

Produce Safety Alliance Grower Training Knowledge Assessment Results

L. Narine¹, A. Enderton², M. Bengé³, E. Bihn⁴, S. Brown⁵, J. Kovacevic⁶, E. Newbold⁷, K. Schneider⁸, A. Shaw⁹

Abstract

The Produce Safety Alliance (PSA) Grower Training (GT) curriculum was developed to provide produce growers with training that meets requirements in §112.22(c) of the Food Safety Modernization Act (FSMA) Produce Safety Rule (PSR). Four regional food safety centers evaluated course attendees' knowledge change over four years (2019 to 2022) using a pre- and post-test quiz. Knowledge assessment results showed (a) respondents gained knowledge on each of the seven modules presented in the curriculum; (b) the curriculum content had a large effect on knowledge gain; (c) knowledge gain differed significantly between years, but the differences were not of practical importance; (d) remote participants learned significantly more than in-person participants, but differences were not of practical importance; and (e) the quiz consisted of low and moderate difficulty questions (no questions were high difficulty) and had generally acceptable discriminant properties. Implementation of the FSMA PSR has progressed from an educational to a regulatory phase, and the authors recommend replacing the knowledge assessment with a tool that measures how the PSA course prepares growers for compliance.

Article History










Received: February 16, 2024

Accepted: March 19, 2024

Published: April 16, 2024

Keywords

evaluation; fruits; vegetables; knowledge; assessment; grower training; Food Safety Modernization Act; produce safety; SDG 3: good health and wellbeing

-
1. Lendel Narine, Extension Associate Professor, Utah State University, 4800 Old Main Hill, Logan, UT 84322, lendel.narine@usu.edu,  <https://orcid.org/0000-0001-6962-2770>
 2. Arlene Enderton, Program Specialist, Iowa State University Extension and Outreach, 2625 N Loop Drive, Suite 2430, Ames, Iowa 50011, arlene@iastate.edu,  <https://orcid.org/0009-0004-7416-6203>
 3. Matt Bengé, Director, Program Development and Evaluation Center (PDEC), University of Florida, PO Box 112060, Gainesville, FL 32610, mattbenge@ufl.edu,  <https://orcid.org/0000-0002-5358-3233>
 4. Elizabeth Bihn, Director, Produce Safety Alliance, Cornell University, 665 W. North Street, Food Research Lab – NYSAES, Geneva, NY 14456, eab38@cornell.edu,  <https://orcid.org/0000-0001-5806-3491>
 5. Stephanie Brown, Food Safety Specialist, Oregon State University, 1207 NW Naito Parkway, Portland, OR 97209, stephanie.brown@oregonstate.edu,  <https://orcid.org/0000-0001-8682-9984>
 6. Jovana Kovacevic, Associate Professor and Extension Food Safety Specialist, Oregon State University, 1207 NW Naito Parkway, Portland, OR 97209, jovana.kovacevic@oregonstate.edu,  <https://orcid.org/0000-0003-1254-4294>
 7. Elizabeth Newbold, Assistant Director, NE Center to Advance Food Safety, University of Vermont, 310 Main St, PO Box 559, Bennington, VT 05201-0559, Elizabeth.Newbold@uvm.edu,  <https://orcid.org/0000-0002-0862-4593>
 8. Keith Schneider, Professor, University of Florida, 359 FSHN Bldg., Newell Dr., Gainesville, FL 32611, keiths29@ufl.edu,  <https://orcid.org/0000-0003-0145-3418>
 9. Angela Shaw, Professor, Texas Tech University, Box 42141, Lubbock, TX 79409-2141, Angela.Shaw@ttu.edu,  <https://orcid.org/0000-0001-5908-0429>

Introduction and Problem Statement

The Produce Safety Alliance (PSA) Grower Training (GT) curriculum was developed to provide produce growers with training that meets requirements in §112.22(c) of the Food Safety Modernization Act (FSMA) Produce Safety Rule (PSR). Concurrently, four regional centers, funded by United States Department of Agriculture (USDA) and United States Food and Drug Administration (FDA), were established to help develop networks of food safety professionals who could deliver the PSA GT courses to ensure training was widely available to produce growers, regulatory personnel, educators, and others. As of December 31, 2023, PSA Trainers have delivered 2,859 PSA GTs domestically since 2016, with 53,992 participants. Prior to the COVID-19 pandemic, PSA GTs were exclusively delivered in person. During the pandemic, real-time remote courses delivered by PSA Trainers using remote conference technologies was allowed. This remote delivery policy was temporary, but due to the popularity of the option, an assessment of training evaluations was conducted to ensure it was as effective a modality as in-person courses (Bugingo et al., 2023). In 2023, the remote course delivery policy became permanent, allowing PSA Trainers to teach PSA GTs in either course format.

To help evaluate the effectiveness of PSA GTs, pre-training (pre-test) and post-training (post-test) knowledge assessment tests were developed by the Southern Center (SC) for Food Safety Training, Outreach, and Technical Assistance to measure the immediate knowledge change of participants. The four regional centers have been collating these data from within their regions to gauge effectiveness, impact, and remaining training needs.

Voluntarily reported knowledge assessment data collected by the four regional centers from January 2019 to June 2022 are described herein. Data from in-person and remote delivery courses were analyzed to assess whether the PSA GTs resulted in short-term knowledge gain.

Theoretical and Conceptual Framework

Since the regional centers were launched, the Targeting Outcomes of Programs (TOP) model has been used to assess program performance. The TOP model, an expansion of Bennett's hierarchy (Bennett 1975; Bennett 1976), was first developed in 1994 to evaluate program outcomes in planning, implementation, and evaluation (Harder, 2009; Rockwell & Bennett, 2004). This model includes a two-sided hierarchy (program development and program performance), with seven levels shared between the two sides. These include (a) resources; (b) activities; (c) participation; (d) reactions; (e) the knowledge, attitudes, skills, and aspirations of participants (KASA); (f) practices; and (g) social, economic, and environmental conditions (SEE; Harder, 2009; Rockwell & Bennett, 2004). Notably, the model allows for multiple evaluation strategies (process and outcomes evaluation) to measure programmatic performance (one side of the hierarchy) (Harder, 2009; Rockwell & Bennett, 2004).

In an initial effort to evaluate knowledge gain from programs delivering standardized FSMA trainings (e.g., PSA GTs), a plan for sharing basic training information (e.g., average knowledge

assessment scores) at a national level was developed in collaboration with the Lead Regional Coordination Center and other regional centers. A common set of quantitative indicators (i.e., pre- and post-tests) were also created and used across trainings delivered by regional center partners. Although the same set of indicators were used, each center approached knowledge assessment and additional training-related data collection (for the KASA-related portion of the TOP model) differently. To start understanding the short-term training impacts of produce safety trainings on a national scale, this manuscript focuses on the evaluation of knowledge assessment data collected during trainings delivered from January 2019 to June 2022. Moving forward, these data will serve as guiding points to further the standardization of national evaluation efforts, including the assessment of medium- and long-term impacts, and re-evaluation of current quantitative indicators for PSA GTs. This framework will also serve as a guide for evaluating other standardized FSMA trainings provided by food safety professionals within the regional center networks (e.g., the Food Safety Preventive Controls Alliance Preventive Controls for Human Food participant course).

Purpose

The purpose of this study was to assess the short-term knowledge outcomes of the PSA GT course over a four-year period. The objectives were to (a) assess knowledge gain for each module of the PSA GT course, (b) assess overall knowledge gain by participants, (c) compare net knowledge gain across each year of the course, (d) examine differences in knowledge gain by delivery modality, and (e) assess the quality of each item in the knowledge assessment via the difficulty index and discrimination index. The study's results will be used to inform strategies for improving program implementation and optimizing program evaluation instruments.

Methods

This study examined primary data through quantitative methodology. Quantitative research is used to test theory through numerical evaluation by observing the relationships among variables (Ary et al., 2006; Creswell, 2014) and generates knowledge by examining phenomena affecting individuals (Allen, 2017). Secondary data is information that existed prior to a study, and it was not collected by the researcher solely for the purpose of their study (Stewart & Kamis, 1992). Zimmerman and Kahl (2018) explained that the collection of preexisting data can inform Extension programs and provide an increased understanding of the community by "putting individual Extension program impacts into a broader perspective." Preexisting PSA data were collected from the four regional centers for this study.

The PSA GT knowledge test contained 25 questions within seven curricular modules. Each test question consisted of four multiple-choice options, and PSA GT participants took the test before and after the training. Each PSA trainer either scored the tests and shared the pre- and post-knowledge test scores with their specific regional evaluation liaison or sent quizzes (electronically or via paper mail) to their regional evaluation liaison for data entry. Each regional

evaluation team holds its own Institutional Review Board approval, and demographic data were not collected.

Each regional evaluation team gathered their regional PSA knowledge assessment data collected from January 2019 through June 2022 for this study. The total number of cases collected across all four regions was 7,185. The aggregated regional data were reviewed, and cases with incomplete pre- or post-test scores were removed from the dataset, which yielded a final data set of 6,583 cases. For this study, the data were additionally coded for three independent variables: (a) year, (b) delivery modality, and (c) U.S. region. A paired samples *t*-test was used to address objectives (a) and (b), a one-way ANCOVA was used to answer objectives (c) and (d), and a difficulty index based on the proportion of correct responses and discrimination index via a point-biserial correlation (Millman & Green, 1989) were calculated for objective (e).

Findings

Objective (a): Module-Level Knowledge Gain

There are seven modules in the PSA GT course. Scores were standardized to the percent of correct responses in each module. Table 1 shows the average pre- and post-test scores across each module. There were statistically significant increases ($p < 0.001$) in participant scores between the pre- and post-test scores across all seven modules. Based on Cohen's *d*, the effect size between pre-and-post-test scores was small for module 2, moderate for modules 1 and 3, and large for 4, 5, 6, and 7.

Table 1

Module-Level Knowledge Scores

Module	<i>n</i>	%		<i>t</i>	(One-tailed) <i>p</i>	Cohen's <i>d</i>
		<i>Pre-test</i>	<i>Post-test</i>			
1	5,147	72	82	34.55	<0.001	0.48
2	5,147	94	97	10.20	<0.001	0.14
3	5,150	74	86	29.00	<0.001	0.40
4	5,161	54	77	48.66	<0.001	0.68
5	5,165	58	83	57.13	<0.001	0.80
6	5,138	52	77	53.85	<0.001	0.75
7	5,153	44	69	59.38	<0.001	0.83

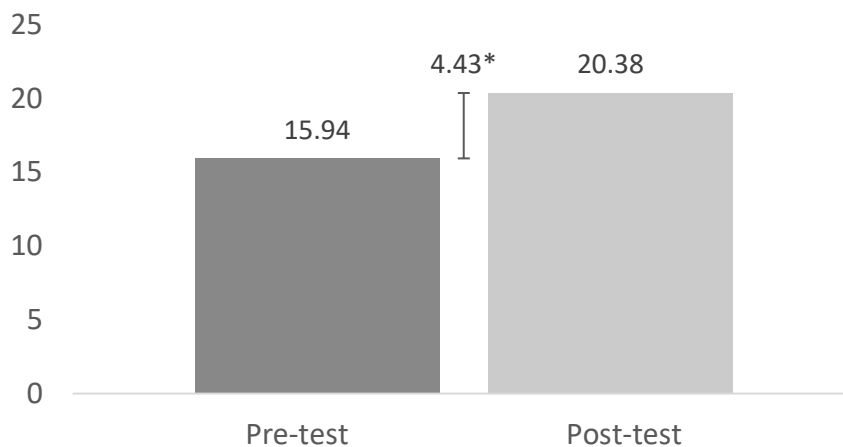
Note. The Southern Region did not gather module-level data. Table 1 excludes data from that region.

Objective (b): Overall Knowledge Gain

Participants demonstrated an increase in their knowledge of the PSR after completing the course. Based on a paired t -test, there was a statistically significant increase in participants' ($n = 6,583^*$) pre-test and post-test scores ($t = 108.39, p < 0.001$). On average, participants scored 15.94 ($SD = 3.62$) or 64% on the pre-test and 20.38 ($SD = 3.93$) or 82% on the post-test. Based on Cohen's d , the PSA course had a large effect on participants' knowledge of the PSR ($d = 1.34$).

Figure 1

Pre- and post-test knowledge from the PSA Course.



* The error bar represents the mean difference between groups.

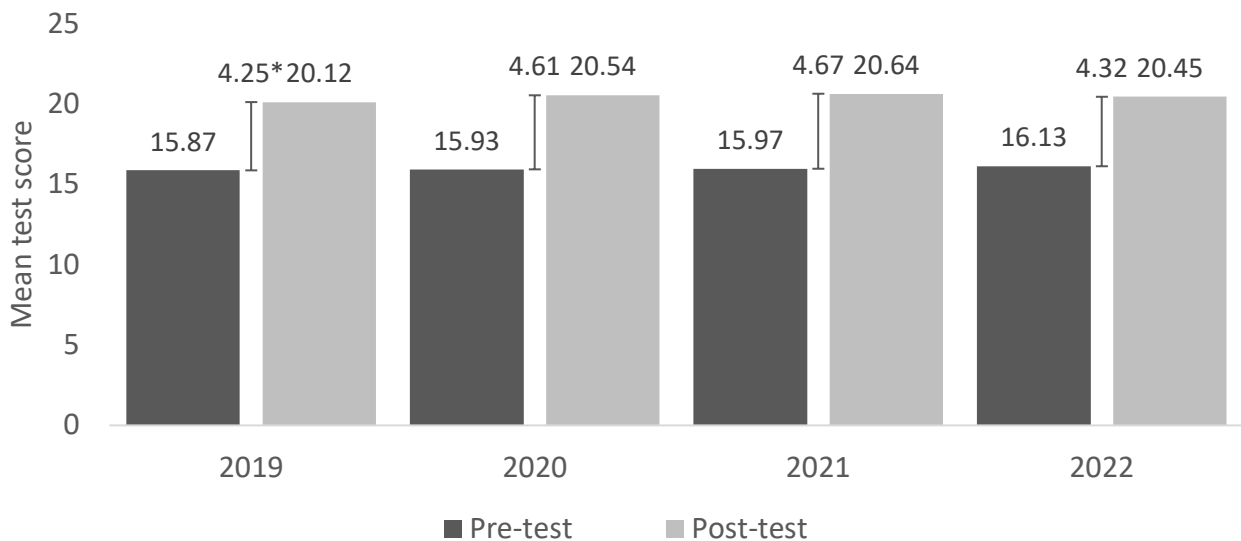
Objective (c): Net Knowledge Gain by Year

A one-way ANCOVA was used to estimate the net change in knowledge by program year while controlling for the effects of pre-test scores. The pre-test was statistically correlated with the post-test ($r = 0.617$), but the correlation was below the recommended threshold for its effect on the dependent variable ($r < 0.80$) and, therefore, served as a valid covariate in the model. Results indicated a statistically significant difference in mean knowledge change by year ($t = 10.34, p < 0.001$) when controlling for pre-test scores. However, the partial eta squared was very small ($\eta_p^2 = 0.005$), indicating that the differences in mean knowledge change across each program year was small and likely significant due to the large sample size. Therefore, while statistically significant, on a practical level, participants had similar knowledge gain between 2019 to 2022 when controlling for pre-test scores.

Mean post-test scores for each year were as follows; 2019 = 20.12 ($SD = 3.53, n = 2,635$); 2020 = 20.54 ($SD = 3.60, n = 1,720$); 2021 = 20.64 ($SD = 3.71, n = 1,263$); 2022 = 20.45 ($SD = 3.73, n = 695$). Figure 2 provides a comparison of pre-and-post-test scores between 2019 to 2022. Corroborating results of the ANCOVA model, Figure 2 shows that participants across each program year had similar steady levels of knowledge gain.

Figure 2

Change in knowledge scores between 2019 to 2022.



*The error bars represent the mean difference between groups.

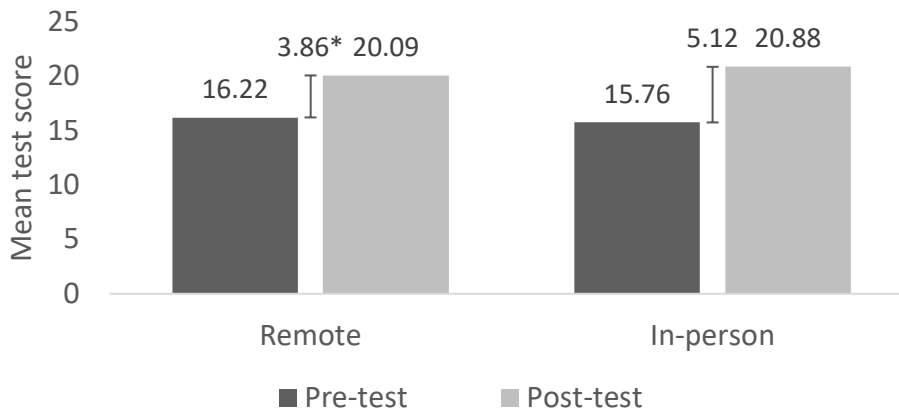
Objective (d): Knowledge Gain by Delivery Modality

Most participants attended the PSA GT course remotely in 2020 and 2021. A one-way ANCOVA was used to examine the differences in net knowledge gain by delivery modality while controlling for the effect of pre-test scores. There was a statistically significant difference in knowledge gain between remote and in-person delivery ($t = 51.65, p < 0.001$). Yet, the partial eta squared was small ($\eta_p^2 = 0.008$), indicating the difference in mean knowledge gain based on delivery modality was not practically significant and likely due to the large sample size. Therefore, while statistically significant, participants in remote and in-person classes had similar levels of knowledge gain.

The mean post-test score was 20.09 ($SD = 3.62$) for in-person participants ($n = 4,228$) and 20.88 for remote participants ($n = 2,305$) across all years. Figure 3 shows the pre- and post-test scores between in-person and remote participants. Consistent with findings from the ANCOVA model, remote and in-person participants experienced a similar level of knowledge gain from the PSA course as shown in Figure 3. While pre-test scores were controlled in the ANCOVA model, the figure shows remote participants entered the course with more content knowledge compared to in-person participants.

Figure 3

Pre- and post-test differences between delivery modes.



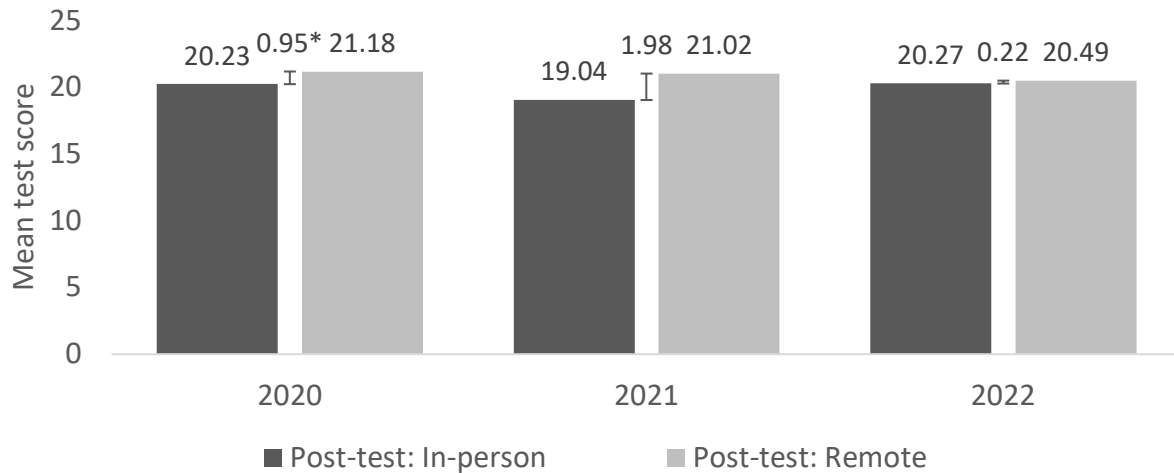
*The error bars represent the mean difference between groups.

Prior to March 2020, the PSA GT course was only delivered in person. In 2020, 67% of participants ($n = 1,302$) completed the PSA course in-person, while 33% completed it remotely ($n = 629$). In 2021, 22% of participants ($n = 294$) completed the course in-person and 78% completed it remotely ($n = 1,064$). In 2022, 18% ($n = 201$) completed the course in-person and 81% completed the course remotely ($n = 876$).

A two-way ANCOVA was conducted to assess the difference in net knowledge gain by delivery year (2020 to 2022) and delivery modality (in-person vs. remote) while accounting for the effects of pre-test scores. Results showed a weak but statistically significant difference in knowledge based on the interaction between delivery modality and program year ($t = 3.10$, $p < 0.01$, $\eta_p^2 = 0.001$). This suggests there were minor changes in knowledge gain for remote and in-person participants each year during the pandemic. Given a minor interaction effect, a Bonferroni post hoc test was used to identify statistically significant differences in knowledge based on the interaction between year and delivery modality. Using adjusted family-wise p-values, results showed remote participants in 2020 and 2021 scored significantly higher on the post-test (2020: $M = 21.18$, $SD = 3.27$; 2021: $M = 21.02$, $SD = 3.57$) compared to in-person participants (2020: $M = 20.23$, $SD = 3.72$; 2021: $M = 19.04$, $SD = 3.92$). However, remote and in-person participants experienced similar levels of knowledge gain in 2022 when controlling for the effects of pre-test scores. Figure 4 provides a descriptive summary of net knowledge gain by delivery modality and program years. Scoring a net two-point difference, remote participants performed noticeably better than in-person participants in 2021.

Figure 4

Net knowledge gain between remote and in-person participants during the first years of the pandemic.



*The error bars represent the mean difference between groups.

Objective (e): Item Difficulty and Discrimination Index

The PSA GT knowledge assessment contains 25 multiple-choice questions and was developed by a team of food safety specialists and Extension educators led by Catherine Shoulders of the University of Arkansas. The difficulty index and discrimination index were calculated using item-level data from participants who completed the PSA training between 2019 to 2022 in three out of the four regional centers ($n = 5,195$, excluding the Southern region). Table 2 shows the corresponding indices for all items in the assessment.

The Difficulty index represents the proportion of people who answered the item correctly. Ranging from 0 to 1, items with higher scores are easier, while lower scores are more difficult. From Table 2, 14 out of the 25 items (or 56% of items) were categorized as easy (> 0.80), while 11 were moderately difficult (0.30 to 0.80). No item had a high level of difficulty. The discrimination index measured the effectiveness of an item to distinguish between high and low performers on a scale of -1 to 1. An effective test item should provide sufficient evidence to differentiate between students who attained topic mastery and those who did not. Most items had a good discrimination index (> 0.30), while three had relatively low discrimination power (Q5, Q6, and Q16). Overall, while the items had generally acceptable discriminant properties, the test was easy for respondents.

Table 2*Item Analysis of the PSA GT Knowledge Assessment*

Item	Difficulty index	Discrimination index: Point-biserial correlation
Q1	.95	.35
Q2	.81	.39
Q3	.75	.45
Q4	.94	.36
Q5	.98	.19
Q6	.96	.29
Q7	.86	.43
Q8	.86	.47
Q9	.89	.41
Q10	.83	.37
Q11	.74	.51
Q12	.80	.44
Q13	.61	.37
Q14	.95	.31
Q15	.78	.43
Q16	.95	.28
Q17	.76	.43
Q18	.69	.39
Q19	.90	.39
Q20	.86	.46
Q21	.55	.44
Q22	.59	.46
Q23	.91	.32
Q24	.58	.34
Q25	.70	.51

Conclusions, Discussion, and Recommendations

The hierarchy that serves as the basis of the TOP model integrates evaluation into the overall program development process. The pre- and post-test tool was developed, in part, to assess the immediate knowledge gain of the PSA GT attendees. Looking at the results from pre- and post-tests nationally from 2019 – 2022, participants experienced an increase in their knowledge of the FSMA PSR after completing the course.

The key takeaway from these results is consistency in knowledge gain over the four-year period, regardless of delivery type or year of implementation. The differences in knowledge gain by implementation and year were statistically significant mainly because of the large sample size. While statistically significant, some differences (e.g., a 1.25-point difference in

knowledge change scores between remote and in-person participants) are not practically significant. For example, a 1- or 2-point difference suggests very little practical knowledge change, even though it is statistically different due to sample size and power, it is not practically different. Another important takeaway is the internal consistency within the dataset because the test was reliable in most cases, and there was an improvement in knowledge. Furthermore, the questions had sufficient discrimination power to differentiate between the low and top performers. Another key takeaway is that the questions were designed, in part, to measure immediate knowledge gain of produce safety concepts, and the practical significance of the data demonstrates that the objective has been met. Per the TOP model, there is a need to further the national evaluation model to assess medium- and long-term impacts of the PSA GT beyond knowledge gain and to measure synthesis and application of content. It is important to note the pre- and post-tests still have value in demonstrating program effectiveness. Trainers may continue to use the tool for other reporting purposes; regardless, the need for national aggregation of the data has been met.

The FSMA PSR has evolved since its inception in 2015. Some portions of the rule have been finalized, while requirements for pre-harvest agricultural water and application intervals of untreated biological soil amendments of animal origin (BSAAO) have not. FSMA PSR inspections began in 2018, with growers moving beyond the need for produce safety knowledge to a need to implement that knowledge on their farms. Seven years later (2023), the focus has shifted to ensuring growers know how to implement the PSR on their farms.

In conclusion, this research has shown that the initial objective of the assessment tool was met, and the PSA GT curriculum resulted in a significant knowledge gain in each module of the curriculum. With evolving educational needs based on PSR inspection findings and revised regulatory requirements, there is a need for a new assessment tool to measure knowledge change as well as growers' capacity to incorporate food safety behaviors and processes. Specifically, the research team suggests replacing this assessment tool and developing a new tool capable of answering the following questions: 1) are growers learning what they need to know about the rule? 2) are growers learning how to produce safe food? and 3) is the assessment tool meaningful for regulatory application and education effectiveness measures?

Acknowledgments

Thank you to Dr. Amy Harder, former evaluator for the Southern Center for Food Safety Training, Outreach, and Technical Assistance, for conceptualizing and leading efforts to develop the PSA GT knowledge assessment tool. Thanks to Donna Pahl Clements and Ellen Johnsen for reviewing the manuscript and suggesting edits. This work was funded by the Food Safety Outreach Program from the USDA National Institute of Food and Agriculture, agreement numbers 2021-70020-35740, 2021-70020-35732, 2021-70020-35753, and 2021-70020-35497.

L. Narine - data curation, formal analysis, funding acquisition, methodology, project administration, resources, visualization, writing-original draft, writing-review and editing; **A.**

Enderton - conceptualization, data curation, project administration, visualization, writing-original draft, writing-review and editing; **M. Bengé** – Data curation, writing-original draft; **E. Bihn** - writing-original draft, writing-review and editing; **S. Brown** - data curation, funding acquisition, methodology, project administration, resources, writing-original draft, writing-review and editing; **J. Kovacevic** - conceptualization, data curation, funding acquisition, methodology, project administration, resources, writing-original draft, writing-review and editing; **E. Newbold** - conceptualization, data curation, funding acquisition, investigation, project administration, resources, writing-original draft, writing-review and editing; **K. Schneider** - funding acquisition, project administration, writing-review and editing; **A. Shaw** - data curation, formal analysis, funding acquisition, investigation, methodology, resources, validation, visualization, writing-review and editing.

References

- Allen, M. (Ed.). (2017). *The SAGE encyclopedia of communication research methods*. Sage Publications, Inc. <https://doi.org/10.4135/9781483381411.n293>
- Ary, D., Jacobs, L. C., Razavieh, A., & Sorensen, C. (2006). *Introduction to research in education* (7th ed.). Thompson-Wadsworth.
- Bennett, C. (1975). Up the hierarchy. *Journal of Extension*, 13(2), 7-12. <https://archives.joe.org/joe/1975march/1975-2-a1.pdf>
- Bennett, C. (1976). *Analyzing impacts of extension programs (ESC-575)*. United States Department of Agriculture, Science, and Education Administration. <https://ia600708.us.archive.org/12/items/analyzingimpacts57benn/analyzingimpacts57benn.pdf>
- Bugingo, C., Stoeckel, D., Clements, D., George, L., Saunders, T., Blasini, D., Humiston, M., Melville, T., Way, R., Acuna-Maldonado, L., & Bihn, E. A. (2023). *Remote PSA grower training delivery policy: Lessons learned and recommendations for future courses*. Produce Safety Alliance. <https://resources.producesafetyalliance.cornell.edu/documents/Remote-PSA-Grower-Training-Delivery-Policy-Lessons-Learned-and-Recommendations-for-Future-Courses.pdf>
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Sage Publications, Inc.
- Harder, A. (2019). *Using the TOP model to measure program performance: A pocket reference* (WC092). University of Florida Department of Agricultural Education and Communication UF/IFAS Extension. <https://edis.ifas.ufl.edu/publication/WC092>

Millman, J., & Green, J. (1989). The specification and development of tests of achievement and ability. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 335-366). Macmillan Publishing Co, Inc; American Council on Education.

Rockwell, K., & Bennett, C. (2004). Targeting outcomes of programs: A hierarchy for targeting outcomes and evaluating their achievement. *Agricultural Leadership, Education, and Communication Department: Faculty Publications*, 48.
https://digitalcommons.unl.edu/aglecfacpub/48/?utm_source=digitalcommons.unl.edu%2Faglecfacpub%2F48&utm_medium=PDF&utm_campaign=PDFCoverPages

Stewart, D. W., & Kamins, M. A. (1993). *Secondary research: Information sources and methods* (2nd ed., Vol. 4). Sage Publications, Inc.

Zimmerman, J. N., & Kahl, D. (2018). Finding publicly available data for extension planning and programming: Developing community portraits. *The Journal of Extension*, 56(3), Article 1. <https://doi.org/10.34068/joe.56.03.01>

© 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/>).