

# Development and Validation of a High School Agricultural Literacy Assessment

R. Judd-Murray<sup>1</sup>, B. K. Warnick<sup>2</sup>, D. C. Coster<sup>3</sup>, M. L. Longhurst<sup>4</sup>

## Abstract

The National Agricultural Literacy Outcomes (NALOs) are knowledge benchmarks for school-aged youth and are used to improve agricultural literacy (National Agriculture in the Classroom, 2014; National Center for Agricultural Literacy, 2017). Despite educational efforts, prior research indicated that high school populations remained at low or deficient literacy levels. Additionally, no agricultural literacy assessment instruments using the NALOs as a standardization benchmark have been developed for the 9-12th grades. The purpose of the study was to validate a summative NALO-centered assessment that could provide baseline data on agricultural literacy following the completion of secondary education (12<sup>th</sup> grade). The study followed the framework established by Longhurst et al. (2020) for similar assessments in elementary grades. A Delphi team produced 45 items for validation that were reviewed using a convenience sample of Utah State University undergraduate students. Those items were evaluated using factor, item, and discriminant analysis. Results finalized two 15-item assessments and determined both had acceptable reliability, were adequate for model fit, and were valid for the NALOs and the three proficiency levels. The instruments are critical tools for providing a standardized approach to evaluation efforts. Researchers and educators should use these instruments to provide comparable agricultural literacy data across populations to better identify trends, program needs, and meaningful inferences.

## Article History





Received: October 12, 2023

Accepted: May 20, 2024

Published: June 18, 2024

## Keywords

National Agricultural Literacy Outcomes; agricultural education; agricultural knowledge; high school; secondary education

- 
1. Rose Judd-Murray, Assistant Professor, Utah State University, 6000 Old Main Hill, Logan, UT 84322-6000, [rose.juddmurray@usu.edu](mailto:rose.juddmurray@usu.edu),  <https://orcid.org/0000-0002-2045-1015>
  2. Brian K. Warnick, Department Head and Professor, Utah State University, 6000 Old Main Hill, Logan, UT 84322-6000, [brian.warnick@usu.edu](mailto:brian.warnick@usu.edu),  <https://orcid.org/0009-0001-8608-1897>
  3. Daniel C. Coster, Professor, Utah State University, 3900 Old Main Hill, Logan, UT 84322-3900, [dan.coster@usu.edu](mailto:dan.coster@usu.edu),  <https://orcid.org/0000-0002-4801-5637>
  4. Max L. Longhurst, Professional Practice Associate Professor, Utah State University, Old Main Hill, Logan, UT 84322-2805, [max.longhurst@usu.edu](mailto:max.longhurst@usu.edu),  <https://orcid.org/0000-0003-3227-800X2805>

## Introduction and Problem Statement

Agricultural literacy efforts prepare K-12 students to recognize and interpret information relevant for determining adult decisions regarding their health, global environment, public policy, and economic benefits (Hess & Trexler, 2011; Lawson & Weser, 1990; Redmond & Griffith, 2003). Agricultural literacy also influences positive perceptions and attitudes about agriculture (Specht et al., 2014). Due to its importance, literacy assessments were developed using a variety of benchmarks and methods for K-12 students (Frick, 1993; Leising et al., 1998, 2000; Powell et al., 2008). While these instruments provided relevant data, Brandt (2016) and Longhurst et al. (2020) noted that older frameworks and definitions did not meet current needs. The lack of consistency in instrumentation contributed to a lack of replication and the ability to compare results with other populations nationwide. Both suggested that using the National Agricultural Literacy Outcomes (NALOs) (Spielmaker & Leising, 2013) K-12 grade-level-banded benchmarks and the National Agricultural Literacy Logic Model (Spielmaker et al., 2014), a validated educational framework, could provide the consistency and uniformity necessary for a benchmark-based agricultural literacy assessment instrument. Moving toward uniformity can provide a pathway for agricultural literacy assessment that is more reliable, comparable, and applicable across different studies, disciplines, and contexts.

The absence of a standardized tool to measure agricultural literacy during and after high school limits the understanding of agricultural literacy levels in different populations. Without a consistent assessment, it is difficult to identify national knowledge gaps, design targeted education interventions, and track progress in educational efforts. Consequently, this study aimed to develop a summative high school (12th grade) agricultural literacy assessment using the NALOs as a foundational tool for future research.

## Theoretical and Conceptual Framework

The framework for this study is based on Longhurst et al. (2020) who developed and validated a 3-5th grade NALO-based agricultural literacy assessment. Their work provided a replication model for us; their success was centered upon two essential frameworks.

### National Agricultural Literacy Outcomes Framework

Cosby et al. (2022) stated that the NALOs “provide the most comprehensive learning framework across the globe against which to measure student agricultural literacy...and they provide benchmarks to increase uniformity across the national education system in the USA” (p. 10). The NALO benchmarks were developed via Delphi using a rigorous integration of national grade level benchmarks and national standards for science, social studies, and health—organized through the lens of agricultural literacy (Spielmaker & Leising, 2013). The NALOs reflect prior research and five cross-disciplinary themes: (a) Agriculture and the environment, (b) Plants and animals for food, fiber & energy, (c) Food, health & lifestyle, (d) STEM, and (e) Culture, society, economy & geography (National Center for Agricultural Literacy [NCAL], 2017). They were designed as a logic model component for K-20 assessment and program evaluation for National Agriculture in the Classroom programs and the National Center for Agricultural

Literacy (Spielmaker et al., 2014). The NALOs align with the AAAE Research Values (American Association for Agricultural Education [AAAE], 2023), providing a measure for determining the impacts of educational outreach. Within this study, the NALOs guided instrument item development through the question: *What must students know or be able to do with the information they have learned to be proficient in the NALO standards?*

### **Programme for International Student Assessment (PISA) Framework**

Generally, summative assessments are cumulative to determine what students do or do not know. A significant limitation is determining only a pass or failing score, where a failing score may convey that the student lacks any understanding (Boud & Falchikov, 2006). The National Research Council (2009) suggested using assessments that showed a progression sequence because it identifies what a person can do within stages of development. Therefore, the PISA framework served as a guide because it “assesses students and uses the outcomes of that assessment to produce estimates of students’ proficiency in relation to the skills and knowledge being assessed in each domain” (OECD: Programme for International Student Assessment [OECD: PISA], 2016, p. 276). The PISA framework has well-defined parameters. The domains used within this framework were the five NALO themes. The domain skills (assessment items) were developed from very low levels of proficiency to very high levels. Following the structure, the easiest items focused on content knowledge and the relation to agricultural phenomena. The most difficult items drew upon interrelated ideas and concepts that required “an understanding of events, consequences, or processes” (OECD: PISA, 2016, p. 282). Within this context, a student’s agricultural literacy level determined their place on a sliding proficiency scale, ranked by how frequently they answered questions correctly. This premise followed the central dogma of the PISA assessment: “if a student’s proficiency level exceeds the item’s difficulty, the probability that the student can complete that item is high, and if the student’s proficiency is lower than what is required by the item, the probability for student success on that item is low” (OECD: PISA, 2016, p. 279). Most importantly, the NALOs were constructed in grade-banded levels that interrelated and overlapped, ensuring that students advanced from primary to advanced content, practice, and examples of complexities as they moved toward more sophisticated curricula. The integration of both frameworks established a well-defined model for (a) developing questions that represented an increase in skill and ability for better understanding student proficiency and (b) providing data that were representative of progression toward literacy.

## **Purpose**

Guided by the conceptual frameworks, the study aimed to develop and validate a summative high school (12th grade) NALO-centered instrument that could assess proficiency levels of agricultural literacy.

The study addressed the following research question: Is the instrument a valid and reliable measure of the five Grade 12 NALO themes and the proficiency stages of agricultural literacy (i.e., exposure, factual literacy, and applied proficiency)?

## Methods

There were three phases for defining the quantitative development and validation of the agricultural literacy assessment instrument.

### Phase One: Instrument Construction

Longhurst et al. (2020) showed effectiveness in determining instrument items via a Delphi model because of the complexity of the content. Goodman (1987) noted that if the experts participating in the development process were representative of the area of knowledge, then content validity and reliability could be assumed. Messick (1995) and Sireci (1998) clarified that content validation added verification and critical mechanisms of construct validity. Literature also indicated that committee selection was an essential part of the process because it determined the quality of the items (Jacobs, 1996; Judd, 1972; Taylor & Judd, 1989). Therefore, the consideration of experts who participated in the Delphi construction of items was paramount. Individuals were direct experts in secondary agricultural education, curriculum development, agricultural policy, communications, cooperative extension and outreach, agribusiness, and STEM education; they were selected from multiple states and possessed advanced degrees or teaching certificates. In all, twelve members participated in item construction. Delbecq et al. (1975) suggested that ten to fifteen members were sufficient if the background of the Delphi subjects were homogenous. We promoted homogeneity in the subject selection to best represent the processing capability.

### Phase Two: Data Collection

The convenience sample population for validation was  $N = 600$  Utah State University students. Convenience sampling of college students is a prevalent approach for data collection in educational research (Hanel & Vione, 2016). Undergraduate student samples can be a legitimate solution when strongly justified, and problems can be minimized through conscientious research design and execution (Bello et al., 2009; Winton & Sabol, 2022). To carefully address these parameters, we incentivized college students with extra credit because obtaining data from end-of-senior-year high school students was extremely difficult due to our state ethics review opt-in-only rules for minors and the willingness of that population to participate in a survey at that specific time. We prioritized recruiting first-year, first-semester students but allowed older students to participate to ensure that the sample size could accurately accommodate the factor analysis that Comrey and Lee (1992) estimated for factor analysis ( $N = 500+$ ) to be very good or excellent. Survey items were accessed via Qualtrics. We monitored the survey for three weeks; email reminders were sent to students weekly throughout the collection period.

### Phase Three: Instrument Validation

We analyzed the data following procedures and processes outlined by Longhurst et al. (2020). First, data were organized, cleaned for non-response, and dummy coded. The highest and partial scores were calculated, and then an Exploratory Factor Analysis (EFA) was conducted in SAS (Version 9.4). The frequencies of the relationships between the proficiency stages determined the latent constructs. Following EFA, item analyses were conducted on items with

varying frequencies. Ultimately, the best items were identified and analyzed using Confirmatory Factor Analysis (CFA) and Discriminant Analysis (DA). We concluded our analysis by identifying the final items to construct two separate instruments.

### Limitations

College samples tend to exhibit homogeneity toward diversity, and students may fall within the higher spectrum of cognitive skills (Stevens, 2011). We underscored the importance of lived experiences that could contribute to agricultural literacy proficiency over time. Additionally, the survey items were directly associated with the NALO benchmarks, resulting in correlation, lack of independence, and multicollinearity risks. Measures of covariance among the latent variables were analyzed, but CFA results should be treated with caution. Finally, using DA enabled determining whether differences existed between the proficiency stages. The use of DA defined the degree to which the instrument differentiated between the constructs.

## Findings

### Phase One: Instrument Construction

The Delphi team developed survey items by integrating item content, relevance to the NALO demands, and effectiveness guidelines for summative assessment. Each team member was asked to create between three and five questions (including answers) for each NALO theme. The questions had to be identified by one of the three proficiency levels. The first three rounds identified which of the 64 questions best represented the NALO theme and the appropriate proficiency level. Table 1 shows an example of how a construct analysis clarified the requirements of the NALO benchmarks and the parameters for each proficiency level. It is an example of how the team combined the defined measures of the proficiency scale parameters from PISA, and the proficiency level descriptors. The process provided a point-by-point evaluation for each determinant factor required for a valid summative assessment of the 12th-grade NALOs. From there, the fourth round eliminated questions with the lowest rankings and sent the remaining items to be refined by the group for the final two iterations.

**Table 1***Construct Analysis: Examples of Theme Two Items*

Item number & Proficiency level	Assessment item content	NALO demands	Proficiency level identifiers
2.12 Exposure level	Identify examples of organic nutrients.	Lifecycles of plants and animals; distinguish between renewable and non-renewable resources; the importance of soil nutrients; compare natural cycles in comparison to managed lifecycles within agriculture; how organic and inorganic nutrients affect plant growth and development.	Students can recognize terms; recall singular facts, especially ones that draw upon their personal or familiar experiences; recognize simple cause-and-effect relationships; select simple explanations with relevant or cueing support.
2.12 Factual literacy level	Identify the factors (including cost, culture, convenience, access, and taste) that affect the population's food choices.	The variety of year-round food choices; food distribution and transportation systems; major factors in food choices for people and animals are cost, culture, convenience, and access; viewpoints on production methods and practices; impacts of transporting food due to location, climate, and geography; consumer demand influences production, processing, and marketing; consumer choices influence food production systems.	Students can order, sort, analyze, and move/transfer knowledge from one area of application to another; draw upon moderately complex facts and ideas to construct explanations; make simple predictions; identify the relevancy of facts in context.
2.12 Applied proficiency level	Determine agricultural practices that balance production and conservation (e.g., using modern science and technology).	Importance and stewardship of natural resources in delivering agricultural products and maintaining the environment; understand the concept of stewardship for soil, water, plants, and animals; examine viewpoints on production methods and practices.	Students can recognize, articulate, and evaluate what they have learned; can use abstract ideas or concepts to explain a complex phenomenon; demonstrate competency in the information that may be unfamiliar or novel; draw on a range of inter-related ideas; can construct complex predictions; internalize the significance of facts about 'real-world' application.

*Note.* Proficiency levels adapted from the works of Joplin (1981), Roberts (2006), and the PISA Technical Report (OECD: PISA, 2016).

Thus, the team finalized 45 items (three questions for each proficiency level in each NALO theme). Longhurst et al. (2020) showed that 15 questions were sufficient for the final instrument, but more questions were reviewed to increase the probability of a valid question in each theme and proficiency level.

Based on the commitment to high-quality Delphi development, replication of successful methods established in prior research, and connection to the best practices for summative evaluation, we showed that the items summatively assessed the Grade-12 NALO benchmarks and provided content and construct validity for each survey item.

### Phase Two: Data Collection

The undergraduate sampling resulted in 71% of participants having completed less than one year of college and 89% less than two years ( $n = 468$ ), with only 11% ( $n = 47$ ) having completed three to four years but being younger than 23 years old. Qualtrics reported that 580 students accessed the survey,  $N = 515$  completed the survey, and 48 did not complete the survey (89% response rate). We proceeded toward validation based on Comrey and Lee (1992) and MacCallum et al. (2001) who determined an acceptable level of  $N$  was dependent upon (a) the commonality of the variables, (b) the degree of overdetermination of the factor, (c) the size of the loading, and (d) model fit ( $f$ ). These boundaries provided a conservative measure for our sample size, with priority given to the requirements of the factor analysis due to its importance in the study.

### Phase Three: Instrument Validation

We coded 1 or 0 for correct or non-correct responses. Each possible response option was also scored as correct (1) or non-correct (0) for items with more than one correct response. This allowed for the allocation of partial scores for each overall item based on the percentages of correct responses selected by a respondent. The 45 survey items were first measured for total correct response ( $max = 34$ ,  $min = 4$ ,  $M = 21.34$ ,  $SD = 5.44$ ,  $N = 515$ ). A maximum score was used to determine initial participant proficiency stages based on PISA literature (OECD: PISA, 2016, pp. 280–281), testing parameters, and statistical best practices. Partial total correct scoring was used to determine if a survey item was too difficult or if there were only poor or too difficult portions. An item analysis, difficulty index, and correct partial percentages were critical indicators for establishing the baseline measures before factor and item analysis.

### Factor and Item Analysis

We used a structural linear equation model for EFA and CFA. Three latent factors representing the proficiency stages were analyzed against the items from each NALO theme. The factor loadings determined the influence of the proficiency groups on the scores associated with each survey item. The EFA measured the strength of the relationships between the proficiency stages (factors) and items using the percentage of correct and incorrect responses as indicator variables representing the NALO themes. Items that targeted EFA loading and frequency correct ranges were then examined with item analysis. This resulted in the construction of two 15-item assessment instruments. Each instrument contained three questions for each of the five NALO themes. Based on the allocated proficiency level, those three questions were staged from the



easiest to the most challenging item. Following EFA, we determined questions that were too easy, difficult, or poor and eliminated them based on frequency results. Item analyses were then conducted on items with varying frequencies, which were also used to determine if the EFA frequencies improved when specific poor answer choices were removed. We carefully ensured that option changes did not affect the question context.

Separately, we conducted a CFA for each of the two 15-item assessment instruments to determine if the model fit was adequate. For each instrument, each item was loaded on its assigned factor (proficiency level) to determine if the underlying correlational structure of the independent variables (the five NALO themes) represented each latent factor. Table 2 shows that the linear structural equation estimation indicated that both instruments fit adequately.

**Table 2**

*Confirmatory Factor Analysis Fit Summary Based on Total Correct Items*

Fit Summary	Instrument 1	Instrument 2
Chi-square $\chi^2$	131.80	124.26
Chi-square <i>df</i>	87	87
Variance estimate $\chi^2/df$	1.51	1.43
Adjusted Goodness-of-Fit (GFI)	.95	.96
RMSEA estimate	.03	.03
RMSEA lower 90% confidence limit	.02	.02
RMSEA upper 90% confidence limit	.04	.04
Bentler Comparative Fit Index	.94	.93
Bentler-Bonett Non-normed Index	.93	.92

*Note.* Root mean square error of approximation (RMSEA)

Additionally, the CFA analysis determined that indicator variables significantly loaded on their respective proficiency stage factor (all *p*-values below .001), indicating that differences between loadings and zero were significant. Collectively, it identified an almost non-existent shared variance among the variables—or a considerable amount of unique variance was seen among them.

The Cronbach's coefficient across proficiency stages was measured for Instrument I (*N* = 515): Exposure (Total  $\alpha$  = .46, Partial  $\alpha$  = .55); Literacy (Total  $\alpha$  = .58; Partial  $\alpha$  = .62); Proficiency (Total  $\alpha$  = .37; Partial  $\alpha$  = .65) and Instrument II (*N* = 515): Exposure (Total  $\alpha$  = .48, Partial  $\alpha$  = .50); Literacy (Total  $\alpha$  = .47; Partial  $\alpha$  = .54); Proficiency (Total  $\alpha$  = .29; Partial  $\alpha$  = .38). The reliability coefficients are low; however, Taber (2018) noted that alpha values vary greatly by discipline. Additionally, high reliability may indicate that items are redundant, and the length of the instrument (less than 20 items) limits the alpha and complicates the process of unpacking internal reliability. The partial scores have higher alpha measures because they have a greater



range of possible responses. They are relevant because they identify that when questions are not scored strictly right or wrong, they lead to a greater understanding of where respondent understanding is. The alpha numbers are likely low due to multiple themes for each factor. We corroborated these results with Pearson's product-moment correlation and Difficulty Indices. Results indicated an acceptable internal consistency and reliability level for both instruments because our goal was to produce non-redundant instrument items that could discriminate skill levels. Ultimately, the CFA showed enough evidence to substantiate the model as fitting adequately with a small or weak relationship between the proficiency stages.

### ***Discriminant Analysis & Summary***

DA was used to clarify the CFA results. Table 3 indicates that the cross-validation percentages for both instruments were extremely accurate and well within the range of  $p < .05$ . Equally strong re-substitution percentages were as good or better than the cross-validation results. The DA was the most definitive conclusion that the items aligned correctly for the five NALO themes, indicating that users can accurately administer either assessment to determine students' proficiency levels in agricultural literacy. Users, however, should not "mix and match" questions between instruments because both have been independently validated in this study. We concluded our analysis by finalizing items to construct two separate assessment instruments.

**Table 3**

#### *Discriminant Analysis: Cross-validation Summary Using Linear Discriminant Functions*

Proficiency Stage	<i>n</i>	Cross-validation %	Cross-validation error estimation
<b>Instrument 1</b>			
Exposure	74	97.37	.03
Factual Literacy	261	98.86	.01
Applicable Proficiency	175	100.0	.00
Total	515	100.0	.009*
<b>Instrument 2</b>			
Exposure	90	91.84	.08
Factual Literacy	317	97.24	.03
Applicable Proficiency	91	93.41	.00
Total	515	100.0	.04**

*Note.* \* $p < .01$ ,  $df = 514$ ; \*\* $p < .05$ ,  $df = 514$

### **Findings Summary**

The findings showed that both instruments were valid and reliable for measuring the 12th-grade NALO theme benchmarks and determining an agricultural literacy proficiency level. Future users need to understand how to use the instruments effectively.

### ***Using the Instruments***

Determining the proficiency level of a participant is an essential part of assessment analysis. Practitioners can identify the proficiency stages of the two instruments by listing participants with a score  $\geq 12$  (out of 15) as applicably proficient, those with a score of  $8 \leq 11$  as factually literate, and those  $\leq 7$  at the exposure level.

Scores can be interpreted individually or using a group's mean, median, or mode. Total correct scores are as helpful as partial correct scores. Partial correct scores can be obtained by examining individual assessments to determine which NALO items were incorrect, then using that information to identify gaps in thematic content, misinformation, or analysis related to experience or agricultural exposure. If the NALO themes are used for program achievement goals, and students do not show consistent growth across all five themes, the score can indicate curricula or instructional gaps. These instruments were designed to show cumulative assessment for K-12 agricultural literacy development. Ideally, students who have been instructed throughout their primary and secondary education should be applicably proficient at the end of twelfth grade. Proficiency levels that are less than ideal for high school graduates give educators and agricultural stakeholders information that can be used to understand where adult consumers may need additional information to make informed agricultural decisions. Furthermore, although this study sought to provide summative assessment, educators' use of the tools as a formative measurement is encouraged. Using the instruments formatively, in combination with a qualitative interview, could be the most exact way to determine how participants perceived or misperceived a correct answer.

## **Conclusions, Discussion, and Recommendations**

This study provided two standardized instruments that can measure agricultural literacy nationwide. There are now NALO-based assessments for elementary, middle, and high school students. By addressing the absence of a standardized tool for high school students and high school graduates, we fill a gap in existing literature and enhance the reliability, comparability, and applicability of future research in agricultural literacy. We recommend using these instruments to unify efforts to identify national knowledge gaps, better target educational initiatives, and increase study replication using consistent instrumentation. While not a comprehensive assessment, these instruments can impact how we implement and evaluate formal and nonformal agricultural education. Additionally, program planners and evaluators should use data from these assessments to determine the efficacy of their programs, hopefully leading to initiatives driven by program impacts rather than program outputs. Doerfert (2003) maintained that the true implications of agricultural literacy could only be seen as we study populations and programs over time. Instrument use within the same or similar programs can provide a roadmap of program efficacy that showcases which areas of agricultural literacy have improved over time and through which methods of instruction.

Practitioners should work with researchers to identify populations beyond K-12 students or formal classroom settings (Warnick, 2022). These instruments can improve learning opportunities for youth and adults in community-driven events associated with agritourism, 4-

H, community gardening, farmer's markets, and career awareness fairs. The length of the assessment makes it digitally accessible in a variety of environments via a smartphone and may open opportunities for greater discussions on agricultural topics with event participants. Additionally, Minkler and Salvatore (2012) outlined that collaborative research and evaluation processes contributed to greater success within community-engaged programs. Researchers working in tandem with communities can assess agricultural literacy and tailor interventions based on feedback and data from community members. Land grant institutions and Cooperative Extension can fulfill pivotal roles in enhancing their communities by leveraging their resources, expertise, and outreach capabilities toward agricultural literacy assessment. Using these tools at local levels provides data on knowledge and fosters relationships that can promote trust, limit misinformation about agriculture, and encourage informed consumer choices. By bridging the gap between research and communities, assessment data can contribute to the prosperity of agricultural sectors through programs that empower individuals to make informed decisions that positively impact their well-being and the broader community.

## Acknowledgments

**R. Judd-Murray** - formal analysis, investigation, writing- original draft and editing; **B. Warnick** - investigation, writing- review and editing; **D. Coster** - formal analysis, investigation, writing- review and editing; **M. Longhurst** - investigation, writing- review and editing.

## References

- American Association for Agricultural Education (AAAE). (2023). *Research values of the American Association for Agricultural Education*. [https://aaaeonline.org/resources/Documents/FOR%20ONLINE%20\(8.5%20x%2011\)%20-%20AAAE%20Research%20Values.pdf](https://aaaeonline.org/resources/Documents/FOR%20ONLINE%20(8.5%20x%2011)%20-%20AAAE%20Research%20Values.pdf)
- Bello, D., Leung, K., Radebaugh, L., Tung, R. L., & van Witteloostuijn, A. (2009). From the editors: Student samples in international business research. *Journal of International Business Studies*, 40(3), 361–364. <https://doi.org/10.1057/jibs.2008.101>
- Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, 31(4), 399–413. <https://doi.org/10.1080/02602930600679050>
- Brandt, M. R. (2016). *Exploring elementary students' agricultural and scientific knowledge using evidence-centered design* [Thesis]. University of Nebraska Lincoln. <https://digitalcommons.unl.edu/natresdiss/131/>
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Lawrence Erlbaum.

- Cosby, A., Manning, J., Power, D., & Harreveld, B. (2022). New decade, same concerns: A systematic review of agricultural literacy of school students. *Education Sciences*, 12(4), Article 4. <https://doi.org/10.3390/educsci12040235>
- Delbecq, A. L., Van de Ven, A. H., & Gustafson, D. H. (1975). *Group techniques for program planning*. Scott, Foresman, and Co.
- Doerfert, D. L. (2003). Agricultural literacy: An assessment of research studies published within the agricultural education profession. *Proceedings of the 22nd Annual Western Region Agricultural Education Research Conference*.
- Frick, M. J. (1993). Developing a national framework for a middle school agricultural education curriculum. *Journal of Agricultural Education*, 34(2), 77–84. <https://doi.org/10.5032/jae.1993.02077>
- Goodman, C. M. (1987). The Delphi technique: A critique. *Journal of Advanced Nursing*, 12(6), 726–734. <https://doi.org/10.1111/j.1365-2648.1987.tb01376.x>
- Hanel, P. H. P., & Vione, K. C. (2016). Do student samples provide an accurate estimate of the general public? *PLoS ONE*, 11(12), 1–10. <https://doi.org/10.1371/journal.pone.0168354>
- Hess, A. J., & Trexler, C. J. (2011). A qualitative study of agricultural literacy in urban youth: Understanding for democratic participation in renewing the agri-food system. *Journal of Agricultural Education*, 52(2), 151–162. <https://doi.org/10.5032/jae.2011.02151>
- Jacobs, J. M. (1996). *Essential assessment criteria for physical education teacher education programs: A Delphi study* [Doctoral dissertation, West Virginia University]. West Virginia University The Research Repository. <https://researchrepository.wvu.edu/cgi/viewcontent.cgi?article=10089&context=etd>
- Joplin, L. (1981). On defining experiential education. *Journal of Experiential Education*, 4(1), 17–20. <https://doi.org/10.1177/105382598100400104>
- Judd, R. C. (1972). The use of Delphi methods in higher education. *Technological Forecasting and Social Change*, 4(2), 173–186. [https://doi.org/10.1016/0040-1625\(72\)90013-3](https://doi.org/10.1016/0040-1625(72)90013-3)
- Lawson, A. E., & Weser, J. (1990). The rejection of nonscientific beliefs about life: Effects of instruction and reasoning skills. *Journal of Research in Science Teaching*, 27(6), 589–606. <https://doi.org/10.1002/tea.3660270608>
- Leising, J. G., Igo, C. G., Heald, A., Hubert, D., & Yamamoto, J. (1998). *A guide to food and fiber systems literacy*. W. K. Kellogg Foundation and Oklahoma State University.

- Leising, J. G., Pense, S. L., & Igo, C. (2000). An assessment of student agricultural literacy knowledge based on the food and fiber systems literacy framework. *Journal of Southern Agricultural Education Research*, 50(1), 146–151. <http://jsaer.org/category/journal/vol-50/>
- Longhurst, M. L., Judd-Murray, R., Coster, D. C., & Spielmaker, D. M. (2020). Measuring agricultural literacy: Grade 3-5 instrument development and validation. *Journal of Agricultural Education*, 61(2), 173–192. <https://doi.org/10.5032/jae.2020.02173>
- MacCallum, R. C., Widaman, K. F., Preacher, K. J., & Hong, S. (2001). Sample size in factor analysis: The role of model error. *Multivariate Behavioral Research*, 36(4), 611–637. [https://doi.org/10.1207/S15327906MBR3604\\_06](https://doi.org/10.1207/S15327906MBR3604_06)
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741. <https://doi.org/10.1037/0003-066X.50.9.741>
- Minkler, I. M., & Salvatore, A. L. (2012). *Study design and analysis in dissemination and implementation research* (2nd Edition). Oxford University Press.
- National Agriculture in the Classroom. (2014). *Agricultural literacy*. National Agriculture in the Classroom. <https://agclassroom.org/get/literacy/>
- National Center for Agricultural Literacy. (2017). Multistate research: W3006. Utah State University. <https://www.agliteracy.org/research/multistate/>
- National Research Council. (2009). *Transforming agricultural education for a changing world*. The National Academies Press. <https://doi.org/10.17226/12602>
- OECD: Programme for International Student Assessment. (2016). *PISA 2015: Technical report*. Organization for Economic Co-operation and Development (OECD). <http://www.oecd.org/pisa/sitedocument/PISA-2015-technical-report-final.pdf>
- Powell, D., Agnew, D., & Trexler, C. (2008). Agricultural literacy: Clarifying a vision for practical application. *Journal of Agricultural Education*, 49(1), 85–98. <https://doi.org/10.5032/jae.2008.01085>
- Redmond, E. C., & Griffith, C. J. (2003). Consumer food handling in the home: A review of food safety studies. *Journal of Food Protection*, 66(1), 130–161. <https://doi.org/10.4315/0362-028X-66.1.130>
- Roberts, T. G. (2006). A philosophical examination of experiential learning theory for agricultural educators. *Journal of Agricultural Education*, 47(1), 17–29. <https://doi.org/10.5032/jae.2006.01017>

- Sireci, S. G. (1998). The construct of content validity. *Social Indicators Research*, 45(1), 83–117. <https://doi.org/10.1023/A:1006985528729>
- Specht, A. R., McKim, B. R., & Rutherford, T. (2014). A little learning is dangerous: The influence of agricultural literacy and experience on young people's perceptions of agricultural imagery. *Journal of Applied Communications*, 98(3), 63–74. <https://doi.org/10.4148/1051-0834.1086>
- Spielmaker, D. M., & Leising, J. G. (2013). *National agricultural literacy outcomes*. Utah State University, School of Applied Sciences & Technology. <https://cdn.agclassroom.org/nat/data/get/NALObooklet.pdf>
- Spielmaker, D. M., Pastor, M., & Stewardson, D. M. (2014). A logic model for agricultural literacy programming. *Proceedings of the 41st annual meeting of the American Association for Agricultural Education*. National Center for Agricultural Literacy. <https://www.agliteracy.org/research/logic/>
- Stevens, C. K. (2011). Questions to consider when selecting student samples. *Journal of Supply Chain Management*, 47(3), 19–21. <https://doi.org/10.1111/j.1745-493X.2011.03233.x>
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273-1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Taylor, R. E., & Judd, L. L. (1989). Delphi method applied to tourism. In S. Witt & L. Moutinho (Eds.), *Gazing into the oracle: The Delphi method and its application to social policy and public health* (pp. 56–88). Jessica Kingsley Publishers.
- Warnick, B. K. (2022). *W3006: Multistate agricultural literacy research* [2020 Annual Report]. Utah State University. <https://www.nimss.org/projects/view/mrp/outline/18611>
- Winton, B. G., & Sabol, M. A. (2022). A multi-group analysis of convenience samples: Free, cheap, friendly, and fancy sources. *International Journal of Social Research Methodology*, 25(6), 861–876. <https://doi.org/10.1080/13645579.2021.1961187>

© 2024 by authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).