2019

Development and Testing of Program Evaluation Instruments for the iCook 4-H Curriculum

Douglas R. Mathews
University of Maine

Zachary J. Kunicki
University of Rhode Island

Sarah E. Colby
University of Tennessee, Knoxville

Lisa Franzen-Castle
University of Nebraska-Lincoln

Kendra K. Kattelmann
South Dakota State University

See next page for additional authors

Research Article

Development and Testing of Program Evaluation Instruments for the *iCook 4-H* Curriculum

Douglas R. Mathews, PhD, RD[1,†]; Zachary J. Kunicki, PhD, MS[2]; Sarah E. Colby, PhD, RD[3]; Lisa Franzen-Castle, PhD, RD[4]; Kendra K. Kattelmann, PhD, RDN, LN, FAND[5]; Melissa D. Olfert, DrPH, RDN[6]; Adrienne A. White, PhD, RDN, FAND[1]

ABSTRACT

Objective: To develop and test the validity of program outcome evaluation instruments for cooking, eating, and playing together for obesity prevention during *iCook 4-H*.

Design: Instrument development for both youth and adults through pre-post testing of items newly constructed and compiled to address key curriculum constructs. Testing occurred throughout program intervention and dissemination to determine dimensionality, internal consistency and test-retest reliability, and validity.

Setting: A 5-state out-of-school program in cooperative extension and other community sites.

Participants: Youths aged 9−10 years; adults were main food preparers; the first phase involved 214 dyads and the second phase, 74 dyads.

Main Outcome Measure(s): Youth measures were cooking skills, culinary self-efficacy, physical activity, and openness to new foods. Adult measures were cooking together, physical activity, and eating together.

Analysis: Exploratory factor analysis to determine initial scale structure and confirmatory factor analysis to confirm factor structures. Longitudinal invariance tests to see whether the factor structure held over time. Test-retest reliability was determined by Pearson *r* and internal consistency was determined by coefficient Ω and Cronbach α. Validity testing was determined by Pearson *r* correlations.

Results: Youth cooking skills, openness to new foods, and adult eating together and cooking together showed strong evidence for dimensionality, reliability, and validity. Youth physical activity and adult physical activity measures showed strong evidence for dimensionality and validity but not reliability. The youth culinary self-efficacy measure showed strong evidence for reliability and validity but weaker evidence for dimensionality.

Conclusions and Implications: Program outcome evaluation instruments for youths and adults were developed and tested to accompany the *iCook 4-H* curriculum. Program leaders, stakeholders, and administrators may monitor outcomes within and across programs and generate consistent reporting.

Key Words: cooking, dyad interventions, *iCook 4-H*, program evaluation, youth (*J Nutr Educ Behav.* 2019; 51:S21−S29.)

Accepted October 24, 2018. Published online November 22, 2018.

INTRODUCTION

iCook 4-H was a 6-year study for youth (aged 9−10 years) and adult (primary meal preparer) dyads that included a control−treatment 2-year intervention study with assessments at 0, 4, 12, and 24 months followed by a test of dissemination with assessments at 0 and 4 months.[1] During the intervention phase, the treatment group participated in a 12-week, 6-session, face-to face program curriculum for dyads to cook, eat, and play together, with additional activities between months 4 and 24. The curriculum,[2] which was based on the Social Cognitive Theory (SCT)[3] and the 4-H experiential learning approach,[4] was developed for out-of-school youth programming, primarily within 4-H/extension with

[1]School of Food and Agriculture, University of Maine, Orono, ME
[2]Department of Psychology, University of Rhode Island, Kingston, RI
[3]Department of Nutrition, University of Tennessee, Knoxville, TN
[4]Nutrition and Health Sciences Department, University of Nebraska-Lincoln, Lincoln, NE
[5]Department of Health and Nutritional Sciences, South Dakota State University, Brookings, SD
[6]Division of Animal and Nutritional Sciences, Davis College of Agriculture, Natural Resources, and Design, West Virginia University, Morgantown, WV
†Dr Mathews was a graduate student at the time this study was conducted.

Conflict of Interest Disclosure: The authors have not stated any conflicts of interest.
The publication of this supplement to the *Journal of Nutrition Education and Behavior,* including this article, was supported by the National Institute of Food and Agriculture, US Department of Agriculture, under award number 2012-68001-19605. This article underwent the usual peer-review process followed by the journal.

Address for correspondence: Adrienne A. White, PhD, RDN, FAND, School of Food and Agriculture, University of Maine, 5735 Hitchner Hall, Orono, ME 04469-5735; E-mail: awhite@maine.edu

https://doi.org/10.1016/j.jneb.2018.10.014

the ability to adapt to other program venues. The SCT was used because of the emphasis of *iCook 4-H* on having an impact on personal, environmental, and behavioral factors with a focus on reciprocal role modeling. The overarching goals were improving culinary skills, the occurrence of family meals, physical activity for both youths and adults, and goal-setting behavior in youth. The dissemination phase was designed to test the program transition from a research to a community setting with minimal assistance from the researchers.[1] For the dissemination phase, the curriculum was modified to be 14 weeks with 8 sessions.[2] The study was implemented at the 5 land-grant universities in Maine, Nebraska, South Dakota, Tennessee, and West Virginia.

Evaluation instruments surrounding the 3 core components of the *iCook 4-H* curriculum (cooking, eating, and playing) were designed and tested during the intervention phase and further confirmed during the dissemination phase. The goal was to have tested evaluation measures as recommended by researchers[5] that were designed specifically for the curriculum when the *iCook 4-H* program was published for national distribution. The objective of this study was to describe the development and testing of the youth and adult program instruments that occurred over the intervention and dissemination study phases.

METHODS

Study Design

The program evaluation instruments were developed in an online survey format (Qualtrics, Provo, UT; 2013). Instrument development occurred at the baseline and 4-month assessments during the intervention phase. Final instrument testing was completed at baseline during the dissemination phase. The instruments were developed in a Likert-style response format. The researchers constructed survey items to address cooking, eating, and playing together, which were the most important constructs in the curriculum. Several statistical techniques were conducted to develop the evaluation instruments during the intervention and dissemination phases. The researchers used exploratory factor analysis (EFA) to determine item inclusion and potential subscales; the factor structure was further confirmed using confirmatory factor analysis (CFA). The recommended coefficient Ω was used for internal consistency; Cronbach α was also provided.[6,7] In addition, test-retest reliability was calculated for consistency over baseline and 4 months using Pearson r correlation.[8] During the intervention phase, sample sizes were 214 control and treatment dyads at baseline and 54 control group dyads at month 4. Control and treatment group dyads were analyzed separately for test-retest reliability to avoid bias of the treatment condition. For the dissemination phase, sample sizes were 74 youths and 76 adults. The dissemination phase was used as a second independent sample to conduct the CFAs, for which minimum sample sizes of n = 200 were recommended.[9] Owing to the sample size issue, the researchers decided to analyze the dissemination data separately for youths and adults (ie, incomplete dyads were used opposite the procedure employed for the intervention phase data) to obtain as large a sample size as possible. The dissemination study followed the same procedures as the intervention study, but only the baseline data were used in the analyses.

Protection of Human Subjects

The iCook 4-H Study procedures were approved at each phase of the study by the Institutional Review Boards for the Protection of Human Subjects at all five universities associated with the project.

Study Participants and Recruitment

Sample 1: Intervention phase. Dyads (n = 228) consisted of youths (mean age, 9.4 [SD, 0.7] years) and their primary adult meal preparers (mean age, 39 [SD, 8] years). Reported youth demographics were 55% female and 63% white, 14% Hispanic, 12% black, and 13% other. Of the adults, 43% had a high school degree or less and 40% reported participating in food assistance programs such as the *Supplemental Nutrition Assistance Program*. After data cleaning for program evaluation data, 214 dyads were included in instrument development analysis. Dyads were recruited between May and August, 2013 using standardized materials, which included flyers targeted directly to 4-H youth and for use in schools, clinics, and grocery stores, scripts for media sources, and a Facebook page with ads targeted to cities where the study was occurring. All materials included the study purpose, time commitment, eligibility criteria, and participant incentive information, which was $10/youth and adult at the baseline and 4-month assessments. Eligibility criteria were that youths were aged 9–10 years; had no dietary restrictions, food allergies, or activity-related medical conditions; and had access to a computer with the Internet.

Sample 2: Dissemination phase. Participants consisted of youths (n = 74) (control = 39 [53%] and treatment = 35 [47%]), who were mainly female (68%) (mean age, 9.5 [SD, 0.8] years), and their primary adult (n = 76) meal preparers, who were female (96%) (mean age, 38 [SD, 6.6] years). Of the adults, 77% were married, 71% were employed, 30% had a high school degree or less, and 26% reported participating in food assistance programs such as the *Supplemental Nutrition Assistance Program*. While treatment dyads completed the survey instruments at the *iCook 4-H* sessions, control adults were e-mailed survey links with instructions about completing the surveys and assisting youth, as needed, if their youths had not completed the survey while in the 4-H programs (non-*iCook* program). Youth and adults each received stipends of $10 at the 2 assessments.

Development of instruments. The SCT informed the development of the *iCook 4-H* curriculum.[2] Sessions designed for youths and adults to cook, eat, and play together were developed to provide opportunities for observational learning and reciprocal role-modeling, along with

building behavioral capability, culinary self-efficacy, and reinforcement of behavior both in the sessions and at home between sessions. Based on a thorough review of the curriculum, these constructs were carried over into the development of the program evaluation instruments, so that the key components of cooking, eating, and playing behavior, role modeling, culinary self-efficacy, and family activities within the home environment provided the basis for items. The researchers also conducted a careful review of the literature.[10,11] The research team, including cooperative extension faculty, statistical consultants, and key stake holders, assisted with item identification. The goal of the instruments was to measure change in *iCook 4-H* focal areas that were built on the tenets of the SCT. For youths, the constructs were cooking skills, openness to new foods, culinary self-efficacy, family mealtime, physical activity, and goal setting. For adults, the constructs were cooking with youths, shopping with youths, family meals, and physical activity. The researchers pretested items with youths to ensure comprehension and conducted pilot-testing. For the youth instrument, response options were based on 1 of 3 5-point Likert scales to test (1) skills, by asking *Can you* …, ranging from 1 = never to 5 = always; (2) openness to new foods, by asking *How willing are you* …, ranging from 1 = very unwilling to 5 = very willing; and (3) culinary self-efficacy, by asking *I am sure* … The Likert scale for self-efficacy questions ranged from 1 = strongly agree to 5 =strongly disagree and was reverse-coded for analysis. For the adult instrument, response options were based on a Likert scale ranging from 1 = never to 5 = always with focal areas of cooking eating and playing together.

Statistical analysis. Using a classical test theory approach,[8] the youth and adult measures were tested for dimensionality, reliability, and validity. Tests for dimensionality were conducted using a minimum average partial (MAP) test to guide the number of factors to extract using EFA.[12,13] The goal-setting measure was not included in testing because it had only 2 items (*How often do you set healthy goals for yourself?* and … *meet healthy goals?*), but the items were retained because of their importance in the curriculum. If >1 factor was recommended to be extracted by the MAP test, Promax rotation was used because it was expected that there would be correlations among the factors.[9] Factor extraction was done using the maximum likelihood method; if any item did not load on a factor above |0.29| or was a complex loading of ≥|0.29| on ≥1 factor, the item was dropped and the EFA was reconducted.[9,14,15] Based on the EFA results, CFA was used to validate the dimensionality of the sample further. For the CFA, fit indices of a nonsignificant χ^2 test, χ^2–degree of freedom (DF) ratio < 5.0 for acceptable fit, 3.0 for better or 2.0 for best fit, comparative fit index (CFI), and root mean square error of approximation (RMSEA) < 0.10 for acceptable, 0.08 for better, and 0.05 for best fit were all consulted.[9,16–19] However, the χ^2 is highly sensitive and a significant finding does not necessarily mean an unacceptable model fit.[9,15] After checking model fit indices, *z*-tests were used to see whether the items significantly loaded onto the factors. In the CFA models, the factor loading of the first item was fixed to 1.0 to identify the metric used for the remaining factor loadings.[9,16] Moreover, longitudinal invariance testing was conducted by fitting the same CFA model at baseline and 4 months and specifying correlations between the factors across time points as well.[16] Longitudinal invariance tests were also interpreted using the χ^2 test, χ^2-DF ratio, CFI, and RMSEA indices. Reliability estimates were calculated using coefficient Ω, Cronbach α, and test-retest reliability at baseline and 4 months using Pearson *r* correlations.[6,7] Values > 0.70 indicate ideal internal consistency reliability shown by Ω and α. However, because there was an intervention component to this study, the researchers did not expect test-retest reliability values > 0.70 to be found in the intervention sample, because the goal of the intervention was to change the constructs under study. The final step of the analyses was to create a correlation matrix among the validated measures to test for validity. It was expected that related scales (eg, the youth and adult physical activity measures) would be significantly positively correlated, which would provide evidence for convergent validity, whereas unrelated measures (eg, the youth self-efficacy measure and youth eating openness to new foods measure) would have no significant relation, providing evidence for discriminant validity. All analyses were conducted in R (version 3.5.1, R Core Team, Vienna, Austria, 2017) using the lavaan, MissMech, and psych packages.

RESULTS

The first step of the analyses was to conduct missing data diagnostics. In the intervention sample, the amount of missing data was minor (1.5%), and Little's missing completely at random (MCAR) was not significant, which provided evidence that the data were MCAR. Thus, all missing data were imputed using the maximum likelihood imputation with the expectation-maximization algorithm for the EFA and reliability analyses and full-information maximum likelihood for longitudinal invariance testing. After missing data diagnostics, the researchers checked assumptions. No issues of nonnormality emerged. In the dissemination sample, the results of missing data diagnostics also suggested that the data were MCAR; similar approaches to the intervention data were used for the missing data using full-information maximum likelihood imputation for the CFA.

The second step of the analyses was to calculate a MAP test to determine the number of factors to extract for the subsequent EFAs. In all, from 5 youth measures (cooking skills by yourself [7 items], cooking skills with help [7 items], physical activity [3 items], openness to new foods [3 items], culinary self-efficacy [6 items], and eating together [5 items]) the MAP test suggested extracting a single factor. Similarly, in all 3 adult scales, cooking together (5 items), eating together (3 items), and physical activity (3 items), the MAP test suggested extracting a single factor as well. Because of the size of the

Table 1. Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) Results: Youth Cooking Skills Measure

Items	Loadings		
	By Yourself: EFA (Intervention) (n = 214)	With Help: EFA (Intervention) (n = 214)	No Modifier[b]: CFA (Dissemination) (n = 74)
Can you cook foods to the right temperature . . .?	0.72	0.82	0.54[a]
Can you use an oven for cooking . . .?	0.67	0.68	0.73[a]
Can you use herbs and spices when cooking . . .?	0.60	0.81	0.39[a]
Can you use a blender . . .?	0.59	0.62	0.66[a]
Can you measure ingredients for a recipe . . .?	0.51	0.73	0.56[a]
Can you use a knife to cut foods . . .?	0.48	0.55	0.65[b]
Can you store foods the right way . . .?	0.40	0.58	0.28[a]

[a]$P < .001$ for the CFA *z*-tests, confirming factor loading; [b]Fixed for model identification.
Note: Modifiers *by yourself* and *with help* were dropped for items in dissemination for response burden.

dissemination sample, this analysis was done only on the intervention sample.

The third step of the analyses was to conduct an EFA for all 5 youth scales and all 3 adult scales, again using only the intervention sample. All EFA loadings were extracted by the maximum likelihood method; no rotation was done because the MAP tests suggested extracting a single factor. If any item loaded <|0.30|, the item was dropped and the EFA was reconducted.[9] Starting with the youth measures, in the cooking skills by yourself measure, EFA showed that all 7 items loaded >|0.29|, the eigenvalue was 2.96, and the factor structure explained 33% of the variance. Table 1 shows the loadings. The cooking skills with help measure also retained all items, the eigenvalue was 3.86, and the results explained 48% of the variance, with loadings shown in Table 1. In the physical activity measure, all items were retained, the eigenvalue was 1.29, and the factor explained 15% of the variance. The loadings are shown in Table 2. In the openness to new foods measure, all items were retained, the eigenvalue was 2.11, and the factor explained 56% of the variance. The loadings are shown in Table 3. In the culinary self-efficacy measure, all items were retained, the eigenvalue was 3.28, and the single factor explained 46% of the variance. Table 4 shows the loadings. For the final youth measure, eating together, based on the EFA results, these items were dropped owing to low loadings of −0.02, 0.19, and 0.29, respectively: *How often is it stressful to eat as a family? How* often *do you help shop for groceries?* and *How often do you help cook meals?* Because eating together resulted in only 2 items (*How often do you eat as a family?* and *Eat at a table with no distractions?*) no further analyses were completed on this measure. In the adult measures, the EFA on the physical activity measure suggested retaining all items and an eigenvalue of 1.66, and the factor explained 33%

Table 2. Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA): Youth Self-Efficacy and Adult Cooking Together Measures

Items	Loadings	
	EFA (Intervention) (n = 214 Youths, Adults)	CFA (Dissemination) (n = 74 Youths, 76 Adults)
Youth culinary self-efficacy measure		
I am sure I can use a stovetop.	0.82	0.77[a]
I am sure I can use an oven.	0.78	0.73[a]
I am sure I can make food safely to avoid getting sick.	0.65	0.53[a]
I am sure I can follow a recipe.	0.64	0.37[a]
I am sure I can cook.	0.59	0.60[b]
I am sure I can use a knife safely.	0.54	0.32[a]
Adult cooking together measure		
How often does your child help you cook meals?	0.68	0.72[b]
How often do you enjoy making meals with your child?	0.64	0.85[a]
How often do you enjoy making meals?	0.50	0.57[a]
How often do you feel confident with your kitchen skills?	0.44	0.14
How often does your child help in meal planning?	0.34	0.52[a]

[a]$P < .001$ for the CFA *z*-tests, confirming factor loading; [b]Fixed for model identification.

Table 3. Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA): Youth Openness to New Foods and Adult Eating Together Measures

Items	Loadings	
	EFA (Intervention) (n = 214 Youths and Adults)	CFA (Dissemination) (n = 74 Youths and 76 adults)
Youth openness to new foods measure		
How willing are you to try foods in new and interesting ways?	0.82	0.76[a]
How willing are you to taste new foods you have not tried?	0.79	0.83[b]
How willing are you to cook new foods that you have not tried?	0.63	0.55[a]
Adult eating together		
How often do you make eating together a family a priority?	0.92	1.00[b]
How often does your family eat together each week?	0.75	0.63[b]
How often do the topics of conversation at mealtimes include all family members?	0.58	0.37[a]

[a]$P < .001$ for the CFA *z*-tests, confirming factor loading; [b]Fixed for model identification.

of the variance. Loadings are shown in Table 4. The eating together scale EFA also suggested retaining all items, it had an eigenvalue of 2.11, and the factor explained 58% of the variance. Results are shown in Table 3. The final adult measure was on cooking together; based on the initial EFA, 4 items were dropped (using a grocery list [weekly meal planning], *Does child help shop for groceries?* and preferring to eat out over cooking, owing to loadings <|0.30|. In the second iteration, the remaining 5 items were retained, the eigenvalue was 2.10, and the factor explained 29% of the variance. The results are shown in Table 2.

The fourth step of the analyses was to conduct CFAs on all of the measures based on the EFA results, but using the dissemination control/treatment dataset (n = 74 youths and 76 adults) to verify the dimensionality findings in a second independent sample. However, because the cooking skills by yourself and cooking skills with help youth measures were worded so similarly, to reduce participant burden, these measures were collapsed into a single measure and reworded to ask *Can you . . .?* with no modifier of . . . *by yourself* or . . . *with help* at the end of each item. Thus, the CFA analysis was still guided by the EFA results, but only 1 set of loadings is displayed in Tables 1–5.

Starting with the youth measures, the cooking skill measure showed acceptable fit to the data based on the χ^2-DF ratio, CFI, and RMSEA values. Table 1 shows standardized loadings. The physical activity CFA showed great fit to the data based on the χ^2-DF ratio and CFI values, but the RMSEA value was above the acceptable cutoff guideline of 0.10. In the physical activity measure, the factor variance was fixed to 1.0 to provide 1 DF for model identification. Standardized loadings are shown in Table 2. The openness to new foods CFA also showed great fit to

Table 4. Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA): Physical Activity Measures for Youths and Adults

Items	Loadings	
	EFA (Intervention) (n = 214 Youths and Adults)	CFA (Dissemination) (n = 74 Youths and 76 Adults)
Youth physical activity measure		
When you think about each day of the week, how often are you physically active for at least 60 min/d?	0.45	0.94[c]
When you think about each day of the week, how often does your heart pump hard and you sweat when you are being physically active?	0.35	0.30[a]
How often does your family play actively together?	0.35	0.32[a]
Adult physical activity measure		
When you think about each day of the week, how often is your child physically active for at least 60 min/d?	0.62	1.00[c]
When you think about each day of the week, how often are you physically active for at least 30 min/d?	0.59	0.27[b]
How often does your family actively play together?	0.51	0.49[a]

[a]$P < .001$ for the CFA *z*-tests; [b]$P < .05$; [c]Fixed for model identification.

Table 5. Confirmatory Factor Analysis Fit Indices to Test Dimensionality

Measure	χ^2	Degrees of Freedom	χ^2– Degrees of Freedom Ratio	Comparative Fit Index	Root Mean Square Error of Approximation (90% Confidence Interval)
Youth measures					
Cooking skills	34.86[b]	14	2.49	0.90	0.10 (0.06–0.14)
Physical activity	2.08	1	2.08	0.91	0.12 (0.00–0.36)
Openness to new foods	0.31	1	0.31	1.00	0.00 (0.00–0.25)
Culinary self-efficacy	22.51[a]	9	2.50	0.85	0.14 (0.07–0.22)
Adult measures					
Physical activity	12.64[b]	1	12.64	0.67	0.39 (0.22–0.60)
Eating together	1.87	1	1.87	0.98	0.11 (0.00–0.35)
Cooking together	9.02	5	1.80	0.95	0.10 (0.00–0.21)

[a] $P < .01$; [b] $P < .001$.

the data based on all fit indices; it also had the factor variance fixed to 1.0 for model identification. Table 3 lists standardized loadings. The culinary self-efficacy CFA showed acceptable fit based on the χ^2-DF ratio but it did not achieve acceptable fit based on the CFI and RMSEA values Standardized loadings for the culinary self-efficacy CFA are provided in Table 4. Table 6 lists the fit indices for all CFA models.

For the adult measures, the physical activity model showed a relatively unacceptable fit to the data based on the χ^2-DF ratio, CFI, and RMSEA values. In the physical activity CFA, the variance of the question *How often does your family actively play together?* was fixed to 0 for model identification owing to a nonsignificant, negative variance for item in the initial model. The standardized loadings are in Table 5. The eating together CFA showed great fit to the data based on the χ^2-DF ratio and CFI values, but the RMSEA was above the acceptable fit cutoff of 0.10.[16,19,20] The eating together CFA had the variance of the question *How often do you make eating together as a family a priority?* fixed to 0 for model identification owing to a negative variance in the initial model again; the loadings are shown in Table 3. Finally, the cooking together CFA showed an acceptable fit to the data based on all fit indices reaching the acceptable cutoff guidelines. Table 2 displays the loadings. Table 6 provides a summary of fit indices for all of the CFA models.

The fifth step of the analyses was testing for longitudinal invariance to examine whether the factor structure was consistent over time. The baseline assessments for the youth and adult measures were compared with assessments at 4 months. Longitudinal invariance testing was done on the intervention sample because the sample size of the dissemination sample was too small to justify the complex analyses. Among the youth measures, the cooking skills and openness to new food measures showed evidence for longitudinal invariance, but the remaining measures did not. Among the adult measures, the eating together measure showed evidence for longitudinal invariance but the physical activity and cooking together measures did not. Table 6 lists the fit indices.

The sixth step of the analyses was to test for internal consistency and test-retest reliability (Table 7). Internal consistency testing was done by coefficient Ω and Cronbach α. The researchers

Table 6. Longitudinal Invariance (Consistency Over Time) Fit Indices

Measure	χ^2	Degrees of Freedom	χ^2– Degrees of Freedom Ratio	Comparative Fit Index	Root Mean Square Error of Approximation (90% Confidence Interval)
Youth measures					
Cooking skills (by yourself)	182.93[a]	76	2.41	0.86	0.08 (0.07–0.10)
Cooking skills (with help)	154.67[a]	76	2.04	0.94	0.07 (0.05–0.09)
Physical activity	41.80[a]	8	5.23	0.56	0.14 (0.10–0.18)
Openness to new foods	14.15	8	1.77	0.99	0.06 (0.00–0.11)
Cooking self-efficacy	210.50[a]	53	3.97	0.83	0.12 (0.10–0.14)
Adult measures					
Physical activity	78.61[a]	8	9.83	0.71	0.20 (0.16–0.25)
Eating together	29.45[a]	8	3.68	0.95	0.11 (0.70–0.16)
Cooking together	214.29[a]	34	6.30	0.65	0.16 (0.14–0.18)

[a] $P < .001$.

Table 7. Reliability Results for Youth and Adult Measures

Measure	Ω	Cronbach α	Test-Retest (Treatment)	Test-Retest (Control)
Youth measures				
Cooking skills (by yourself)	0.83	.77	0.65 (0.54–0.74)	0.74 (0.61–0.83)
Cooking skills (with help)	0.90	.86	0.52 (0.39–0.63)	0.65 (0.49–0.76)
Physical activity	0.31	.34	0.32 (0.17–0.46)	0.38 (0.16–0.56)
Openness to new foods	0.77	.79	0.48 (0.34–0.60)	0.47 (0.27–0.63)
Cooking self-efficacy	0.88	.83	0.40 (0.26–0.53)	0.67 (0.51–0.78)
Adult measures				
Physical activity	0.56	.59	0.63 (0.51–0.72)	0.57 (0.39–0.71)
Eating together	0.77	.79	0.46 (0.32–0.58)	0.70 (0.56–0.80)
Cooking together	0.73	.62	0.66 (0.55–0.74)	0.63 (0.47–0.76)

Note: All test-retest correlations were significant at $P < .001$.

calculated test-retest reliability using Pearson *r* correlation with 95% confidence intervals separately between the intervention treatment and control groups at baseline and 4 months. The youth measures had acceptable reliability except for the physical activity and eating together measures. Similarly, the adult physical activity measure did not achieve acceptable internal consistency reliability levels, but it showed good test-retest reliability. The remaining 2 adult measures had good reliability results.

The final step in the analysis was to construct a correlation matrix among all of the measures in the intervention sample to provide evidence for convergent and/or discriminant validity. The results are shown in Table 8. The youth cooking skills measure showed evidence for convergent validity between the 2 forms of the measure and weaker evidence for discriminant validity by the weak but significant correlation with the youth physical activity measure. The youth openness to new foods measure also showed convergent validity with the youth cooking skills measure, and again weaker evidence by the weak yet significant correlation with the youth physical activity measure. The youth culinary self-efficacy measure showed evidence for convergent validity by a negative correlation with the cooking skills measure and evidence for discriminant validity by no significant relation with the physical activity measure. The youth eating together measure showed evidence for convergent validity with the youth cooking skills measure and evidence for discriminant validity by no relation with the youth openness to new foods measure. The youth and adult physical activity measures showed evidence for convergent validity by a significant positive relation with each other. Finally, the adult eating together and cooking together measures showed evidence for convergent validity based on the significant, positive relation with each other.

DISCUSSION

The results of this study must be interpreted within the scope of the limitations. Acceptable fit indices for the CFA models were not always achieved, and in several of the CFA models, not all of the factor loadings were significant based on the *z*-tests or had to be fixed to 1.0 owing to a negative variance. This is likely because of the low sample size in the dissemination dataset, but given that the dissemination data were collected after the intervention, it did not seem appropriate to flip analyses to conduct the CFAs on the intervention data based on EFA results from the dissemination data. Although there was some evidence for dimensionality based on the EFA results and the longitudinal invariance results, in future research, the dimensionality of these measures might be explored through larger sample sizes and testing new questions. Whereas analyses for validity were conducted among the measures developed in this study, stronger evidence for validity would come from using previously established measures.[21–24] The 2-item youth goal-setting measure and eating together measure that were reduced to a 2-item measure may be useful in practice and are important to assess in the *iCook 4-H* curriculum but are hard to validate via traditional statistical methods.

The specific aims of this study were to develop youth and adult measures based on the *iCook 4-H* study. Evidence was found for dimensionality, reliability, and validity for most of the measures, although only the youth openness to new foods measure and the adult eating together measures reached acceptable levels across all analyses. The other measures did not show evidence for longitudinal invariance, a measure of how consistent the factor structure was over time. However, this may be explained by analyzing the treatment and control samples together, and it may be that the factor loadings changed over time based on the impact of the *iCook 4-H* intervention for the treatment group.

IMPLICATIONS FOR RESEARCH AND PRACTICE

In the future, although the instrument was designed specifically for the *iCook 4-H* curriculum, the scales could be used or adapted for a variety of programs with similar key constructs. Testing the scales might confirm or provide evidence for modification to improve the dimensionality, reliability, and validity. Using these scales with the *iCook 4-H* curriculum could provide a comparison of outcomes across program administration in a variety of settings.

Table 8. Correlations Among Adult and Youth Measures With 95% Confidence Intervals

Variable	1	2	3	4	5	6	7
1. Cooking skills yourself (youth)	–	–	–	–	–	–	–
2. Cooking skills with help (youth)	0.30*** (0.18 to 0.42)	–	–	–	–	–	–
3. Physical activity (youth)	0.16[a] (0.03 to 0.29)	0.13 (−0.00 to 0.26)	–	–	–	–	–
4. Openness to new foods (youth)	0.23[c] (0.10 to 0.36)	0.34[c] (0.21 to 0.45)	0.14[a] (0.01 to 0.27)	–	–	–	–
5. Culinary self-efficacy (youth)	−0.55[c] (−0.64 to −0.45)	−0.31[c] (−0.42 to −0.18)	−0.12 (−0.25 to 0.02)	−0.28[c] (−0.40 to −0.15)	–	–	–
6. Physical activity (adult)	0.01 (−0.13 to 0.14)	−0.02 (−0.15 to 0.11)	0.20[b] (0.07 to 0.33)	−0.09 (−0.22 to 0.05)	0.06 (−0.08 to 0.19)	–	–
7. Eating together (adult)	−0.01 (−0.14 to 0.13)	0.01 (−0.13 to 0.14)	0.27[c] (0.14 to 0.39)	−0.01 (−0.15 to 0.12)	−0.01 (−0.15 to 0.12)	0.33[c] (0.21 to 0.45)	–
8. Cooking together (adult)	0.11 (−0.03 to 0.24)	−0.00 (−0.14 to 0.13)	0.07 (−0.06 to 0.20)	0.05 (−0.08 to 0.18)	0.01 (−0.13 to 0.14)	0.34[c] (0.22 to 0.46)	0.34[c] (0.22 to 0.46)

[a]$P < .05$; [b]$P < .01$; [c]$P < .001$.

Note: Values in parentheses indicate 95% confidence interval for each correlation.

ACKNOWLEDGMENTS

This material is based on work that is supported by the National Institute of Food and Agriculture, US Department of Agriculture, under Award No. 2012-68001-19605. Other funding is from US Department of Agriculture Experiment Stations for the first 5 authors. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results. Data were taken from the dissertation of the first author. The work was done while he was a graduate student at the identified university affiliation and was part of the multistate study conducted by all authors.

REFERENCES

1. White AA, Colby SE, Franzen-Castle L, et al. The *iCook 4-H* study: an intervention and dissemination test of a youth/adult out-of-school program. *J Nutr Educ Behav.* 2019;51:S2-S20.
2. Franzen-Castle L, Colby SE, Kattelmann KK, et al. Development of the *iCook 4-H* curriculum for youth and adults: cooking, eating, and playing together for childhood obesity prevention. *J Nutr Educ Behav.* 2019;51:S60-S68.
3. Bandura A. *Social Foundations of Thought and Action: A Social Cognitive Theory.* Englewood Cliffs, NJ: Prentice Hall; 1986.
4. National Institute of Food and Agriculture, US Department of Agriculture. Experiential Learning Model. 2016. https://nifa.usda.gov/sites/default/files/resource/Experiential-Learning-Model.pdf. Accessed November 15, 2018.
5. Taylor-Powell E, Steele S, Douglah M. *Planning a Program Evaluation.* Madison, WI: University of Wisconsin Cooperative Extension; 1996.
6. Dunn T, Baguley T, Brunsden V. From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. *Br J Psychol.* 2013;105:399-412.
7. Revelle W, Zinbarg R. Coefficients alpha, beta, omega, and the glb: comments on Sijtsma. *Psychometrika.* 2009;75:145.

8. Nunnally J. *Psychometric Theory*. New York: McGraw-Hill; 1978.
9. Harlow L. *The Essence of Multivariate Thinking*. New York, NY: Routledge; 2014.
10. Baranowski T, Watson K, Bachman C, et al. Self-efficacy for fruit, vegetable and water intakes: expanded and abbreviated scales from item response modeling analyses. *Int J Behav Nutr Phys Activity*. 2010;7:25.
11. Barton K, Wrieden W, Anderson A. Validity and reliability of a short questionnaire for assessing the impact of cooking skills interventions. *J Hum Nutr Diet*. 2011;24:588-595.
12. Velicer W. Determining the number of components from the matrix of partial correlations. *Psychometrika*. 1976;41:321-327.
13. Zwick W, Velicer W. Comparison of five rules for determining the number of components to retain. *Psychol Bull*. 1986;99:432-442.
14. Merenda P. A guide to the proper use of factor analysis in the conduct and reporting of research: pitfalls to avoid. *Meas Eval Couns Dev*. 1997;3:156-164.
15. Hair J, Black W, Babin B, Anderson R. *Multivariate Data Analysis*. 7th ed. New York, NY: Pearson; 2009.
16. Kline R. *Principles and Practice of Structural Equation Modeling*. 4th ed. New York, NY: Guilford Press; 2015.
17. Bentler P. Comparative fit indexes in structural models. *Psychol Bull*. 1990;107:238-246.
18. Tabachnick B, Fidell L. *Using Multivariate Statistics*. New York, NY: Pearson; 2013.
19. Wheaton B, Muthen B, Alwin D, Summers G. Assessing reliability and stability in panel models. *Sociol Method*. 1977;8:84-136.
20. Gold M, Bentler P. Treatments of missing data: a Monte Carlo comparison of RBHDI, iterative stochastic regression imputation, and expectation-maximization. *Struct Eq Model*. 2000;7:319-355.
21. Bradford T, Serrano EL, Cox RH, Lambur M. Development and testing of a nutrition, food safety, and physical activity checklist for EFNEP and FSNE adult programs. *J Nutr Educ Behav*. 2010;42:123−130.
22. Larson N, Perry C, Story M, Neumark-Sztainer D. Food preparation by young adults is associated with better diet quality. *J Am Diet Assoc*. 2006;12:2001-2007.
23. Saunders R, Pate R, Felton G, et al. Development of questionnaires to measure psychosocial influence on children's physical activity. *Prev Med*. 1997;26:241-247.
24. Pinard C, Uvena L, Quam J, Smith T, Yaroch A. Development and testing of a revised Cooking Matters for Adults survey. *Am J Health Behav*. 2015;8:866-873.