

Developing a Survey to Explore of Sense of Belongingness related to Language Learning Using Rasch Measurement Theory

Desarrollando un cuestionario exploratorio de la sensación de pertenencia relacionado al aprendizaje de idiomas utilizando la teoría de medición educacional de Rasch

Kris Aric Knisely¹ & Stefanie A. Wind²

¹The University of South Dakota

²The University of Alabama

Abstract

This study explored the psychometric properties of a survey designed to measure the construct of belongingness based on a pilot administration of a new instrument. This exploration used Rasch measurement theory to determine the degree to which the instrument could be used to examine this construct within the context of language learning and instruction, and to inform revisions to the instrument prior to its use in future research. Pilot data for the new survey were collected from 249 undergraduate students enrolled at four universities in the Southeastern United States and displayed overall good fit to the Rasch model. Overall, the results suggested that the students reported complex perceptions of their own belonging, indicating neither isolation nor a close community in learning French. Students reported a general lack of pride or shame with complex disclosure patterns. Implications for theory, research, and practice suggest a need for further exploration of the various aspects of sense of belongingness, for the active cultivation of community, as well as for advocacy work.

Keywords: Belongingness Hypothesis; Rasch Measurement Theory; Language Instruction; Language Learning

Post to:

Kris Aric Knisely
Modern Languages and Linguistics
The University of South Dakota
Slagle 306D
414 East Clark Street
Vermillion, South Dakota, 57069
krisknisely@gmail.com

© 2017 PEL, <http://www.pensamientoeducativo.org> - <http://www.pel.cl>

ISSN:0719-0409 DDI:203.262, Santiago, Chile
doi: 10.7764/PEL.54.2.2017.2

Resumen

El presente estudio exploró las propiedades psicométricas de una encuesta diseñada para medir la construcción de pertenencia basada en la administración piloto de un nuevo instrumento. Esta exploración del instrumento utilizó la teoría de medición de Rasch para determinar el grado de usabilidad del instrumento para estudiar el constructo mencionado dentro del contexto de aprendizaje y enseñanza del lenguaje, como también para informar de las revisiones antes de su uso en futuras investigaciones. Los datos del piloto para la nueva encuesta fueron reunidos de 249 estudiantes de pregrado inscritos en cuatro universidades del Sudeste de Estados Unidos y mostraron un buen ajuste al modelo de Rasch. En general, los resultados sugieren que los estudiantes reportan percepciones complejas de su propia pertenencia, indicando que no se sienten aislados ni poseen una comunidad cercana en su aprendizaje del idioma francés. Los estudiantes mostraron, en general, una falta de orgullo o vergüenza en patrones de información complejos. Las implicancias de la teoría, investigación y práctica sugieren una mayor exploración de la construcción de pertenencia en diversos aspectos, para el cultivo activo de la comunidad como para el trabajo de promoción.

Palabras clave: Hipótesis de pertenencia; Teoría de medición de Rasch; enseñanza de idiomas; aprendizaje de idiomas

Language learning involves identity development and negotiation (Huang, 2011; Kramsch, 2009). A key aspect of identity is a sense of belonging: An identity is not an abstract idea or label, such as a title, an ethnic category, or a personality trait. It is a lived experience of belonging (or not belonging). A strong identity involves deep connections with others through shared histories and experiences, reciprocity, affection, and mutual commitments. (Wenger, 2000, p. 239)

Theories of language learning contend that fully investing in language learning entails student investment in their own identities (Castillo & Zaragoza, 2011), and in becoming a member of a particular group. Researchers have argued that students who accept the linguistic and cultural ambiguity necessary for language learning leave behind a former sense of self (Rosello, 1997). Although labels and titles are important for identifying groups and signaling identities, they alone do not constitute identities in and of themselves. Lived identities are understood in this study as experiences of belonging. This paper uses the theoretical lens of the belongingness hypothesis (Baumeister & Leary, 1995), for which the existing measures are limited, to consider the results from a pilot administration of a new measure of belongingness within the context of language learning.

Purpose

The purpose of this study is to explore and evaluate the psychometric properties of a new survey intended to measure the construct of belongingness (Baumeister & Leary, 1995) in terms of fundamental measurement properties, including item and person ordering and measurement precision (i.e., model-data fit, reliability, and item-person targeting). Specifically, these procedures are intended to provide insight into the degree to which the Student Perceptions of Belongingness Survey can be used in future research to measure students' sense of belongingness in the context of language learning, as well as to inform revisions to the instrument prior to its use in future research. This paper focuses on student self-reported perceptions of belongingness collected via the survey instrument rather than upon any absolute truths.

Research Questions

A single overarching research question guides the analyses in this study: To what extent does the Student Perceptions of Belongingness Survey demonstrate adherence to fundamental measurement properties? In order to consider the degree to which Student Perceptions of Belongingness Survey can provide insight into students' self-reports of their sense of belongingness in the context of language learning, and in order to inform revisions to the instrument prior to future administrations, we consider the psychometric properties of the survey items within a measurement framework that is based on fundamental measurement properties. Evidence of adherence to model expectations provides support for the interpretation and use of the instrument. Likewise, departures from model expectations can inform revisions to the instrument.

Using Rasch measurement theory as the guiding measurement framework, the analyses were organized around two sub-questions: (1) how are the students and items ordered on the construct? (2) how precise are the estimates of student and item locations on the construct?

Literature Review

This paper explores the measurement of the articulation of identities in social relationships in the context of second language learning and is grounded in the belongingness hypothesis as articulated by Baumeister & Leary (1995). In this section, the sense of belongingness construct is described. Then this hypothesis is situated within the context of research on language learning.

According to the belongingness hypothesis, "the need to belong is a powerful, fundamental, and extremely pervasive motivation," (Baumeister & Leary, 1995, p. 497). As a part of this need to belong, individuals regularly, freely, and effortlessly form attachments to others and avoid threats to existing social attachments. Further, existing relationships tend to be maintained because enduring, positive, and significant interpersonal relationships are more satisfactory than those that are not enduring, and more satisfactory than interactions with a regularly changing group of individuals. Moreover, a "lack of attachments" corresponds to "a variety of ill effects on health, adjustment, and well-being," (Baumeister & Leary, 1995, p. 497). In turn, the importance of attachments and the role of approval in the creation and maintenance of social bonds carry significant explanatory weight for approval-seeking behavior (Baumeister & Leary, 1995). In fact, much of human behavior, emotions, and thoughts are theorized to be guided by the need to belong (Baumeister & Leary, 1995).

To elaborate, as a fundamental motivation, the need to belong is hypothesized to (a) produce effects readily under all but adverse conditions, (b) have affective consequences, (c) direct cognitive processing, (d) lead to ill effects (such as health or adjustment) when thwarted, (e) elicit goal-oriented behavior designed to satisfy it (subject to motivational patterns such as object substitutability and satiation), (f) be universal in the sense of applying to all people, (g) not be derivative of other motives, (h) affect a broad variety of behaviors, and (i) have implications that go beyond immediate psychological functioning (Baumeister & Leary, 1995, p. 498).

A number of researchers have presented similar hypotheses; however, this work is less centralized around the frequent interaction and persistent caring that characterize the belongingness hypothesis. Examples of such work include the love and belonging needs of Maslow's motivational hierarchy, which are presented only as second to basic needs (Maslow, 1943; 1968). The belongingness hypothesis can be understood to encompass both Maslow's love needs and belonging needs, as well as the next-most-important category of esteem needs. As such, the belongingness hypothesis permits the consideration of social needs across Maslow's hierarchy. Other closely related theories include attachment theory (Bowlby 1969, 1973) and similar postulates (Horney, 1945; Sullivan, 1953; Fromm, 1955, 1956; de Rivera, 1984; Hogan, 1983, Epstein, 1992; Ryan, 1991; Guisinger & Blatt, 1994).

Although students' sense of belongingness, as defined by the belongingness hypothesis, has been thoroughly explored in the broader field of educational studies, this construct has rarely been considered in the domain of language learning motivation. However, the social aspects of language learning have received much consideration (cf. Masgoret & Gardner, 2003), particularly in relation to achievement. Connecting these studies, achievement and the belongingness hypothesis have been linked by Baumeister and Leary (1995), who stated that "people prefer achievements that are validated, recognized, and valued by other people over solitary achievements," (p. 498). The belongingness hypothesis can further be connected to questions of enrollment in language classes; the activities one engages in and are intimately related to the social groups with which one seeks to identify. In accord with Baumeister and Leary's (1995) hypothesis, this study is based on the assumption that the need for belonging is a fundamental motivation. Accordingly, it is necessary to elucidate how the construct can be measured, and to consider the results in terms of their implications for language learning and instruction.

Analytic Approach for Measuring Student Perceptions of Belongingness

In order to explore the psychometric properties of the Student Perceptions of Belongingness Survey, we use a model based on Rasch measurement theory. This analytic approach was selected for several reasons. First, models based on Rasch measurement theory provide a method for empirically investigating the psychometric properties of an instrument in terms of strict underlying measurement requirements related to invariant measurement (Engelhard, 2013). Specifically, when the requirements for invariant measurement are met, item locations on the construct (i.e., item difficulty) can be interpreted independently from student locations (i.e., the "strength" or level of students' sense of belongingness), and vice versa. In the context of the Student Perceptions of Belongingness Survey, evidence of adequate fit to the Rasch model provides support for the interpretation of individual students and items on a common scale that reflects the construct. Accordingly, the items can be considered in terms of differences in their relative difficulty to endorse, and students can be identified who reflect varying levels of a sense of belongingness. Alternatively, evidence that particular items or students do not adhere to the expectations of the Rasch model can be used to identify areas for revision prior to subsequent administrations of the survey.

Because of these invariance properties, estimates from the Rasch model can be used to create a visual display called a variable map (i.e., Wright Map; Wilson, 2011) that illustrates the locations of items and persons on a common linear continuum and facilitates the interpretation and communication of results from Rasch analyses. Because of these useful characteristics, methods based on Rasch measurement theory are frequently used to guide the development and revision of measurement instruments in the social sciences (Engelhard, 2013; Wilson, 2005; Wolfe & Smith, 2007). For a didactic introduction to Rasch measurement theory, please consult Bond and Fox (2015). A detailed discussion of using Rasch approaches to explore affective variables in the context of language learning can be found in Knisely and Wind (2015).

Methods

Instrument

Pilot data for the Student Perceptions of Belongingness Survey were collected using a nine-item instrument designed by one of the researchers as part of a larger study (Knisely, 2015) to capture information about undergraduate student perceptions of their own sense of belongingness in the context of language learning (see Appendix A for the items). The instrument aims to elicit student self-reports of their sense of belongingness rather than any absolute truths. Specifically, the instrument was designed to address students' sense of belongingness in terms of their bonds with other French language learners (Items 3 and 5), sense of being a part of a community of language learners (Item 9), patterns of disclosure (Items 2, 4, and 7), and feelings of pride and shame (Items 1, 6, and 8) with respect to French language study. The survey items are third-person statements about belongingness on which students are asked to rate their level of agreement on a Likert-type scale from 1 (*strongly disagree*) to 6 (*strongly agree*).

Participants

Pilot data for the Student Perceptions of Belongingness Survey were collected from a sample of 294 undergraduate students. The sample included 174 female and 120 male undergraduate students enrolled at four different universities in the Southeastern United States. This survey was created as a part of a larger mixed-methods project examining language attitudes and motivation within the context of French as a second language (Knisely, in press; Knisely, 2016; Knisely & Wind, 2015) via several survey measures, focus groups, and interviews. Thus, at the time of survey administration, 200 of the participants were presently taking French or had taken French in the past.

Data Preparation

Prior to analysis, it was necessary to re-code three items in order to achieve a consistent interpretation of the orientation of the rating scale in terms of students' reported sense of belongingness. As presented to the students in the pilot administration, Item 2, Item 4, and Item 6 had a negative orientation--such that higher ratings on these items reflected a weaker sense of belongingness, and lower ratings reflected a stronger sense of belongingness. The opposite orientation was used for the remaining items (higher ratings indicate stronger sense of belongingness). In order to achieve a consistent alignment between the rating scale categories and the construct, these items were reverse-

coded before any analyses were conducted, such that higher ratings reflected a stronger sense of belongingness.

After the negatively oriented items were recoded, a preliminary analysis of the survey items using the Rasch model identified two main areas for improvement prior to further analysis. First, initial results indicated issues related to the rating scale category functioning. Specifically, preliminary analyses indicated that the original six-category rating scale included more categories than were used by the respondents. Specifically, preliminary analyses revealed infrequent use of categories 2, 3, and 5 (< 10%), and non-modal rating scale category probability curves for these three categories. Further, examination of the average person measure observed for category 2 revealed that this category did not increase monotonically along the logit scale. Based on these findings and recommendations from Wright and Linacre (1992) and Linacre (2002), the six-category scale was collapsed to three categories (Original categories 1 and 2 = New category 1; Original categories 3 and 4 = New category 2; Original categories 5 and 6 = New category 3). After recoding, results indicated three distinct, monotonic rating scale categories and improved model-data fit, based on item and person fit statistics.

Second, preliminary results both before and after collapsing the rating scale categories indicated severe misfit for Item 3 (reverse-coded): *I feel like I would be the only person I know who took French*, based on both unstandardized and standardized Rasch fit statistics. Accordingly, Item 3 was not included in the final analysis. Further investigation related to this item may reveal potential substantive explanations for the poor model-data fit.

Data Analysis

A polytomous formulation of the Rasch model (Rasch, 1960/1980) was employed to examine the psychometric properties of the Student Perceptions of Belongingness Survey using the Facets computer program (Linacre, 2015). Specifically, we used a Rating Scale (RS) formulation of the Rasch model (Andrich, 1978) order to obtain estimates of student and item locations on a common linear scale that represents the construct. We selected the RS formulation, rather than the Partial Credit (PC) formulation (Masters, 1982) based on preliminary analyses with both models, which revealed a comparable rating scale category structure across items based on the PC model. Accordingly we selected the RS model, which is more parsimonious and leads to a simpler interpretation of item difficulties. The RS model can be stated mathematically as follows:

$$\ln \left[\frac{P_{ni(x=k)}}{P_{ni(x=k-1)}} \right] = \theta_n - \beta_i - \tau \quad (1)$$

where

$P_{ni(x=k)} / P_{ni(x=k-1)}$ = probability that student n assigns a rating in category k , rather than category $k - 1$ on item i .

θ_n = location of student n on the construct,

β_i = difficulty of item i , and

τ_k = threshold between categories $k-1$ and k .

This model was used to obtain evidence of the psychometric properties of the Student Perceptions of Belongingness Survey in terms of two major categories of indices: (1) location estimates; and (2) indicators of measurement precision. These indices reflect the ordering of items and students on the construct of belongingness, and the quality of these estimates, respectively.

Location Estimates

The first category of psychometric properties of interest in this study was location estimates for students, items, and rating scale categories. In the context of the Student Perceptions of Belongingness Survey, student location estimates (θ) reflect the degree to which each student endorses (i.e., agrees with) the items, item location estimates (β) reflect the relative difficulty to endorse each of the survey items, and rating scale category location estimates (τ ; i.e., thresholds) reflect the difficulty associated with providing a rating in a particular category. Location estimates for students, items, and thresholds are reported on a common linear scale based on log-odds units (i.e., logits) that reflects the construct of belongingness. For students, the estimates were oriented positively, such that higher locations suggest that students readily endorsed the survey items, indicating a relatively strong sense of belongingness, whereas lower locations suggest that the student reported a relatively less strong sense of belongingness. Items were oriented negatively, such that higher locations suggest that an item is relatively difficult to endorse (i.e., difficult to agree with), and lower locations suggest that an item is relatively easier to endorse.

Rating scale category thresholds were oriented in the same manner as items, where higher locations suggest that a stronger sense of belongingness is required to provide a rating in a particular rating scale category, and lower locations suggest that a less-strong sense of belongingness is required to provide a rating in a particular category. When the rating scale formulation of the Rasch model is applied, threshold estimates are calculated for the number of rating scale categories minus one. Because the collapsed rating scale for the sense of belongingness survey included three categories, two thresholds were calculated for each item: τ_1 reflects the difficulty associated with providing a rating in category 2, rather than category 1; τ_2 reflects the difficulty associated with providing a rating in category 3, rather than category 2.

When adequate model-data fit is observed (discussed below), these estimates can be used to compare individual students and items to other students and items in terms of belongingness. Further, the location estimates can be visually compared using a variable map, which is a graphical display that summarizes the overall calibration of students, items, and rating scale categories on the logit scale.

Indicators of Measurement Precision

In addition to logit-scale calibrations, the Rasch model can be used to obtain additional diagnostic statistics that describe the precision within which items and students have been calibrated on the logit scale. In this study, four categories of measurement precision are considered: (A) standard errors for student and item locations; (B) model-data fit for students and items; (C) targeting between student and item locations; and (D) reliability statistics.

A. Standard Errors. Alongside estimates of logit-scale locations for students and items, most Rasch software packages also provide estimates of standard errors (SE) that reflect the precision with which individual elements within each facet (i.e., individual students and items) are measured. Smaller values of the SE indicate that the logit-scale location estimate for a particular student or item is more precise, and larger values indicate that the estimate is less precise.

B. Model-data Fit Statistics. To examine the psychometric properties of the belongingness survey at the individual item level, model-data fit statistics are examined that describe the degree to which the survey items function as expected by the Rasch model. Two major types of model-data fit statistics are traditionally considered in the context of Rasch models: Infit and Outfit statistics. Both of these fit statistics are summaries of residuals, or differences between observed responses and model-expected responses. Whereas Infit statistics are particularly sensitive to less-extreme unexpected responses, Outfit statistics are more sensitive to extreme unexpected responses. Following Smith, Schumacker, and Bush (1998), this study focuses on the standardized versions of Infit and Outfit. Because these statistics are standardized, their expected value when data fit the model is 0.00, with a standard deviation of 1.00. Values that exceed ± 2 indicate misfit to the model. Specifically, values of model-data fit statistics that fall below -2 suggest less variation in the responses than expected by the model, and values greater than $+2$ suggest more variation in responses than expected. In keeping with this study's focus on the psychometric properties of the belongingness survey items, model-data fit statistics are examined in greater detail for the item facet than for the person facet.

C. Targeting. The third source of evidence of measurement precision is related to the *targeting*, or alignment between the logit-scale locations of students and items. Because the Rasch model provides estimates of students and items on a common linear scale, it is possible to compare the distribution of student and item estimates. Targeting, or a close alignment between these two distributions, suggests that the items are appropriately difficult for the sample of students, and that the students provide meaningful information about items. Correspondingly, measures of students and items are more precise when the item and student distributions share similar locations on the logit scale.

D. Reliability Statistics. Within the context of Rasch measurement theory, reliability is considered using separation statistics, which provide information about the spread of individual students, items, or other facets on the logit scale. Specifically, the reliability of separation statistic (*Rel*) is calculated for each facet as an indicator of the degree to which differences among students and items are observed based on the belongingness survey. When adequate model-data fit is observed (discussed further below), *Rel* for students is comparable to coefficient alpha. For the item facet, *Rel* describes differences in the difficulty to endorse the survey items. In addition to *Rel*, the Facets computer program (Linacre, 2015) calculates a chi-square statistic that describes the degree to which differences among logit-scale locations for students and items are statistically significant.

Results

In this section, results are summarized as they relate to the two sub-questions that were used to guide the analyses for this study: (1) How are the students and items ordered on the construct? And (2) How precise are the estimates of student and item locations on the construct? In order to address the first sub-question, student and item estimates are described in terms of their relative locations on the logit scale. In order to address the second sub-question, indicators of measurement precision are discussed as they relate to the four categories described above: (A) standard errors; (B) model-data fit; (C) targeting; and (D) reliability.

How are Students and Items Ordered on the Construct?

In order to examine the ordering of students and items on the construct, the Rasch model was used to obtain estimates of students and items on a linear scale that represents belongingness. First, these estimates were examined visually using Figure 1, which is a variable map that illustrates the calibration of the students, items, and rating scale categories on the logit scale. This variable map was created using the *WrightMap* package for R (Torres Iribarra & Freund, 2014), and it reflects the general structure of Rasch variable maps that are available from other software packages, including Facets (Linacre, 2015) and Winsteps (Linacre, 2016). Across the figure, the y -axis reflects the logit scale, where higher values reflect a stronger sense of belongingness, and lower values reflect a lower sense of belongingness.

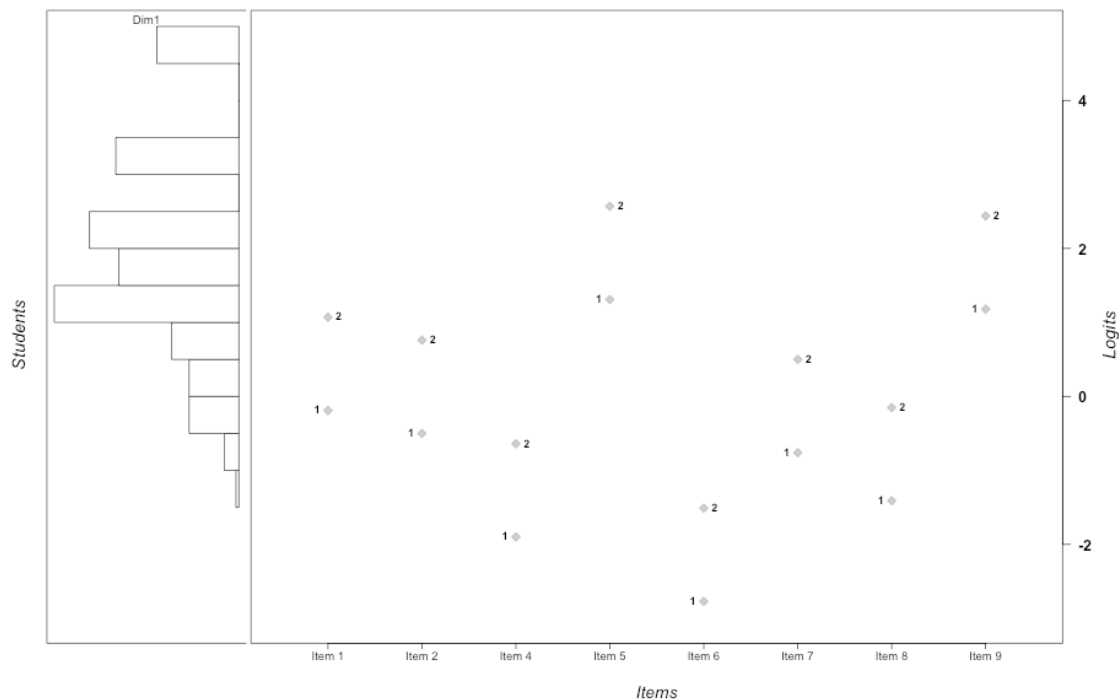


Figure 1. Variable Map

Note. High numbers on the logit scale represent a stronger sense of belongingness. Items 2, 4, and 6 were reverse-coded prior to analysis, such that higher ratings represent a stronger sense of belongingness. Item 3 is not included in the analysis.

Student Locations

In a preliminary analysis, the student locations revealed one student who appeared to be an outlier when viewed alongside the remaining 293 students in the sample (location = -4.64 logits). To clarify the interpretation of person estimates, we removed this student from the final analysis. The panel on the left side of Figure 1 shows the distribution of the remaining 293 student locations on the logit scale using a histogram. Overall, these student calibrations indicate a wide spread of belongingness measures, ranging from -1.07 logits for the student with the weakest reported sense of belongingness (Average rating = 1.62), to 4.65 logits for the student with the strongest sense of belongingness (Average rating = 2.98). Because this study reflects an analysis of data from the pilot administration of the belongingness survey in order to inform item revisions, individual student locations are not examined in detail here.

Item Locations

The panel on the right side of Figure 1 shows the calibration of the eight items in terms of their rating scale category thresholds, which were calculated using the Rasch-Andrich (non-cumulative) formulation (Andrich, 1982; Andrich, 1978). The item numbers are listed along the x -axis, and the location of the two rating scale thresholds (τ_1 and τ_2) on the logit scale are plotted along the y -axis for each item. These threshold locations reflect the difficulty associated with providing a rating in category 2, rather than category 1 (τ_1) and the difficulty associated with providing a rating in category 3, rather than category 2 (τ_2), respectively, based on the collapsed rating scale with three categories. Because a Rating Scale formulation of the Rasch model was used in this study (Andrich, 1978; see Equation 1), the relative distance between the thresholds is consistent across all eight items. However, the location of the thresholds on the logit scale varies across items. Differences in the relative locations of the thresholds across items reflect differences in the difficulty to endorse each item. For this facet, low measures on the logit scale indicate that an item is easy to endorse—that is, items with low measures do not require a strong sense of belongingness to provide a high rating; conversely, high measures on the logit scale indicate that an item is difficult to endorse—that is, items with high measures require a strong sense of belongingness to provide a high rating.

Table 1 includes detailed results related to the estimated locations for the eight survey items; these results correspond to the item locations that are shown in Figure 1. For each item, Table 1 includes an overall location estimate that reflects the balance point across the rating scale category thresholds. These overall locations can be used to examine the relative ordering of the belongingness items on the construct. Across the survey items, Item 6 (recoded prior to analysis): *I would be ashamed of learning French* was the easiest item to endorse (Estimate = -2.14 logits, Average rating = 2.92), followed by Item 4 (re-coded prior to analysis): *I would not tell others that I am learning French* (Estimate = -1.27 logits, Average rating = 2.85). The next four items are ordered in terms of increasing difficulty to endorse as follows: Item 8: *I wouldn't worry about what people think of me learning French* (Estimate = -0.78 logits, Average rating = 2.78); Item 7: *I would tell people that I'm learning French* (Estimate = -0.13 logits, Average rating = 2.66); Item 2 (recoded prior to analysis): *I wouldn't want others to know that I took French* (Estimate = 0.13 logits, Average rating = 2.60); and Item 1: *I would be proud to be learning French* (Estimate = 0.44 logits, Average rating = 2.51). The

second-most difficult item to endorse was Item 9: *I would feel like I'm part of a community of people who learn French* (Measure = 1.81 logits, Average rating = 2.06), followed by Item 5: *I have many friends who take French* (Measure = 1.94 logits, Average rating = 2.01), which was the most difficult item to endorse.

Table 1
Item Calibration Results

Item	Average Rating	Overall Item Estimate (logits)	SE	Std. Infit	Std. Outfit
5: I have many friends who take French.	2.01	1.94	0.10	0.95	1.03
9: I would feel like I'm part of a community of people who learn French.	2.06	1.81	0.10	-0.27	-0.34
1: I would be proud to be learning French.	2.51	0.44	0.11	-0.01	1.23
2R: I wouldn't want others to know that I took French.	2.60	0.13	0.12	1.63	0.80
7: I would tell people that I'm learning French.	2.66	-0.13	0.12	-2.01	-1.82
8: I wouldn't worry about what people think of me learning French.	2.78	-0.78	0.15	1.88	-0.54
4R: I would not tell others that I am learning French.	2.85	-1.27	0.17	0.52	-0.65
6R: I would be ashamed of learning French.	2.92	-2.14	0.25	-0.17	-0.90
<i>Mean</i>	2.55	0.00	0.14	0.32	-0.15
<i>SD</i>	0.34	1.42	0.05	1.24	1.07

Note. Items are ordered by the overall item estimate (logits) from high (difficult to endorse) to low (easy to endorse). Items 2, 4, and 6 were reverse-coded prior to analysis, such that higher measures indicate a stronger sense of belongingness. Item 3 is not included in the analysis.

How precise are the estimates of student and item locations on the construct?

Next, the precision of student and item estimates was considered using four sources of evidence: (A) standard errors; (B) model-data fit; (C) targeting; and (D) reliability. Results related to these four sources of evidence are summarized in Table 2. In keeping with the purpose of this study, although these indices were considered for both students and items, the results related to items are considered in more detail.

Table 2
Evidence of Measurement Precision

Indicator of Measurement Precision		Students ($N=293$)	Items ($N=8$)
A. Standard Error	(Min, Max)	(0.58, 1.86)	(0.10, 0.25)
	Mean	0.85	0.14
	SD	0.37	0.05
B. Model-Data Fit Statistics	Std.	(-2.75, 3.24)	(-2.01, 1.88)
	(Min, Max)		
	Mean	0.01	0.32
	SD	1.04	1.24
	Std.	(-2.75, 3.13)	(-1.82, 1.23)
	(Min, Max)		
C. Targeting (logit scale locations)	Mean	0.11	-0.15
	SD	0.83	1.07
	(Min, Max)	(-4.64, 4.65)	(-2.14, 1.94)
D. Reliability	Mean	1.99	0.00
	SD	1.36	1.42
	Reliability of Separation	0.60	0.99
	Chi-Square	680.9*	716.7*
	df	292	7

* $p < 0.001$

A. Standard Errors

First, SEs were examined for each of the student and item estimates on the logit scale. Summary statistics for student and item SEs are provided in Table 2, Panel A. For students, SEs ranged from 0.58 to 1.86, with an average value of 0.86 ($SD=0.37$). This finding suggests that the precision of student location estimates varied widely across the sample of student who participated in the pilot administration of the Student Perceptions of Belongingness Survey. Furthermore, this finding suggests that additional steps are necessary before the survey items can be used to meaningfully distinguish among individual students in terms of their perceptions of belongingness, including the addition of survey items that are more closely aligned to student locations (discussed further below), as well as the investigation of differential person functioning (DPF).

For items, SEs ranged from 0.10 to 0.25, with an average value of 0.14 ($SD=0.05$; see Table 1 and Table 2). These values suggest that the precision of item estimates was higher and more consistent than the precision of student estimates. However, differences among SEs for the eight survey items suggest that the items vary in terms of the precision with which they are located on the logit scale. As was observed related to SEs for student estimates, revisions to the Student Perceptions of Belongingness Survey are needed in order to improve the precision of item estimates, including the use of additional items that are more closely targeted to student locations and the investigation of differential item functioning (DIF).

B. Model-Data Fit

Next, model-data fit statistics were examined for students and items. Summary statistics for student and item fit are given in Table 2, Panel B. Item-level fit statistics are also provided in Table 1. As noted above, the model-data fit analysis in this study included the use of standardized Infit and

Outfit statistics based on the Rasch model. These statistics are calculated for each student and item as evidence of person fit and item fit to the model, respectively. For both students and items, values between -2.00 and +2.00 suggest that the observed pattern of responses for an individual student across items, or for an individual item across students, reflects the expected patterns based on the model.

As can be seen in Table 2, Panel B, acceptable average model-data fit statistics for students suggest that student response patterns generally adhered to the expectations of the Rasch model. Specifically, average values of both the standardized Infit and Outfit statistics reflect the expected values when data fit the Rasch model (Std. Infit: $M=0.02$, $SD=1.03$; Std. Outfit: $M=0.11$; $SD=0.82$). Among the 293 undergraduate students included in the sample for the pilot administration, five students' standardized Infit statistics were lower than expected (Std. Infit < -2.00), and one student's standardized Outfit statistic was lower than expected (Std. Outfit < -2.00); these lower-than-expected values indicate that these students' response patterns included less variation than expected by the model. On the other hand, 12 students' standardized Infit statistics were higher than expected (Std. Infit > +2.00), and eleven students' standardized Outfit statistics were higher than expected (Std. Outfit > +2.00); these higher-than-expected values indicate that these students' response patterns were more variable than expected by the model. However, the proportion of students identified as misfitting based on these statistics is within the expected range of persons who misfit the model based on chance alone. Additional analyses of these students with lower- and higher-than-expected fit statistics, including DPF analyses and qualitative investigations, are warranted in order to more fully understand these students' interpretation of the survey items.

Table 2, Panel B also indicates acceptable average values of standardized Infit and Outfit statistics for items (Std. Infit: $M=0.32$, $SD=1.24$; Std. Outfit: $M=-0.15$; $SD=1.07$). Furthermore, the item-level model-data fit statistics reported in Table 1 suggest adequate fit to the Rasch model for the items included in the pilot administration of the belongingness survey. Across the survey, only one item had a standardized fit statistic slightly lower than -2.00 (Item 7: *I would tell people that I'm learning French*; Std. Infit = -2.01). Values of the standardized Infit and Outfit statistics for all of the other items fell between -2.00 and +2.00.

C. Targeting

In order to provide a frame of reference for comparing student measures to item measures on the logit scale, the item locations were centered at zero logits ($M = 0.00$ logits). The average student calibration was 1.93 logits ($SD = 1.50$), which is notably higher than the average calibration of the item facet ($M = 0.00$, $SD = 1.33$). This finding suggests that the group of students who responded to the Student Perceptions of Belongingness Survey had a strong overall sense of belongingness relative to the difficulty of the items. Furthermore, this finding suggests that additional items are needed in order to achieve closer targeting between students and items. In particular, additional items that are targeted to students with high locations on the construct would provide a more closely targeted measure of belongingness that would result in more-precise estimates of student and items in terms of this construct.

D. Reliability

The final source of evidence of measurement precision for the Student Perceptions of Belongingness Survey items is related to reliability. In the context of Rasch measurement theory, reliability is considered using *reliability of separation statistics*, which are calculated separately for each facet in the model. In the current study, reliability of separation statistics were calculated for students and item in order to explore the degree to which there were differences in logit scale locations for the individual students and items. The reliability of separation statistic for the student facet was notably lower ($Rel = 0.60$) than the value observed for the items ($Rel = 0.99$). When examined alongside the significant values of the chi-square statistic for the two facets ($p < 0.001$), these findings suggest that there are significant differences among the logit scale locations for individual students and items. The finding of a lower reliability of separation statistic for the student facet suggests that there may be several groups of students with similar levels of belongingness, within which differences among individual students in terms of locations on the construct may be less distinct. Additional items that are more-closely targeted to the student locations would likely provide additional insight into student ordering and lead to a higher value of the reliability of separation statistic for students.

Summary and Discussion

The major purpose of this study was to consider the psychometric properties of a new survey developed to explore the construct of belongingness. Using the rating scale formulation of the Rasch model, the psychometric quality of the Student Perceptions of Belongingness Survey was considered in terms of fundamental measurement properties, including item and student ordering and measurement precision. Together, the analyses provide insight into the degree to which this instrument can be used to measure students' sense of belongingness in the context of language learning, as well directions for revisions to the instrument prior to its use in future research. Overall, the results suggested that the items were functioning as expected based on the Rasch model, and that there were differences in the relative difficulty to endorse the survey items. Differences in item locations provide insight into various aspects of students' sense of belongingness related to language learning that have implications for research, theory, and practice in language instruction. However, the results also revealed several areas for improvement to the Student Perceptions of Belongingness Survey. Evidence of overall acceptable fit to the Rasch model suggests that the results from the belongingness survey items can be used to inform research and theory. A closer examination of the results affords a more-detailed understanding of the construct of belongingness as well as areas of improvement for the instrument before its use in future studies. In order to respond to the overarching research question, this section includes a discussion of the results in terms of the overall psychometric quality of the Student Perceptions of Belongingness Survey based on the pilot administration, followed by a discussion of student and item locations on the variable map. Lastly, implications of the results for research, theory, and practice are discussed.

Psychometric Quality of the Student Perceptions of Belongingness Survey

Results from initial analyses revealed severe model-data misfit for Item 3 (*I feel like I would be the only person I know who took French*), with exclusively positive residuals. This result suggests that all of the unexpected responses to this item occurred as a result of higher-than-expected responses. These results are congruent with the hypothesis that there were strong French programs and communities at the institutions where data were collected. This hypothesis was further supported by the average student location of 1.93 ($SD = 1.50$) being nearly two logits higher than the average item location ($M = 0.00$, $SD = 1.33$). Based on this finding, the group of undergraduate students who responded to the Student Perceptions of Belongingness Survey may be described as having a strong overall sense of community.

Relatedly, infrequent use of categories 2, 3, and 5 on the original 6-point rating scale may further speak to the potentially polarized experiences of individuals. Students' strong reactions towards questions of belongingness, as demonstrated by the frequent use of categories 1, 4, and 6, suggest that many respondents self-reported either very positive or very negative experiences related to belonging. Such findings may relate to Baumeister and Leary's (1995) assertion that a sense of belongingness requires frequent interaction and persisting caring. This connection is proposed because a sense of belongingness may require a certain level of interaction and caring which, if not met, has the potential to correlate with continued self-reports of feelings of not belonging.

Student and Item Locations

In congruence with the hypothesis regarding infrequent use of categories 2, 3, and 5, a simultaneous consideration of the relative ordering of items 3, 5, and 9 suggest a complex belongingness landscape. Although the results regarding item 3 and average student calibrations suggest an overall strong sense of belonging among the individuals in the sample, the finding that item 5 (*I have many friends who take French*) and item 9 (*I would feel like I'm part of a community of people who learn French*) were hardest to endorse offer additional insight into students' perceptions of belongingness related to learning French. Specifically, the relative difficulty of these items suggest that although students report not feeling entirely alone in learning French, they indicate that they do not have a large or close network of friends who study the French language. Thus, it is possible that French language learning is perceived as being an activity that is engaged in by others, but which is not done to the same extent in communities with which students self-identify.

Furthermore, it is informative to consider the finding that item 6R (*I would be ashamed of learning French*) was the easiest item to endorse, as compared to other items, and that item 1 (*I would be proud to be learning French*) was relatively difficult to endorse. Together, these results suggest that, although students report neither instances of extreme exclusion, nor significant shame in learning French, they appear to indicate a general, relative lack of pride in learning the language. In other words, it is easiest for an individual to say "I wouldn't be ashamed" but it relatively difficult to say "I would be proud." Students' responses to items related to disclosure suggest additional complexity in the belongingness construct. To elaborate, the students do not appear to express being worried about what others may think of them learning French, as indicated by the relative ease of

endorsing item 8. However, they simultaneously do not appear to report a clear desire to actively disclose their language learning as demonstrated by the relative ordering of items 2R, 7, 8, and 4R on the variable map from hardest to easiest to endorse. This ordering of items 2R and 7 as the hardest to endorse, appears to indicate that they would not necessarily want others to know that they were learning French (item 2R "*I wouldn't want others to know that I took French*") and that they would not necessarily tell others that they are learning the language (item 7 "*I would tell people that I'm learning French*"). In other words, it is relatively difficult for students to say that they would actively tell others that they are learning French (item 7) and it is also relatively difficult for them to say that they would want others to know that they are learning the language (item 2R). Complexity arises, however, when the relative ease of endorsing items 8 and 4R is considered. Contrary to items 7 and 2R, these items appear to indicate that individuals generally would not worry what others think of them learning French (item 8) and that they would tell others that they are learning the language (item 4R). Considered together, items 7 and 4R ("*I would not tell others that I am learning French*") appear to suggest that, although students do not generally report feeling proud or particularly ashamed to learn French (items 1 and 6R), they may (item 4R) or may not self-disclose (item 7), and although they may not want others to know that they are learning French (item 2R), they are not generally worried about what others think of their language learning (Item 8).

Implications for Research, Theory, and Practice

To summarize, this study proposed an instrument to measure the construct of sense of belongingness through statements regarding sense of community, pride and shame, and disclosure patterns. The findings from this study have important implications for research and theory regarding sense of belongingness in the field of language teaching and learning. With regard to research and theory, findings from the analysis of the pilot administration of the Student Perceptions of Belongingness Survey demonstrated the complexity of undergraduate students' sense of belongingness in language learning, revealing distinctions between peripheral knowledge of community and integration into a community of target language speakers. Such differences between peripheral and integrated community membership echo theories regarding communities of practice (Eckert, 2006). Further, the results suggested that a reported overall sense of community does not necessarily co-occur with pride in learning French or with a distinct desire to actively disclose engagement in French language learning. Future research will need to consider whether a stronger sense of belonging, in which students feel they are part of a community of people who learn French (item 9) and have friends who learn the language (item 5), co-occurs with a clearer sense of pride or a clearer desire to self-disclose French language learning.

In following, with regard to practice, the results suggest that making the creation of community of central importance in language classrooms may have significant benefits. Our current findings suggested that an overall sense of community co-occurred with a lack of extreme exclusion as well as a lack of shame, although future research is needed to determine if a stronger sense of belonging would co-occur with marked pride in and a desire to disclose French language learning. The complex process of community building does not appear to be achieved without effort. Nonetheless, the results from this study, the literature on communities of practice (Eckert, 2006), and the demonstrated importance of intersubjectivity and belonging for successful language learning (Julé,

2004; Kramsch, 2009), together, suggest that this work is of marked importance. This argument is coincides with the theoretical connection between sense of belongingness, achievement (Baumeister & Leary, 1995), and enrollment.

Both the results regarding community and those regarding disclosure patterns have practical applications that extend beyond the classroom. The findings regarding disclosure patterns suggest that there may be a complex relationship between students' sense of belongingness in language learning and broader societal perceptions. Advocacy efforts surrounding the benefits of language learning writ large and of French more specifically may be needed in order to change societal perceptions and open the door for a greater sense of belongingness among French language learners. Specifically, increased belongingness has the possibility of engendering pride in learning French and a clearer desire to disclose language learning that was not seen in this study where only a general sense of community was reported as opposed to a strong sense of belonging characterized by integrated community membership.

Conclusions

This instrument-development study was guided by an overarching research question that considered the degree to which a new instrument designed to measure belongingness in the context of language learning adhered to fundamental measurement properties. With respect to the measurement properties of the Student Perceptions of Belongingness Survey, the Rasch model revealed severe misfit for item 3 which may be indicative of a generally strong sense of belonging to a French-language community among participants. Further, the infrequent use of categories 2, 3, and 5 raised questions regarding potentially polarized feelings of belonging and not belonging. The relative ordering of items on the variable map also suggested the complexity of sense of belongingness as a construct. In contrast to item 3, the difficulty to endorse items 5 and 9 suggested that the undergraduate students who participated in the pilot administration of this instrument reported a general lack of community. Such mixed responses suggest that, although students may know others who learn French, they may feel as though they remain "othered". That is to say, these students may report having a limited or peripheral network of individuals they know or know of who speak French. In following, it is possible that the undergraduate participants in this study considered French language learning to be a pursuit undertaken by individuals outside of their own communities. This complexity was repeated in student self-reports of generally lacking pride as well as shame in learning French. It was further echoed by reported disclosure patterns in which students appeared neither to consistently desire to actively disclose nor to consistently desire to hide their language learning. These findings have multiple implications for theory, research, and practice. Specifically, the three aspects of sense of belongingness (sense of community, pride versus shame, and disclosure patterns) that arose from the data were shown to not necessarily co-occur. Further, results suggest that students' sense of belonging to a community cannot be taken for granted. Accordingly, the collaborative development of strategies for actively building community by researchers and practitioners has the potential to yield significant benefits in terms of student perceptions of their own sense of belongingness. Relatedly, complex disclosure patterns also suggested the need for advocacy work to increase societal support for French language learning and cultivate a sense of pride in learning French among students.

Directions for Future Research

Although this study has made significant contributions to enhancing understanding of the construct of sense of belongingness in the field of French language education in the United States, future research is needed to expand upon its findings. First, in response to the limitations of an eight-item instrument, future work could improve the measure by adding additional survey items to the Student Perceptions of Belongingness Survey, particularly those targeted toward students with high locations on the variable map. Additional items may improve the precision of student estimates. Additional items would also allow for future research to more thoroughly consider the aspects of sense of belongingness as well as differences in individual student locations on this construct. Further, in response to the complexity of interpretation for both survey participants and survey users (e.g., researchers) engendered by negatively-worded items, it is recommended that future studies avoid using negatively-worded statements for both the existing and additional survey items. Second, due to the culturally bound nature of the construct considered, work to investigate generalizability to other languages, contexts, and populations will need to be undertaken. Future research with different samples may provide additional validity evidence.

Third, extending these Rasch analyses to a Many-Facet Rasch Model (Linacre, 1989) including explanatory variables, such as previous language-learning experiences, gender, and race/ethnicity, may allow for further understanding as to whether survey responses are systematically different among individuals with various demographic characteristics or experiences. In particular, differential item and person functioning (DIF and DPF) analyses related to explanatory variables would likely provide valuable insight into students' sense of belongingness related to language learning. Furthermore, DIF and DPF analyses have the potential to elucidate predictors of a strong sense of belongingness. As such, these analyses will likely provide valuable insight into the connection between this belongingness and student retention in language courses. Alternatively, latent regression analyses related to these explanatory variables would shed light on the overall role of these demographic characteristics and experiences that are not specific to individual items and persons. These analyses are particularly promising as a potential method for exploring the complex response patterns observed in the current study with regard to items 2R, 4, 7, and 8.

Finally, the results from this study provide evidence of complex relationships among various aspects of belongingness, including students' sense of community, pride, and willingness to disclose their language learning practices. Because the results suggest that these characteristics do not necessarily co-occur, future research is needed to consider the relationship among these three aspects of belongingness. Additionally, mixed-methods inquiries into participant impressions of the to survey items may serve to deepen understandings of this instrument. In terms of sense of community, future research should more deeply engage with communities of practice (Eckert, 2006) in order to further investigate peripheral and integrated senses of community. Bridging theory, research, and practice, instructors and researchers are encouraged to continue work together to investigate ways to foster a sense of belongingness, given its theoretical link to increased achievement (Baumeister & Leary, 1995), enrollment, and retention.

The original article was received on December 27th, 2016

The revised article was received on October 22nd, 2017

The article was accepted on October 27th, 2017

References

- Andrich, D. A. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43(4), 561–573. <https://doi.org/10.1007/BF02293814>
- Baumeister, R., & Leary, M. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117(3), 497-529.
- Bond, T.G., & Fox, C.M. (2015). *Applying the Rasch Model: Fundamental measurement in the human sciences*. New York: Routledge.
- Bowlby, J. (1969). *Attachment and Loss: Vol.1. Attachment*. New York: Basic Books.
- Bowlby, J. (1973). *Attachment and Loss: Vol. 2. Separation anxiety and anger*. New York: Basic Books.
- Castillo Zaragoza, Elena Desirée. “Identity, motivation and plurilingualism in self-access centers.” *Identity, motivation and autonomy in language learning* (Vol. 54). Ed. Murray, Garold, Xuesong Gao, and Terry Lamb. Bristol: Multilingual Matters, 2011. 91-106.
- De Rivera, J. (1984). The Structure of Emotional Relationships. In P. Shaver (Ed.) *Review of personality and social psychology, Vol.5: Emotions, relationships, and health* (pp. 116-145). Beverly Hills, CA: Sage.
- Eckert, P. (2006). Communities of practice. *Encyclopedia of language and linguistics*, 2(2006), 683-685.
- Engelhard, G. Jr. (2013). *Invariant measurement: Using Rasch models in the social, behavioral, and health sciences*. New York: Routledge.
- Epstein, S. (1992). The Cognitive Self, the Psychoanalytic Self, and the Forgotten Selves. *Psychological Inquiry*, 3, 34-37.
- Fromm, E. (1955). *The Sane Society*. New York: Holt, Rinehart and Winston.
- Fromm, E. (1956). *The Art of Loving*. New York: Harper and Brothers.
- Guisinger, S., & Blatt, S. J. (1994). Individuality and Relatedness: Evolution of a Fundamental Dialectic. *American Psychologist*, 49, 104-111.
- Hogan, R. (1983). *A Socioanalytic Theory of Personality*. In M. Page and R. Dienstbier (Eds.) *Nebraska Symposium on Motivation, 1982* (pp. 55-89). Lincoln: University of Nebraska Press.
- Holland, D., Lachicotte Jr, D., & Skinner, D. D., & Cain, C. (2001). *Identity and agency in cultural worlds*. Harvard University Press.
- Horney, K. (1945). *Our Inner Conflicts: A Constructive Theory of Neurosis*. New York: Norton.
- Huang, Jing. A dynamic account of autonomy, agency and identity in (T) EFL learning. *Identity, motivation and autonomy in language learning* (Vol. 54). Ed. Murray, Garold, Xuesong Gao, and Terry Lamb. Bristol: Multilingual Matters, 2011. 229-246.
- Julé, A. (2004). Intersubjectivity in Language Classrooms. In *Gender, Participation and Silence in the Language Classroom* (pp. 11-18). Palgrave Macmillan UK.
- Knisely, K. A., & Wind, S. A. (2015). Exploring gendered language attitudes using Rasch measurement theory. *Journal of Applied Measurement*, 16(1), 95-112.
- Knisely, K. A. (in press). Exploring the role of culture in the formation of French language ideologies among US post-secondary students. *Contemporary French Civilization*.
- Knisely, K. A. (2015). *Language Learning and the Gendered Self: Learner Identities and French Language Study in a US Context* (Unpublished doctoral dissertation). Emory University, Atlanta, Georgia.
- Knisely, K. A. (2016). Language Learning and the Gendered Self: The Case of French and Masculinity in a US Context. *Gender and Language* 10(2).

- Kramersch, C. (2009) *The multilingual subject: What foreign language learners say about their experience and why it matters*. Oxford University Press.
- Linacre, J. M. (1989). *Many-facet Rasch measurement* (2nd Ed.). Chicago: MESA.
- Linacre, J. M. (2015). Facets Rasch Measurement (Version 3.71.4). Chicago, IL: Winsteps.com.
- Linacre, J. M. (2016). Winsteps Rasch Measurement (Version 3.92.1). Chicago, IL: Winsteps.com.
- Masgoret, A. M., & Gardner, R. C. (2003). Attitudes, motivation, and second language learning: a meta-analysis of studies conducted by Gardner and associates. *Language learning*, 53(1), 123-163.
- Maslow, A. H. (1943). A Theory of Human Motivation. *Psychological Review*, 50, 370-396.
- Maslow, A. H. (1968). *Toward a Psychology of Being*. New York: Van Nostrand.
- Rasch (1960/1980). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research. (Expanded edition, Chicago: University of Chicago Press, 1980).
- Rosello, M. (1994). "Get Out of Here!": Modern Gay Queer Languages in the 1990s. *Canadian Review of Comparative Literature/Revue Canadienne de Littérature Comparée* 21(1), 149-168.
- Ryan, R. M. (1991). The Nature of the Self in Autonomy and Relatedness. In J. Strauss and G. R. Goethals (Eds.) *The Self: Interdisciplinary Approaches* (pp. 208-238). New York: Springer-Verlag.
- Smith, R. M., Schumacker, R. E., & Bush, J. J. (1998). Using item mean squares to evaluate fit to the Rasch model. *Journal of Outcome Measurement*, 2, 66-78.
- Sullivan, H. S. (1953). *The Interpersonal Theory of Psychiatry*. New York: Norton.
- Andrich, D. A. (1982). An extension of the Rasch model for ratings providing both location and dispersion parameters. *Psychometrika*, 47(1), 105-113. <https://doi.org/10.1007/BF02293856>
- Andrich, D. A. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43(4), 561-573. <https://doi.org/10.1007/BF02293814>
- Linacre, J. M. (2015). Facets Rasch Measurement (Version 3.71.4). Chicago, IL: Winsteps.com.
- Linacre, J. M. (2016). Winsteps Rasch Measurement (Version 3.92.1). Chicago, IL: Winsteps.com.
- Masters, G. N. (1982). A rasch model for partial credit scoring. *Psychometrika*, 47(2), 149-174. <https://doi.org/10.1007/BF02296272>
- Torres Irribarra, D., & Freund, R. (2014). *Wright Map: IRT item-person map with ConQuest integration*. Retrieved from <http://github.com/david-ti/wrightmap>
- Wenger, E. (2000). Communities of practice and social learning systems.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. New York, NY: Taylor & Francis.
- Wilson, M. (2011). Some notes on the term: "Wright Map." *Rasch Measurement Transactions*, 25(3), 1331.
- Wright, B. D., & Masters, G. N. (1982). *Rating Scale Analysis. Rasch Measurement*. Chicago: MESA Press.
- Wolfe, E. W. & Smith, E. V., Jr. (2007a). Instrument development tools and activities for measure validation using Rasch models: Part I - Instrument development tools. *Journal of Applied Measurement*, 8, 97-123.
- Wolfe, E. W. & Smith, E. V., Jr. (2007b). Instrument development tools and activities for measure validation using Rasch models: Part II - Validation activities. *Journal of Applied Measurement*, 8, 204-233.

Appendix A: Student Perceptions of Belongingness Survey

People may feel differently about the idea of learning French and how they feel others might feel about them learning French. We would like to understand how you feel and how you feel others might feel. For each of the statements below, circle the number which best represents your answer.

1 = strongly disagree
4 = slightly agree

2 = moderately disagree
5 = moderately agree

3 = slightly disagree
6 = strongly agree

- | | |
|--|-------------|
| 1. I would be proud to be learning French. | 1 2 3 4 5 6 |
| 2**. I wouldn't want others to know that I took French. | 1 2 3 4 5 6 |
| 3*. I feel like I would be the only person I know who took French. | 1 2 3 4 5 6 |
| 4**. I would not tell others that I am learning French. | 1 2 3 4 5 6 |
| 5. I have many friends who take French. | 1 2 3 4 5 6 |
| 6**. I would be ashamed of learning French. | 1 2 3 4 5 6 |
| 7. I would tell people that I'm learning French. | 1 2 3 4 5 6 |
| 8. I wouldn't worry about what people think of me learning French. | 1 2 3 4 5 6 |
| 9. I would feel like I'm part of a community of people who learn French. | 1 2 3 4 5 6 |

* Item not included in analysis; ** Reverse-coded prior to analyses.