Educational Policy Studies and Evaluation Faculty Publications

Educational Policy Studies and Evaluation

9-2020

Applying the Rasch Model to Evaluate the Self-Directed Online Learning Scale (SDOLS) for Graduate Students

Hongwei Yang
University of West Florida

Jian Su
University of Tennessee

Kelly D. Bradley
University of Kentucky, kelly.bradley@uky.edu

Follow this and additional works at: https://uknowledge.uky.edu/epe_facpub

Part of the Educational Assessment, Evaluation, and Research Commons, and the Higher Education Commons

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Applying the Rasch Model to Evaluate the Self-Directed Online Learning Scale (SDOLS) for Graduate Students

Applying the Rasch Model to Evaluate the Self-Directed Online Learning Scale (SDOLS) for Graduate Students

Hongwei Yang, Ph.D.[1], Jian Su, Ph.D.[*,2], and Kelly D. Bradley, Ph.D.[3]
[1]University of West Florida, [2]University of Tennessee, [3]University of Kentucky

Abstract

With the rapid growth of online learning and the increased attention paid to student attrition in online programs, much research has been aimed at studying the effectiveness of online education to improve students' online learning experience and student retention. Utilizing the online learning literature as a multi-faceted theoretical framework, the study developed and employed a new survey instrument. The Self-Directed Online Learning Scale (SDOLS) was used to examine graduate student perceptions of effectiveness of online learning environments as demonstrated by their ability to take charge of their own learning, and to identify key factors in instructional design for effective improvements. The study applied the Rasch rating scale model to evaluate and validate SDOLS through a psychometric lens to establish the reliability and validity of SDOLS. Results from Rasch analysis addressed two research questions. First, evidence was found to generally support the new instrument as being psychometrically sound but three problematic items were also identified as grounds for future improvement of SDOLS. Second, the study assessed the importance of various factors as measured by the SDOLS items in contributing to students' ability to self-manage their own online learning. Finally, the new instrument is expected to contribute to the work of various stakeholders in online education and can serve to improve students' online learning experience and effectiveness, increase online retention rates, and reduce online dropouts.

Keywords: self-directed learning, online teaching and learning, scale development, Rasch analysis

Introduction

Existing research on online education effectiveness has identified essential characteristics of a successful online learning environment (Hone & Said, 2016; Mayes, Luebeck, Ku, Akarasriworn, & Korkmaz, 2011; Palloff & Pratt, 2007). Among them, students' self-directed learning (SDL), or self-management of learning is one consistent and foundational factor recognized in online learning readiness and effectiveness (Prior, Mazanov, Meacheam, Heaslip, & Hanson, 2016; Rovai, Ponton, Wighting, & Baker, 2007). Research indicates SDL contributes to learners' abilities to manage their overall learning activities, to think critically, and to cognitively monitor their learning performance when navigating through the learning process. SDL also helps students better interact and collaborate with the instructor and peers for feedback and support (Beach, 2017; Garrison, 1997, p. 21; Hyland & Kranzow, 2011, p. 15; Kim, Olfman, Ryan, & Eryilmaz, 2014, p. 150).

Conceptual Framework

SDL has been a core theoretical construct in adult education and its research has evolved over time (Garrison, 1997). The existing literature on SDL has established an understanding of SDL as both a process and a personal attribute (Song & Hill, 2007, p. 38).

Knowles (1975) defined SDL as adult students' ability to self-manage their own learning, and his work served as a how-to book for adult students planning to develop competency as self-directed learners (Long, 1977). Next, Caffarella (1993, pp. 25–26) described three principal ideas underlying the SDL process: (a) a self-initiated process of learning, (b) more learner autonomy, and (c) greater control by the learner. Under SDL, learners take primary responsibility for their own learning to meet their unique needs and achieve personal goals. Hiemstra (1994) interpreted self-directed learning as indicating individual adults had the capacity to plan, navigate, and evaluate their own learning on the path to their personal learning goals. By contrast, Garrison (1997) presented a more comprehensive theoretical model of self-directed learning, focused on the learning process itself containing both motivational and cognitive aspects of learning. This model integrated three overlapping dimensions related to learning in an educational setting: (a) external management, (b) internal monitoring, and (c) motivation. Finally, noting that SDL may function differently in different learning situations, Song and Hill (2007) examined various learning contexts (the online context, in particular) where self-direction in learning takes place. They argued that a better understanding of trans-contextual SDL attributes unique to the online setting contributes to better online teaching and learning experiences.

Online Teaching and Learning

The past decades have witnessed a rapid development of technology contributing to the rise of online teaching and learning, which has led to increasing interest in SDL (Chou & Chen, 2008). Known for its flexibility allowing learning to accommodate adult learners' busy schedules, online education has been the right place for them to take charge of their own learning. On one hand, online learning supports the self-management dimension of Garrison's SDL model (1997). Online learning platforms lend themselves to greater learner control and autonomy, and ultimately, intrinsic motivation to learn. Due to being able to

self-direct their own learning, learners more willingly turn what they have learned into professional practice (Beach, 2017). On the other hand, SDL is a critical characteristic a learner should possess for better adjustment and success in online learning, and for improving learning outcomes (Bonk, Lee, Kou, Xu, & Sheu, 2015; Heo & Han, 2018, p. 62; Hyland & Kranzow, 2011; Kim et al., 2014; Loizzo, Ertmer, Watson, & Watson, 2017). With interest, curiosity, and desire for self-improvement being among the most important motivating factors, learners are independent and autonomous in their use of various devices and places to learn, and for meeting their self-directed learning needs at their own pace (Bonk et al., 2015; Heo & Han, 2018, p. 62). Therefore, given the increasing opportunities for online learning, an area of particular interest to online learning researchers addresses the learner's ability to guide and direct his or her own learning (Beach, 2017; Hyland & Kranzow, 2011; Song & Hill, 2007, p. 27).

Existing SDL Measures

The measurement of self-direction in learning has been operationalized in studies that develop and validate instruments measuring various aspects of SDL, and, many times, revalidate these instruments again in culturally relevant settings, in different student populations, and so on.

Many SDL instruments are based on Knowles's andragogic theory (Cadorin, Bressan, & Palese, 2017; Knowles, 1975). First, in Guglielmino (1977), the Self-Directed Learning Readiness Scale (SDLRS) was developed based on Knowles's original concept of self-directed learning. Here, SDL readiness refers to the extent to which the individual possesses the abilities, attitudes, and personality characteristics necessary for self-directed learning (Wiley, 1983, p. 182). The SDLRS purported to measure the complex of attitudes, skills, and characteristics comprising an individual's current level of readiness to manage his or her own learning. Next, also adding to the SDL literature is the Self-Rating Scale of Self-Directed Learning (Williamson, 2007) measuring self-directed learning abilities in five dimensions. The instrument was subsequently revalidated in the Italian context to have a reduced number of items measuring SDL in eight dimensions (Cadorin, Bortoluzzi, & Palese, 2013; Cadorin, Suter, Saiani, Williamson, & Palese, 2010).

Besides SDL instruments designed for the general student population, SDL assessment tools have also been developed for students in specific domains. For example, in nursing education, multiple SDL instruments have been constructed measuring students' SDL skills to enhance the quality of their professional practice, including: (a) Self-Directed Learning Instrument (Cheng, Kuo, Lin, & Lee-Hsieh, 2010); (b) Self-Directed Learning Readiness Scale for Nursing Education (Fisher, King, & Tague, 2001); and (c) Autonomous Learner Index (Abu-Moghli, Khalaf, Halabi & Wardam, 2005).

Finally, many more SDL scales have been developed to serve various purposes and student populations, including: (a) Self-Directed Learning Scale (Lounsbury & Gibson, 2006); (b) Self-Directed Learning Inventory, for elementary school and college students (Jung, Lim, Jung, Kim, & Yoon, 2012; Suh, Wang, & Arterberry, 2015); and (c) Oddi Continuing Learning Inventory (Oddi, 1986). For a comprehensive listing of SDL measures, readers should refer to systematic reviews of SDL scale development studies, such as Cadorin et al. (2017) and Sawatsky (2017).

New SDL Instrument

Despite the existence of multiple SDL instruments, the literature review in this study has not identified any such instrument which is designed specifically for the online learning environment and dedicated to students who have had prior online learning experience. First, there are indeed a few SDL items written for the online environment buried somewhere in large scale surveys that measure multiple aspects of online education, such as items measuring student autonomy in the lengthy, 62-item Online Learning Environment Survey (Trinidad, Aldridge, & Fraser, 2005). A long, complicated survey tends to be associated with a low response rate, and when administered, may not collect any responses to the items specific to SDL. Second, among the existing SDL items for online education, many are formulated as prospective, instead of retrospective. Questions surveying students' future opinions when taking an online course may not always, when administered, be answered by students with prior online education experience. Students with no prior online education experience can also respond to those questions by imagining what their experiences would be like if they were to take an online course, and responses from such students are likely to lack validity.

Taking into account the two issues outlined above, as part of a larger study, a new, concise SDL instrument, Self-Directed Online Learning Scale (SDOLS; Su, 2016) has been developed to use retrospective, instead of prospective, questions aimed to collect responses from only students with prior online learning experience. The instrument measures students' SDL ability after he or she has taken an online course; it helps instructional designers determine if an online course meets the needs of students and identifies grounds for improvement. The construction of the SDOLS items was based on brainstorming, referring to existing SDL measures and adapting items from available SDL instruments (Abu-Moghli et al., 2005; Cheng et al., 2010; Fisher et al., 2001; Garrison, 1997, Guglielmino, 1977; Jung et al., 2012; Lounsbury & Gibson, 2006; Oddi, 1986; Suh et al., 2015; Trinidad et al., 2005; Watkins, Leigh, & Triner, 2004; Williamson, 2007). University faculty members with expertise in scale development and instructional design were also consulted to enhance the content validity of the instrument. Although SDOLS was developed based on the responses of graduate students in one research university in the Southeast US, the items are universal enough as an inquiry into the online learning experiences of students in other universities as well.

Research Questions

This study utilized a non-experimental survey research design, based in a post-positivist worldview (Creswell, 2013; Devlin, 2006) to explore graduate students' self-directed online learning ability. The study aimed to assess the psychometric properties of SDOLS, and examine issues related to graduate student perceptions of their SDL ability. Post-positivism holds "a deterministic philosophy in which causes probably determine effects or outcomes" (Creswell, 2013, p. 7). This study was passive in design, as there was no intent to manipulate any variables. The study was also exploratory, as it provided only preliminary psychometric evidence of the instrument and its use in investigating SDL and served as the foundation for examining future application of the instrument to broader contexts.

A Rasch measurement approach was taken using the rating scale model (RSM) to evaluate the psychometric properties of SDOLS (Bond & Fox, 2015). Rasch modeling and its variants have been used in similar research in online education (Choi, Walters, & Hoge, 2017; Wilson, Gochyyev, & Scalise, 2016). Besides

scale validation, the study also examined students' perceptions of their SDL ability. Specifically, the study addressed two research questions:

1. What are the psychometric properties of SDOLS as measured by the Rasch rating scale model?

2. How important do students think various factors, as documented by SDOLS items, are in terms of contributing to their self-directed online learning ability?

Methodology

Research Context

The draft SDOLS instrument was pilot-tested in the fall semester of 2014. A group of 10 graduate students taking an online course in that semester participated in the pilot testing. They were surveyed through Qualtrics after the conclusion of the semester and provided the feedback which was later incorporated into the final survey instrument. Their feedback revolved around identifying any aspects of the draft instrument that could lend themselves to misunderstanding or logical flow problems in the survey delivery and revising such aspects. After factoring in the feedback, the final instrument had 17 items and was administered to another, larger group of students.

Table 1 presents the final SDOLS instrument; each item is a question related to how students take charge of their learning on a 1 to 5 Likert scale: 1 for *strongly disagree* (SD), 2 for *disagree* (D), 3 for *neutral* (N), 4 for *agree* (A), and 5 for *strongly agree* (SA). The 17 items make up two subscales—autonomous learning (AUL; eight items) and asynchronous online learning (AOL; nine items). Finally, all SDOLS items were worded positively; a higher score indicates a higher level of SDL ability.

Table 1

Self-Directed Online Learning Scale

Item	Item statement	Subscale
Q01	I was able to make decisions about my online learning (e.g., selecting online project topics).	AUL
Q02	I worked online during times I found convenient.	AUL
Q03	I was in control of my online learning.	AUL
Q04	I played an important role in my online learning.	AUL
Q05	I approached online learning in my own way.	AUL
Q06	I was able to complete my work even when there were online distractions (e.g., friends sending e-mails).	AUL
Q07	I was able to complete my work even when there were distractions in my home (e.g., children, television).	AUL
Q08	I was able to remain motivated even though the instructor was not online at all times.	AUL
Q09	I was able to access the discussion forum at places convenient to me.	ASL
Q10	I was able to read posted messages at times that were convenient to me.	ASL
Q11	I was able to take time to think about my messages before I posted them.	ASL

Q12	The process of writing and posting messages helped me articulate my thoughts.	ASL
Q13	My writing skills have improved through posting messages.	ASL
Q14	I was able to ask questions and make comments in online writing.	ASL
Q15	I was able to relate the content of online course materials to the information I have read in books.	ASL
Q16	I was able to understand course-related information when it was presented in video formats.	ASL
Q17	I was able to take notes while watching a video on the computer.	ASL

Research Participants

After securing required Institutional Review Board approval, the study proceeded to obtain a nonprobability convenience sample. The sample consisted of all 909 graduate students in the aforementioned university who were taking online courses during the fall semester of 2014. In January 2015, these 909 graduate students were contacted by e-mail through Qualtrics, inviting them to participate in the study.

To address the possible low response rate issue common in online surveys, the study first sent a mass pre-notification e-mail to all 909 students, informing them of an upcoming solicitation to participate in a study about their online learning experiences during the fall semester of 2014. After the data collection started, several follow-up e-mails were sent to remind the students to complete the survey. This continued until the data collection came to an end in April 2015. As an incentive to participate in the survey, all potential participants were entered into a draw to win one of five gift cards valued at $50 each. In the end, 238 participants provided complete responses to all 17 items, which, despite a low response rate of 26.2%, still led to a high student-item ratio of about 14:1, satisfying the criterion that the sample size should be at least six times the number of items for stable results (Mundfrom, Shaw, & Ke, 2005).

Table 2 provides demographics of the sample of 238 participants. The sample consisted of 50 male and 188 female students. Respondents age ranged from 21 years to 51 years (or older), but almost half (45.8%) were under 30 years old. Regarding ethnicity, there were 22 African American students, 15 Asian students, 5 Hispanic/Latino students, 188 White students, and 8 students who identified as being of more than one race. Finally, regarding marital status, the proportion of students who were married was moderately higher than that of students who were not (58.0% for married vs. 42.0% for not married).

Table 2

Demographics of Student Participants

Category	Variable	n	Percent
Gender	Male	50	21.0
	Female	188	79.0
Age	21–25 years	49	20.6
	26–30 years	60	25.2
	31–40 years	58	24.4
	41–50 years	44	18.5
	51 years or older	27	11.3
Race/Ethnicity	Hispanic/Latino	5	2.1
	Asian	15	6.3
	African American	22	9.2
	White	188	79.0
	More than one race	8	3.4
Marital Status	Married	138	58.0
	Not married	100	42.0
Total		238	100.0

Data Analysis

First, the responses of the 238 students were used to compute two sets of descriptive statistics: (a) cumulative response category percentages from all 17 items, and (b) response category percentages for each individual item. In Figure 1 (consisting of subfigures 1a through 1r) the statistics are presented graphically using bar charts (from left to right: SD, D, N, A and SA). Subfigure 1a represents the cumulative percentages of response categories from all 17 items put together. As is seen, as high as 77.7% of the responses were in the *agree* and *strongly agree* categories, indicating the participants tended to endorse item statements. From subfigures 1b through 1r for individual items, the highest bar is always associated with either the SA or the A category, whereas the SD category is always selected least frequently. Therefore, all 17 items elicited similar response patterns and the participants tended to hold a favorable view of the statement for each item.
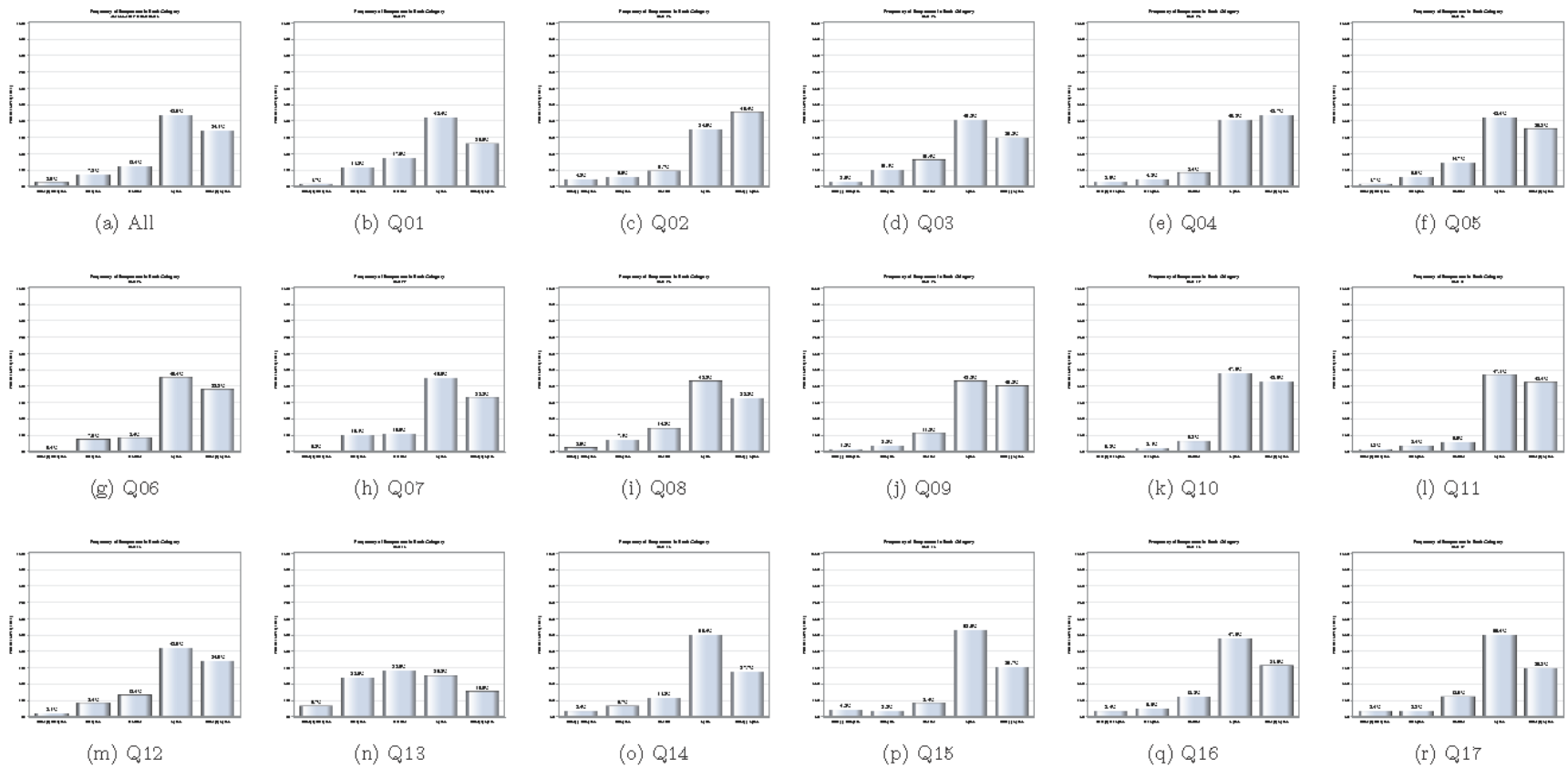
Figure 1. Response frequency distributions for all and individual items.

Next, a unidimensional RSM-based Rasch analysis was conducted within Winsteps 4.1.0 to assess the degree to which students agree with item statements covering various SDL factors (Linacre, 2018). A unidimensional Rasch model assumes the survey items measure only a single underlying construct (e.g., ability to self-manage one's own learning) and establishes the relative difficulty (or relative endorsability) of each item statement with regard to that latent construct (Bond & Fox, 2015).

In the Rasch analysis, several aspects of SDOLS were investigated.

- Unidimensionality: The study investigated the fundamental Rasch analysis assumption of a single underlying construct. The assumption was assessed using a principal component analysis of the correlation matrix of the Rasch standardized residuals.

- Reliability and validity: To identify reliability and validity evidence for SDOLS, person/item reliability and separation statistics were examined. First, person separation classifies people whereas item separation verifies item hierarchy. Next, person and item reliability statistics refer to reproducibility of relative measures or score location (e.g., high reliability of persons (or items) means the probability is high that persons (or items) estimated with high scores actually do have higher scores than persons (or items) estimated with low scores).

- Rating scale effectiveness: The quality of the rating scale was critiqued to see if the response categories functioned as intended, and if students were able to consistently and correctly understand and interpret the response categories.

- Item and person measure quality: Infit and outfit mean square (MNSQ) statistics were used to assess the fit of the 17 items to the Rasch model. According to Linacre (2018, pp. 582-588), the ideal value of an item infit/outfit mean square statistic should fall between 0.50 to 1.50, and values exceeding 2.00 may suggest a noisy problematic item.

- Construct hierarchy: The Wright map was assessed to investigate the construct hierarchy of SDOLS (Wilson, 2005). The map visually presents relative difficulty of items and students' ability to self-manage their learning. From top to bottom, items (to the right of the vertical line) are ranked from the least favorite item (hardest to endorse) to the most favorite item (easiest to endorse); students (to the left of the vertical line) are ranked from those who are most able to self-manage their learning to those who are least able to.

Results

The results support SDOLS as having excellent psychometric properties with the exception of three items. The results also rank-order various factors of SDL regarding their relative importance to students' self-directed online learning ability.

Dimensionality Analysis

Results indicate as high as 51.1% of the raw variance is explained by the Rasch dimension, with 30.9% attributed to persons and 20.2% to items. The largest secondary dimension, indicated by the first contrast under Winsteps, explains only 8.1% of the raw variance with an eigenvalue of 2.8, the strength of at most three items. Therefore, the ratio of the variance explained by items (20.2%) to that explained by the second largest dimension (8.1%) is about 2.50. Despite a possible secondary dimension made up of at best three items, it is also true that virtually all survey datasets consist of multiple dimensions (hardly any dataset is perfectly unidimensional), albeit to varying degrees (Royal & Gonzalez, 2016). Given evidence to support a single, primary underlying construct being measured by the Rasch dimension, the study concludes the unidimensionality assumption is reasonably satisfied for a unidimensional Rasch analysis (Linacre, 2018, pp. 557-558; Royal, Gilliland, & Kernick, 2014).

Reliability and Validity

Person and item separation statistics are, respectively, as high as 2.71 and 4.64. The high person separation statistic indicates SDOLS is sufficiently sensitive to distinguish between individual students with higher and lower levels of SDL ability, and the high item separation statistic suggests the student sample is large enough to confirm item difficulty hierarchy. Overall, these observations support the construct validity of the instrument.

Person reliability is 0.88 (i.e., SDOLS discriminates the sample into enough levels), and item reliability is even higher at 0.96 (i.e., the sample is large enough to precisely locate the items on the latent difficulty continuum). Person reliability being high could be due to ability variance being large. By contrast, item reliability being high could be attributed to large variability in item difficulty and a relatively large number of students.

Rating Scale Effectiveness

First, based on the response category probability curves in Figure 2, each category has a distinctive peak suggesting it is a meaningful endorsement choice for students at a specific ability level. Stated differently, students are able to sufficiently separate one response option from another, thus providing additional evidence of validity.

```
        CATEGORY PROBABILITIES: MODES - Andrich thresholds at intersections
P       -+------+------+------+------+------+------+------+------+-
R   1.0 +                                                         +
O       |
B       |11
A       | 11
B    .8 +   111                                               55+
I       |      1                                            5  |
L       |     11                                          55   |
I       |      1                         4444444        55     |
T    .6 +       1                     44        444    5        +
Y       |      11                   4              4  5         |
     .5 +        1                 4                4*5         +
O       |          *222222     44                 5 4          |
F    .4 +        22 1       22   4               5   44        +
        |       22       1     2 4             55      44       |
R       |      22       1    33**3333    33     5         4     |
E       |     22       1133 4    2    33      55         44    |
S    .2 +     22        331 4      2    33 55         44+
P       | 222         33    *         22    **3       44        |
O       |22          333  44 111      2*55   333   333          |
N       |          33333  4444      11*555 2222    333333       |
S    .0 +**************555555555555 11111111*********************+
E       -+------+------+------+------+------+------+------+------+-
        -4    -3    -2    -1     0     1     2     3     4
        PERSON [MINUS] ITEM MEASURE
```
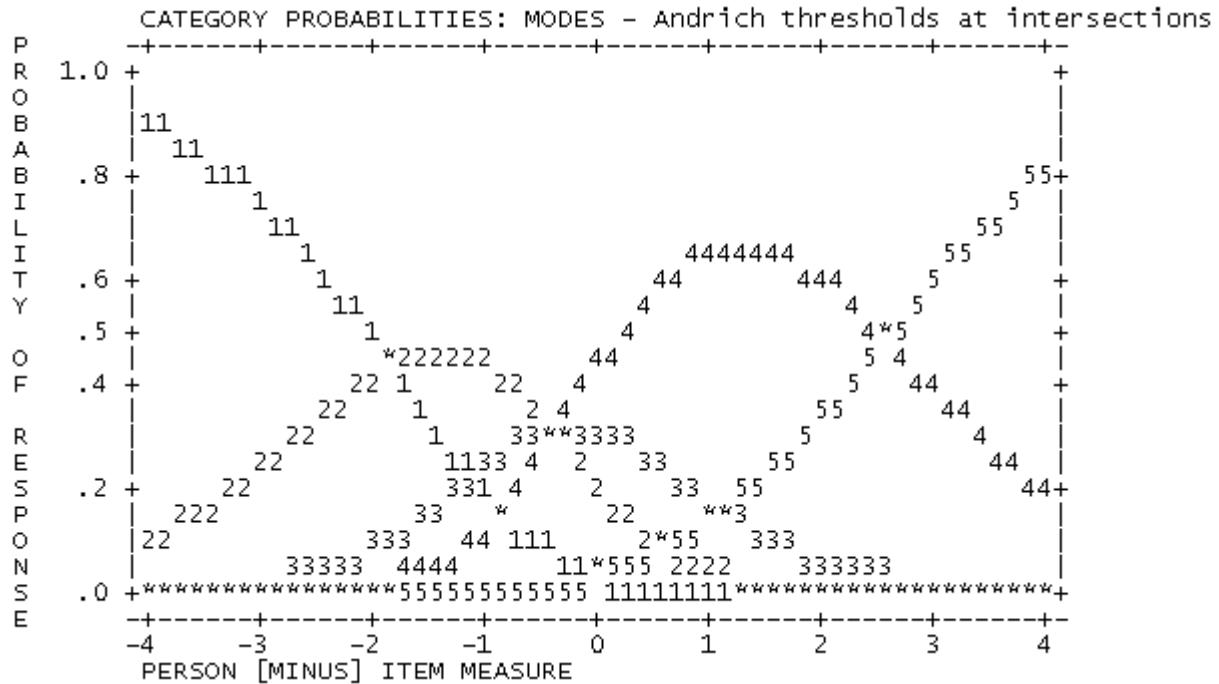
Figure 2. Response category probability curves.

Next, based on the shape of the response category count distribution in Table 3, it is evident that, although students do make full use of all five response categories, they still prefer to select those on the side of agreement (the *agree* category, in particular). Notably, almost all infit and outfit MNSQ statistics fall into the recommended range of 0.50 to 1.50 (Linacre, 2018, pp. 582-588) with only the outfit MNSQ for the SD category being only 0.04 points higher than 1.50. Besides, the category measures and Andrich threshold measures each advance in a stepwise manner, as expected.

Table 3

Category Structure Calibration

Responses		Observed sample		Mean square		Stepwise	
Options	Labels	n	Percentage	Infit	Outfit	Andrich threshold	Category measure
1	SD	104	3	1.14	1.54	None	(-3.10)
2	D	294	7	1.14	1.25	-1.84	-1.40
3	N	502	12	1.11	1.50	-0.43	-0.26
4	A	1,766	44	0.97	0.78	-0.32	1.29
5	SA	1,380	34	0.90	0.90	2.59	(3.73)

Third, the study examined the construct key map for the five response categories (see Figure 3). In the map, items are ordered from the least endorsable item Q13 (top) to most endorsable item Q10 (bottom). Evidently, the ordering of categories remains consistent as 1, 2, 3, 4, and 5 across all 17 items. Such consistency

indicates none of the 17 items might cause misunderstanding or unexpected answers, thus supporting the validity of SDOLS (Ren, Bradley, & Lumpp, 2008). In summary, the results here support that the rating scale structure of SDOLS functioned in the intended way, and the response options were consistently and correctly interpreted by research participants.
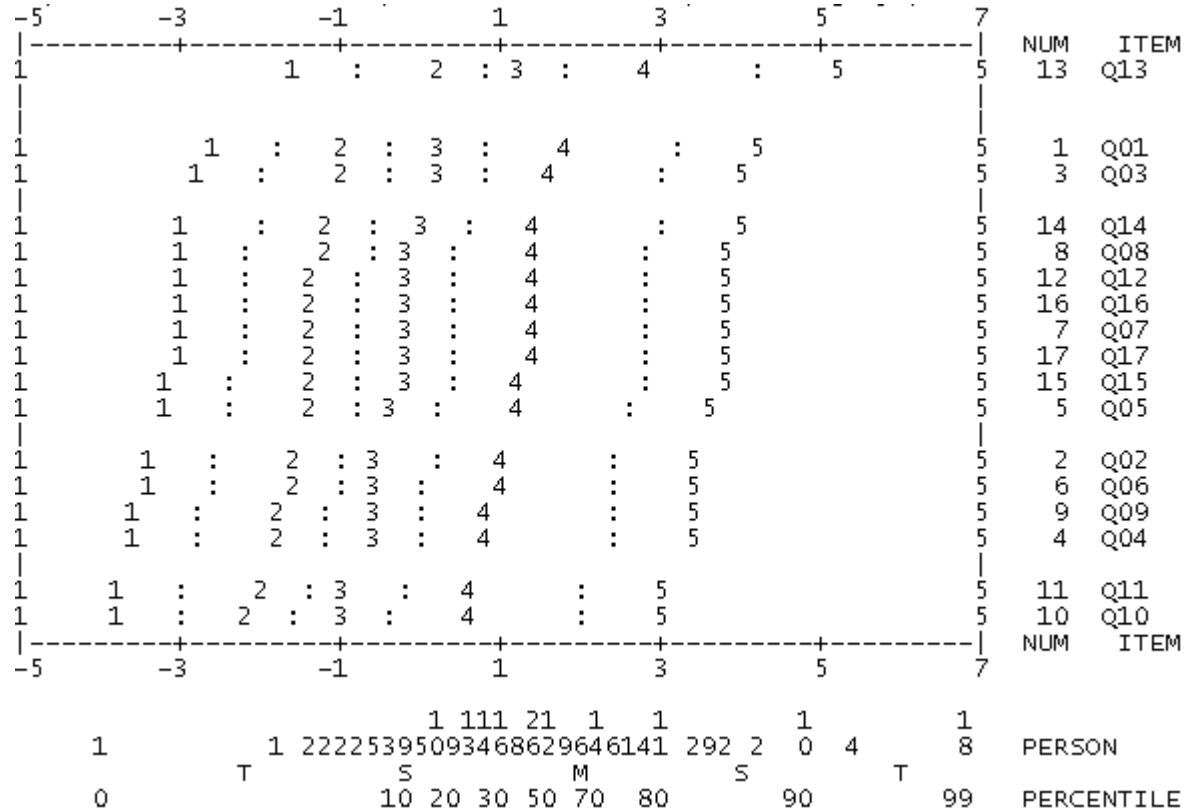
```
-5         -3          -1          1          3          5          7
|--------+---------+---------+---------+---------+---------|  NUM    ITEM
1                1     :     2   : 3    :     4       :      5       5   13   Q13
|                                                            |
1            1   :    2   :  3   :     4         :     5     5    1   Q01
1          1   :    2   :  3   :   4          :    5        5    3   Q03
|                                                            |
1          1   :    2   : 3   :     4         :     5       5   14   Q14
1          1   :    2  : 3   :     4        :      5        5    8   Q08
1          1   :    2  :  3  :      4       :      5        5   12   Q12
1          1   :    2  :  3  :      4       :      5        5   16   Q16
1          1   :    2  :  3  :      4       :      5        5    7   Q07
1          1   :    2  :  3  :       4      :      5        5   17   Q17
1         1   :     2  : 3   :    4         :      5        5   15   Q15
1         1   :     2  : 3   :     4        :     5         5    5   Q05
|                                                            |
1       1   :     2   : 3    :     4      :      5          5    2   Q02
1       1   :     2   : 3    :    4       :     5           5    6   Q06
1      1   :    2   :  3   :   4        :     5             5    9   Q09
1      1   :    2   :  3   :   4        :     5             5    4   Q04
|                                                            |
1       1   :    2  : 3   :     4        :      5           5   11   Q11
1       1   :   2  : 3   :     4       :    5               5   10   Q10
|--------+---------+---------+---------+---------+---------|  NUM    ITEM
-5         -3          -1          1          3          5          7

                    1 111 21  1    1          1          1
         1        1 22225395093468629646141 292 2   0   4    8   PERSON
              T          S           M          S       T
         0            10 20 30 50 70   80         90        99  PERCENTILE
```

Figure 3. Construct key map.

Item Measure Quality

In Table 4, Q13 has an unusually large outfit MNSQ statistic (2.35). Because 2.35 > 2.00, it indicates that, with this item, off-variable noise is greater than useful information. Because this item degrades measurement, it should be revised to remedy the misfit. Besides Q13, Q01 and Q02 have relatively serious misfit issues with inflated infit and outfit MNSQ statistics for each item being greater than 1.50 (ranging from 1.52 to 1.84). These two items may be problematic and thus require further scrutiny to reduce their off-variable noise and improve their fit to the model. Table 4 also indicates all remaining 14 items are productive of measurement, because each item's infit and outfit MNSQ measures fall into the acceptable range of 0.50 to 1.50. Finally, point biserial correlations are all high and positive (ranging from .56 to .72), indicating the orientation of the scoring on each item is consistent with the orientation of the latent variable, and that the items have excellent discriminatory abilities (Linacre, 2018, pp. 526-532).

Table 4

Item Quality Indicators

Item	Total	Measure estimate	Measure SE	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD	Point biserial
Q13	759	1.54	0.08	1.34	3.40	2.35	9.90	.66
Q01	905	0.45	0.09	1.52	4.40	1.84	6.30	.56
Q02	979	-0.25	0.10	1.73	5.40	1.79	5.40	.56
Q17	951	0.04	0.10	1.13	1.20	1.22	1.80	.62
Q04	995	-0.42	0.11	1.17	1.50	1.04	0.30	.64
Q07	951	0.04	0.10	1.13	1.20	1.07	0.70	.66
Q14	934	0.20	0.10	0.98	-0.10	1.11	1.00	.66
Q03	915	0.37	0.09	0.94	-0.60	1.00	0.00	.70
Q05	961	-0.06	0.10	0.95	-0.40	0.96	-0.30	.68
Q09	994	-0.41	0.11	0.89	-0.90	0.83	-1.40	.67
Q06	984	-0.30	0.10	0.85	-1.30	0.87	-1.10	.68
Q16	950	0.05	0.10	0.83	-1.60	0.87	-1.10	.69
Q15	957	-0.02	0.10	0.84	-1.50	0.72	-2.70	.69
Q08	944	0.10	0.10	0.81	-1.90	0.80	-1.90	.71
Q12	946	0.09	0.10	0.81	-1.90	0.77	-2.20	.72
Q11	1014	-0.65	0.11	0.69	-3.10	0.63	-3.30	.69
Q10	1023	-0.76	0.11	0.63	-3.70	0.58	-3.80	.70

Item/Construct Hierarchy

According to the Wright map in Figure 4, students most easily endorse items Q10 and Q11, suggesting students highly value the ability to read posted messages at convenient times and to take time to think about their own messages before posting them. Next, students equally easily endorse items Q04, Q06, and Q09. This indicates students believe discussion forum access at convenient places is a very important factor in online learning. Plus, students believe they take control of their own online learning and are confident of completing their work despite online distractions. Next, students easily endorse Q02, indicating they tend to work online during convenient times. Then, at the average item difficulty level is a group of six items: (a) Q05 (approaching online learning); (b) Q07 (completing work despite home distractions); (c) Q12 (articulating thoughts); (d) Q15 (relating course materials to books); (e) Q16 (understanding course information in video formats); and (f) Q17 (taking notes). Evidently, these are more difficult to endorse than all items already discussed but are easier to endorse than items to be presented next. Next, Q08, Q14, Q01, and Q03 follow closely with virtually identical endorsability measures. Students find it relatively difficult to (a) stay motivated, (b) ask questions and make comments, (c) make decisions, and (d) stay in control in online learning. Finally, the hierarchy continues upward until it reaches the most difficult items to endorse, Q13 staying away from all other items (i.e., there is a large gap between Q13 and all other items in the Wright map), indicating students hardly agree their writing skills have improved through posting messages.

```
MEASURE      PERSON - MAP - ITEM
                    <more>|<rare>
      6     ########  +
                      |
                  ##  |
      5             +
              #####  |
                   T|
                    |
                #   |
      4             +
                #   |
             .####  |
                #   |
                    |
      3      .##### S+
                ##  |
                 .  |
               ###  |
            #######  |
      2      ###    +
             .####   |
            ###### M|    Q13
      #############  |
              ####   |
      1     ######## +T
            #######  |
            .######  |S
              .####   |    Q01      Q03
              #####   |    Q08      Q14
      0        .## S+M Q05      Q07      Q12      Q15      Q16      Q17
              .####   |    Q02
               .#    |    Q04      Q06      Q09
              .##    |S Q11
               #     |    Q10
     -1         #    +T
               #     |
               #     |
                   T|
                .    |
     -2             +
                    |
                    |
                    |
     -3             +
                    |
                    |
                    |
                    |
     -4         .   +
              <less>|<freq>
EACH "#" IS 2: EACH "." IS 1
```
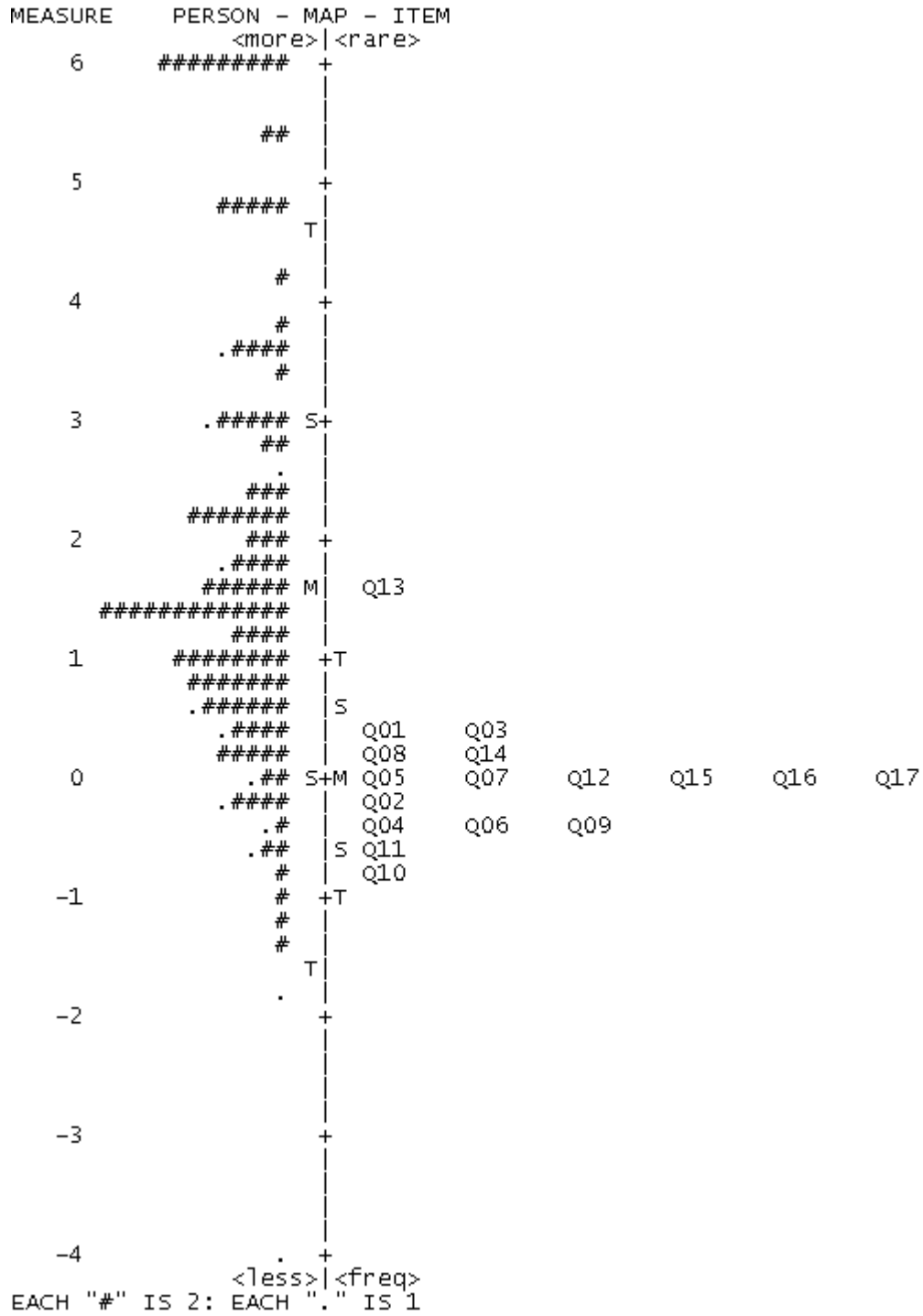
Figure 4. Wright hierarchy map.

Discussion

The study assessed the psychometric properties of SDOLS for measuring students' perceptions of their self-directed online learning ability. Overall, the study supports SDOLS as having decent psychometric properties. Next, the study rank-ordered items regarding students' level of endorsement to offer insights into how important the attributes are for facilitating students' online, self-directed learning.

Psychometric Properties of SDOLS

Regarding psychometric properties of SDOLS, the study was based on the validity framework by Messick (1989) which has been implemented in the Rasch literature (Long, Wendt, & Dunne, 2011, pp. 388–389; Royal & Elahi, 2011, p. 369; Royal et al., 2014, pp. 458–459). According to Messick, validity is the integration of any evidence that impacts the interpretation or meaning of a score. Messick's framework is made up of six unique aspects of validity: (a) substantive, (b) content, (c) generalizability, (d) structural, (e) external, and (f) consequential. The Rasch analysis findings are discussed next within four of these six aspects of validity.

First, the fundamental assumption of unidimensionality is satisfied from a practical perspective due to the existence of a single, primary Rasch dimension which explains over 50% of the raw variance. This finding supports the substantive aspect of validity. Next, measures of reliability are extremely close to or above .90, which serves as evidence of the generalizability aspect of validity. Also, a diagnostic of the rating scale effectiveness indicates the response categories of SDOLS functioned as intended, and the participants were able to correctly and consistently interpret response options, which supports the structural aspect of validity. An assessment of the item fit measures indicated the vast majority of the 17 items provided an adequate fit to the Rasch model; this finding speaks to the content aspect of validity. In summary, multiple pieces of evidence under Messick's validity framework supported SDOLS as being psychometrically sound, thus indicating the instrument is able to produce high-quality data.

Next, the analysis of item misfit reveals three items (Q13, Q01, and Q02, presented in order of misfit per Table 4) that did not provide adequate fit to the Rasch model. These items should be either removed or revised in future iterations of SDOLS.

Implications for Online Teaching and Learning

The SDOLS instrument addresses many issues associated with students' ability to self-manage their learning in online education. Because SDOLS offers insights into online students' feelings regarding various aspects of their SDL ability, the instrument is likely to be relevant to various stakeholders in online education including students, instructors, administrators, instructional designers, researchers, and so on. For example, instructional designers may use the SDOLS data to identify grounds for improvements to an online learning environment, and as a guide in their work to improve their designs. In another instance, they may use the instrument as a diagnostic tool to measure online learners' readiness, and screen for learners whose self-directed learning ability is likely to be weak, before tailoring course designs in a way that improves online learners' success. On the other hand, data collected through the SDOLS instrument will enable instructors, administrators, and researchers to better understand how students' self-directed learning characteristics may relate to their success in online courses and completion of online programs,

thus effectively contributing to improving online course and program designs. In summary, the study recommends the SDOLS instrument should be used for improving student online learning experience and effectiveness, increasing online retention rates, and reducing online dropouts.

Limitations and Future Research

The study is not without limitations, but limitations could be directions of future research. First, the research data could have been subject to self-selection bias, due to the self-selected sample being non-probabilistic and therefore lacking in randomization, and to nonresponse bias exacerbated by a low response rate of 26.2%. Graduate students who chose to complete the online survey could be different demographically and behaviorally than those who chose not to. Second, the study has not assessed SDOLS on two other aspects of validity under Messick's framework. On one hand, because the study is the first introducing and validating SDOLS, there is no way to investigate the consequential aspect of validity, since the instrument was not previously used. On the other hand, findings from the study have not been correlated with those from others, so the external aspect of validity has not been evaluated. Third, the study is yet to examine the extent to which items remain invariant across various subgroups (e.g., by gender). In future research, a differential item functioning analysis could further assess whether SDOLS items function differently across these subpopulations. Finally, given the limitations described above, although the findings here support the scale as having potential, they are still preliminary regarding the issues in the two research questions. Fortunately, the research design and the analytic methodology are straightforward to implement, which makes it easier for future researchers to replicate the study under broader research contexts.

Conclusion

The study develops and validates SDOLS measuring students' ability to self-manage their online learning with a secondary goal of understanding their perceptions of various SDL factors. First, the study finds validity evidence for SDOLS from multiple perspectives under the Messick framework as well as evidence supporting SDOLS as a reliable instrument. The study also finds three problematic items (Q13, Q01, and Q02) based on criteria from the Rasch literature and suggests they should be revised or removed. Second, the study provides insights into students' perceptions of various SDL factors regarding contributions to their SDL ability.

As a final reflection, SDOLS is designed to survey students with prior online learning experience regarding their perceptions of SDL ability under the unique nature and features of the online education environment. The preliminary results here indicate SDOLS can be administered with confidence to students for a reliable and valid measurement of their SDL ability. Because these characteristics of students ultimately determine whether self-directed learning will take place, the instrument is expected to help researchers better understand students' self-directedness in learning within the online environment, which in turn will contribute to the call for adequate social and academic support to enhance students' online learning experience and to reduce the rate of attrition. Besides, despite limited coverage in this study, SDOLS can be used for diagnostic purposes by analyzing the Wright map to identify, characterize, and rank-order learners regarding their level of self-directedness in learning (i.e., distinguishing students who are more

independent learners good at determining their learning needs, planning, and implementing their own learning from students who feel more comfortable with more structured learning options such as traditional classroom environments). This diagnostic use of SDOLS is valuable because, until very recently, there have been few validated tools for identifying the self-directed learners (Sahoo, 2016, p. 167).

References

Abu-Moghli, F. A., Khalaf, I. A., Halabi, J. O., & Wardam, L. A. (2005): Jordanian baccalaureate nursing students' perception of their learning styles. *International Nursing Review, 52*(1), 39–45. https://doi.org/10.1111/j.1466-7657.2004.00235.x

Beach, P. (2017). Self-directed online learning: A theoretical model for understanding elementary teachers' online learning experiences. *Teaching and Teacher Education, 61*, 60–72. https://doi.org/10.1016/j.tate.2016.10.007

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model* (3rd. ed.). New York, NY: Routledge.

Bonk, C. J., Lee, M. M., Kou, X., Xu, S., & Sheu, F. R. (2015). Understanding the self-directed online learning preferences, goals, achievements, and challenges of MIT OpenCourseWare subscribers. *Journal of Educational Technology & Society, 18*(2), 349–368. Retrieved from https://drive.google.com/file/d/1Rrj6XQ5uTWpVR2UH2EvLWI1xvEFAnhcB/view

Cadorin, L. Bortoluzzi, G., & Palese, A. (2013). The self-rating scale of self-directed learning (SRSSDL): A factor analysis of the Italian version. *Nurse Education Today, 33*(12), 1511–1516. https://doi.org/10.1016/j.nedt.2013.04.010

Cadorin, L., Bressan, V., & Palese, A. (2017). Instruments evaluating the self-directed learning abilities among nursing students and nurses: A systematic review of psychometric properties. *BMC Medical Education, 17.* https://doi.org/10.1186/s12909-017-1072-3

Cadorin, L., Suter, N., Saiani, L. Williamson, S. N., & Palese, A. (2010). Self-rating scale of self-directed learning (SRSSDL): Preliminary results from the Italian validation process. *Journal of Research in Nursing, 16*(4), 363–373. https://doi.org/10.1177/1744987110379790

Caffarella, R. (1993). Self-directed learning. *New Directions for Adult and Continuing Education, 57,* 25–35. https://doi.org/10.1002/ace.36719935705

Cheng, S. F., Kuo, C. L., Lin, K. C., & Lee-Hsieh, J. (2010). Development and preliminary testing of a self-rating instrument to measure self-directed learning ability of nursing students. *International Journal of Nursing Studies, 47*(9), 1152–1158. https://doi.org/10.1016/j.ijnurstu.2010.02.002

Choi, J., Walters, A., & Hoge, P. (2017). Self-reflection and math performance in an online learning environment. *Online Learning Journal, 21*(4), 79–102. http://dx.doi.org/10.24059/olj.v21i4.1249

Chou, P.-N., & Chen, W.-F. (2008). Exploratory study of the relationship between self-directed learning and academic performance in a web-based learning environment. *Online Journal of Distance Learning Administration, 11*(1), 1–11. Retrieved from https://www.westga.edu/~distance/ojdla/spring111/chou111.html

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: Sage.

Devlin, A. S. (2006). *Research methods: Planning, conducting and presenting research*. Belmont, CA: Wadsworth/Thomson Learning.

Fisher, M., King, J., & Tague, G. (2001). Development of a self-directed learning readiness scale for nursing education. *Nurse Education Today, 21*(7), 516–525. https://doi.org/10.1054/nedt.2001.0589

Garrison, D. R. (1997). Self-directed learning: Toward a comprehensive model. *Adult Education Quarterly, 48*(1), 18–33. https://doi.org/10.1177/074171369704800103

Guglielmino, L. M. (1977*). Development of the self-directed learning readiness scale* (Doctoral dissertation). *Dissertation Abstracts International*, *38*, 6467A.

Heo, J., & Han, S. (2018). Effects of motivation, academic stress and age in predicting self-directed learning readiness (SDLR): Focused on online college students. *Education and Information Technologies, 23*(1), 61–71. https://doi.org/10.1007/s10639-017-9585-2

Hiemstra, R. (1994). Self-directed learning. In T. Husen & T. N. Postlethwaite (Eds.), *The International Encyclopedia of Education* (2nd ed.). Oxford: Pergamon Press.

Hone, K. S., & Said, G. R. (2016). Exploring the factors affecting MOOC retention: A survey study. *Computers & Education, 98*, 157–168. https://doi.org/10.1016/j.compedu.2016.03.016

Hyland, N., & Kranzow, J. (2011). Faculty and student views of using digital tools to enhance self-directed learning and critical thinking. *International Journal of Self-Directed Learning, 8*(2), 11–27. Retrieved from https://6c02e432-3b93-4c90-8218-8b8267d6b37b.filesusr.com/ugd/dfdeaf_3ddb1e07e0984947b0db3e39a2be19a1.pdf

Jung, O. B., Lim, J. H., Jung, S. H., Kim, L. G., & Yoon, J. E. (2012). The development and validation of a self-directed learning inventory for elementary school students. *The Korean Journal of Human Development, 19*(4), 227–245.

Kim, R., Olfman, L., Ryan, T., & Eryilmaz, E. (2014). Leveraging a personalized system to improve self-directed learning in online educational environments. *Computers & Education, 70,* 150–160. https://doi.org/10.1016/j.compedu.2013.08.006

Knowles, M. S. (1975). *Self-directed learning: A guide for learners and teachers*. Englewood Cliffs, NJ: Prentice Hall/Cambridge.

Linacre, J. M. (2018). *A user's guide to Winsteps® Ministeps Rasch model measurement computer program user's guide*. Beaverton, OR: Winsteps.com. Retrieved from https://www.winsteps.com/a/Winsteps-Manual.pdf

Loizzo, J., Ertmer, P. A., Watson, W. R., & Watson, S. L. (2017). Adult MOOC learners as self-directed: Perceptions of motivation, success, and completion. *Online Learning, 21*(2). http://dx.doi.org/10.24059/olj.v21i2.889

Long, C., Wendt, H., & Dunne, T. (2011). Applying Rasch measurement in mathematics education research: Steps towards a triangulated investigation into proficiency in the multiplicative conceptual field. *Educational Research and Evaluation, 17*(5), 387–407. https://doi.org/10.1080/13803611.2011.632661

Long, H. B. (1977). [Review of the book *Self-directed learning: A guide for learners and teachers*, by M. S. Knowles.]. *Group & Organization Studies, 2*(2), 256–257. https://doi.org/10.1177/105960117700200220

Lounsbury, J. W., & Gibson, L.W. (2006). *Personal style inventory: A personality measurement system for work and school settings*. Knoxville, TN: Resource Associates.

Mayes, R., Luebeck, J., Ku, H. Y., Akarasriworn, C., & Korkmaz, Ö. (2011). Themes and strategies for transformative online instruction: A review of literature and practice. *Quarterly Review of Distance Education, 12*(3), 151–166.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 13–103). New York, NY: Macmillan.

Mundfrom, D. J., Shaw, D. G., & Ke, T. L. (2005). Minimum sample size recommendations for conducting factor analyses. *International Journal of Testing, 5*(2), 159–168. https://doi.org/10.1207/s15327574ijt0502_4

Oddi, L. F. (1986). Development and validation of an instrument to identify self-directed continuing learners. *Adult Education Quarterly, 36,* 97–107. https://doi.org/10.1177/0001848186036002004

Palloff, R. M., & Pratt, K. (2007). *Building online learning communities: Effective strategies for the virtual classroom* (2nd ed.). San Francisco, CA: Wiley & Sons.

Prior, D. D., Mazanov, J., Meacheam, D., Heaslip, G., & Hanson, J. (2016). Attitude, digital literacy and self efficacy: Flow-on effects for online learning behavior. *The Internet and Higher Education, 29*, 91–97. https://doi.org/10.1016/j.iheduc.2016.01.001

Ren, W., Bradley, K. D., & Lumpp, J. K. (2008). Applying the Rasch model to evaluate an implementation of the Kentucky Electronics Education project. *Journal of Science Education and Technology, 17*(6), 618-625. https://doi.org/10.1007/s10956-008-9132-4

Rovai, A. P., Ponton, M. K., Wighting, M. J., & Baker, J. D. (2007). A comparative analysis of student motivation in traditional classroom and e-learning courses. *International Journal on E-Learning, 6*(3), 413–432.

Royal, K. D., & Elahi, F. (2011). Psychometric properties of the Death Anxiety Scale (DAS) among terminally ill cancer patients. *Journal of Psychosocial Oncology, 29*(4), 359–371. http://www.tandfonline.com/doi/abs/10.1080/07347332.2011.582639

Royal, K. D., Gilliland, K. O., & Kernick, E. T. (2014). Using Rasch measurement to score, evaluate, and improve examinations in an anatomy course. *Anatomical Science Education, 7*(6), 450–460. https://doi.org/10.1002/ase.1436

Royal, K. D., & Gonzalez, L. M. (2016). An evaluation of the psychometric properties of an advising survey for medical and professional program students. *Journal of Educational and Developmental Psychology, 6*(1), 195–203. https://doi.org/10.5539/jedp.v6n1p195

Sahoo, S. (2016). Finding self-directed learning readiness and fostering self-directed learning through weekly assessment of self-directed learning topics during undergraduate clinical training in ophthalmology. *International Journal of Applied and Basic Medical Research, 6*(3), 166–169. https://doi.org/10.4103/2229-516X.186959

Sawatsky, A. (2017). *Instruments for measuring self-directed learning and self-regulated learning in health professions education: A systematic review*. Paper presented at the 2017 annual meeting of the Society of Directors of Research in Medical Education (SDRME), Minneapolis, MN.

Song, L., & Hill, J. R. (2007). A conceptual model for understanding self-directed learning in online environments. *Journal of Interactive Online Learning, 6*(1), 27–42.

Su, J. (2016). *Successful graduate students' perceptions of characteristics of online learning environments* (Unpublished doctoral dissertation). The University of Tennessee, Knoxville, TN.

Suh, H. N., Wang, K. T., & Arterberry, B. J. (2015). Development and initial validation of the Self-Directed Learning Inventory with Korean college students. *Journal of Psychoeducational Assessment, 33*(7), 687–697. https://doi.org/10.1177/0734282914557728

Trinidad, S., Aldridge, J., & Fraser, B. (2005). Development, validation and use of the Online Learning Environment Survey. *Australasian Journal of Educational Technology, 21*(1), 60–81. https://doi.org/10.14742/ajet.1343

Watkins, R., Leigh, D., & Triner, D. (2004). Assessing readiness for E-learning. *Performance Improvement Quarterly, 17*(4), 66–79.

Wiley, K. (1983). Effects of a self-directed learning project and preference for structure on self-directed learning readiness. *Nursing Research 32*(3), 181–185. Retrieved from https://journals.lww.com/nursingresearchonline/Abstract/1983/05000/Effects_of_a_Self_Directed_Learning_Project_and.11.aspx

Williamson, S. N. (2007). Development of a self-rating scale of self-directed learning. *Nurse Researcher, 14*(2), 66–83. https://doi.org/10.7748/nr2007.01.14.2.66.c6022

Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: Erlbaum.

Wilson, M., Gochyyev, P., & Scalise, K. (2016). Assessment of learning in digital interactive social networks: A learning analytics approach. *Online Learning Journal, 20*(2), 97–119. Retrieved from http://files.eric.ed.gov/fulltext/EJ1105928.pdf

Athabasca University