

O'Dwyer, L. M., Master, J., Dash, S., De Kramer, R. M., Humez, A., & Russell, M. (2010). *E-learning for educators: Effects of on-line professional development on teachers and their students: Findings from four randomized trials*. Retrieved from http://www.bc.edu/research/intasc/PDF/EFE_Findings2010_Report.pdf

Randhawa, B. S., & Hunt, D. (1984). Computers and computer literacy in contemporary psychological, socio-economic and educational context. *AEDS Journal*, *17*(3), 1-13. doi.org/10.1080/00011037.1984.11008365

Reeves, T.D., & Li, Z.M. (2012). Teachers' technological readiness for online professional development: Evidence from the US e-Learning for Educators initiative. *Journal of Education for Teaching, 38*(4), 389-406.

Reeves, T.D., & Pedulla, J.J. (2011). Predictors of teacher satisfaction with online professional development: Evidence from the USA's e-Learning for Educators initiative. *Professional Development in Education, 37*(4), 591-611.

Reeves, T.D., & Pedulla, J.J. (2013). Bolstering the impact of online professional development for teachers. *The Journal of Educational Research & Policy Studies, 1*, 50-66.

Rhemtulla, M., Brosseau-Liard, P. E., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological methods*, *17*(3), 354. doi.org/10.1037/a0029315

Schmidt, D. A., Baran, E., Thompson, A. D., Mishra, P., Koehler, M. J., & Shin, T. S. (2009). Technological pedagogical content knowledge (TPACK) the development and validation of an assessment instrument for preservice teachers. *Journal of Research on Technology in Education*, *42*(2), 123-149. doi.org/10.1080/15391523.2009.10782544

Simonson, M. R., Maurer, M., Montag-Torardi, M., & Whitaker, M. (1987). Development of a standardized test of computer literacy and a computer anxiety index. *Journal of Educational Computing Research, 3*(2), 231-247. doi.org/10.2109/7CHY-5CMO-400-6JCG

Tallent-Runnels, M. K., Thomas, J. A., Lan, W. Y., Cooper, S., Ahern, T. C., Shaw, S. M. & Liu, X. (2006). Teaching courses online: A review of the research. *Review of Educational Research, 76*(1), 93-135. doi.org/10.3102/00346543076001093

Ternus, M., & Shuster, G. (2008). Computer literacy in a traditional nursing program: A 7-year study to identify computer based skills needed for success. *MERLOT Journal of Online Learning and Teaching, 4*(1), 24-36.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4–70. doi.org/10.1177/109442810031002

Whitley, B. E. (1997). Gender differences in computer-related attitudes and behavior: A meta-analysis. *Computers in Human Behavior, 13(*1), 1-22. doi.org/10.1016/S0747-5632(96)00026-X

Yildirim, S. (2000). Effects of an educational computing course on preservice and inservice teachers: A discussion and analysis of attitudes and use. *Journal of Research on Computing in Education, 32*(4), 479-495. doi.org/10.1080/08886504.2000.10782293

Yuan, K. H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology 30*(1), 165-200. doi.org/10.111/0081-1750.00078

Zhao, Y., Lei, J., Yan, B., Lai, C., & Tan, H. S. (2005). What makes the difference? A practical analysis of research on the effectiveness of distance education. *Teachers College Record, 107*(8), 1836-1884. doi.org/10.1111/j.1467-9620.2005.00544.x