

Northeastern Illinois University

**NEIU Digital Commons**

---

Health Sciences and Physical Education Faculty  
Publications

Health Sciences and Physical Education

---

5-23-2024

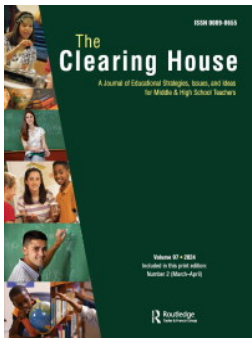
## **The Social Justice Educational Assessment Scale: Filling a Gap in Social Justice Education Assessment and Evaluation**

Jennifer R. Banas

Sarah Gershon

Follow this and additional works at: <https://neiudc.neiu.edu/hpera-pub>

---



# The Clearing House: A Journal of Educational Strategies, Issues and Ideas

ISSN: (Print) (Online) Journal homepage: [www.tandfonline.com/journals/vtch20](http://www.tandfonline.com/journals/vtch20)

## The Social Justice Educational Assessment Scale: Filling a Gap in Social Justice Education Assessment and Evaluation

Jennifer R. Banas & Sarah Gershon

**To cite this article:** Jennifer R. Banas & Sarah Gershon (23 May 2024): The Social Justice Educational Assessment Scale: Filling a Gap in Social Justice Education Assessment and Evaluation, The Clearing House: A Journal of Educational Strategies, Issues and Ideas, DOI: [10.1080/00098655.2024.2338051](https://doi.org/10.1080/00098655.2024.2338051)

**To link to this article:** <https://doi.org/10.1080/00098655.2024.2338051>



Published online: 23 May 2024.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

RESEARCH ARTICLE



## The Social Justice Educational Assessment Scale: Filling a Gap in Social Justice Education Assessment and Evaluation

Jennifer R. Banas<sup>a</sup> and Sarah Gershon<sup>b</sup>

<sup>a</sup>Northeastern Illinois University, Chicago, IL, USA; <sup>b</sup>Deerfield High School, Deerfield, IL, USA

### ABSTRACT

Nationally recognized social justice standards guide educators in developing social justice education. Absent from the guidance are tools to conduct initial formative assessment or to measure the impact of related instruction. To fill that gap, an academic researcher and 10th-grade teacher used a 3-phased, 9-step process to develop, pilot test, and evaluate a self-assessment tool called the Social Justice Educational Assessment Scale (SJEAS). In Phase 1, the team created the SJEAS items, aligning each with the Learning for Justice 9th–12th-grade social justice learning outcomes to ensure content validity. In Phase 2, they pretested the questions, revised the scale, and administered it to 322 student participants. Next, they conducted inter-item and total-item correlation tests and principal component analysis to ensure internal consistency. In Phase 3, they evaluated the SJEAS. Preliminary results indicate the SJEAS could provide educators and curriculum coordinators with a practical formative assessment and impact evaluation tool for social justice education aligned with nationally recognized learning outcomes. Complementary data, including classroom assignments and student discussions, would further enhance its value.

### KEYWORDS

Social justice education, formative assessment, measurement, evaluation, secondary schools

Knowing how to survive in a pluralistic, diverse, multicultural, and often inequitable world is essential for young people to thrive (ABC Task Force, 1989). This skill requires gaining insight into one's identity and others' perspectives, understanding the role of lived experiences and basic human needs, valuing societal differences, possessing emotional conviction and critical thinking skills, and managing difficult situations even when they seem unfair (Derman-Sparks et al. 2019). These cognitive, social, and emotional competencies and skills are necessary for young people to navigate the social-political contexts in which they live, work, and play (Adams et al. 2022). Social justice education can help.

Social justice education raises students' consciousness about inequity in everyday social, environmental, economic, and political situations. It provides the lens to recognize and the skills to interrupt societal inequitable patterns and practices (Hammond 2021). It also includes helping students explore how their experiences and intersectionalities fit into the larger national and global story (Torres 2015). With the support of

their schools, teachers, and evidence-based educational strategies, young people can develop the insight, skills, and strength to live in a complex society alongside others who may or may not share their identity groups (Derman-Sparks et al. 2019; Hammond 2021).

Social justice education initiatives range from in-class modules or lessons to school-wide curricula (Adams et al. 2022). They also might take the form of extracurriculars, such as book groups and advocacy clubs, or as a component of service work (Adams et al. 2022). Whatever the format, initial formative assessment or impact evaluation tools frequently remain absent or lacking in some contexts. Without this feedback, educators lack the information to respond to students' needs and adapt and evaluate social justice education curricula accordingly. Furthermore, how can they demonstrate the difference the curricula make and justify to administrators and community members the time spent implementing social justice education?

This article focuses on our creation and preliminary evaluation of an assessment tool called

the Social Justice Educational Assessment Scale (SJEAS). The tool is a self-assessment designed to capture students' perceived social justice competency concerning nationally recognized social justice learning outcomes. While self-assessments are subjective, they provide valuable insight into a learner audience's mindset. Educators can use this information to tailor instruction or examine changes from pre- to post-instruction as action research. We, an academic researcher and a 10th-grade teacher, describe the context for developing the tool and the steps to evaluate its validity and reliability. We conclude with a discussion regarding opportunities for its use and continued refinement.

## Methods

In the academic years 2020–2022, we created a social justice education module for health education classes at a suburban high school outside of Chicago, IL. We aligned the module with the Learning for Justice (LFJ) 9th–12th grade social justice learning outcomes. These outcomes are divided into four domains: identity, diversity, justice, and action (Learning for Justice 2022). Contextually, *identity* is the collective behavioral and personal characteristics by which someone is known (Yang 2019). *Diversity* refers to including different types of people (e.g. races or cultures) in groups or organizations (Lli 2019). *Justice* is the quality of being and acting just, impartial, or fair (Davis 2020). Finally, *action* consists of the steps to honor and celebrate identity and diversity and bring about justice (Liv 2021).

Absent from the guidance and literature citing the LFJ learning outcomes are ways to conduct initial formative assessments or to measure the impact of related instruction. While social justice education measurement tools exist, most are not designed for classroom settings, and none align with the LFJ learning outcomes (Sebastianelli et al. 2020; Chang and Cochran-Smith 2022; Hicks et al. 2023). Communication with LFJ confirmed this (Author, 2020). For this reason, we developed and evaluated our own tool, the SJEAS.

To create the SJES, we adopted Boateng et al.'s (Boateng et al. 2018) 3-phased, 9-step scale development process. Describing that process is a

spiral of methods and findings. To render the process more clearly, we present each step's methodology and our associated findings together in Results section.

## Results

### Phase 1: Item development

The item development phase focuses on developing the questions that will make up the assessment tool. This phase includes 1) identifying the domain(s) and item generation and 2) assuring content validity.

**Step 1. Identification of the domain and item generation.** Step 1 includes specifying the boundaries of the domain and facilitating item generation. As indicated earlier, we aligned the SJEAS with the LFJ 9th–12th grade learning outcomes. LFJ developed these outcomes to help educators develop age-appropriate content that engages students in social justice issues and establishes related cognitive, social, and emotional competency and skill benchmarks. We used the outcomes as boundaries for scale development because our curriculum aligned with these outcomes. Because the learning outcomes are phrased as I-statements, they are well-suited for self-assessment tools. However, many outcomes were double-barreled or needed to be written in student-friendly language. For example, we identified the following outcome as double-barreled and needing rewording:

I know my family history and cultural background and can describe how my own identity is informed and shaped by my membership in multiple identity groups.

We broke apart the outcome into these two items:

1. I know my family history and cultural background
2. I can describe how my identity is influenced by belonging to different social groups.

Ultimately, we generated 26 items from the 20 learning outcomes.

The next decision we made was the response choices. We modeled ours after the Social Skills Improvement System Social-Emotional Learning survey, validated with 3rd–12th graders (Gresham and Elliot 2008; Gresham et al. 2018). Those choices are: *a lot true for me* (4), *generally true for me* (3), *a little true for me* (2), and *not true for me* (1).

**Step 2: Content validity.** Content validity is the extent to which a tool measures the domain of interest (Vissoci et al. 2022). For Step 2, Boateng suggests enlisting experts to evaluate the content validity of the items generated and conducting cognitive interviews with the target population to ensure that the items adequately measure the domain of interest. We did not do this as our focus was strictly on the LFJ outcomes. Furthermore, the LFJ outcomes were generated by a panel of experts on social justice education (Learning for Justice 2019), and each of our items was forged from those outcomes. Content validity also specifies relevance and representation by asking, “Do the items reflect the audience’s experience?” Our scale items were relevant because they invited students to self-assess their cognitive, social, and emotional competency and skills in relation to the LFJ learning outcomes with which the instruction was aligned.

### Phase 2: Scale development

The scale development phase includes four steps. These steps involve pretesting, administering the scale to enough of the “right” people, reducing the number of scale items as needed, and extracting factors.

**Step 3: Pretesting questions.** Pretesting ensures that scale items and response choices are meaningful. Typically, pretesting entails administering the scale with a small group to eliminate poorly worded items (Boateng et al. 2018). We distributed the SJEAS to a few classes the year before official data collection. The teacher read each question aloud, and students could ask clarifying questions and offer feedback as part of a whole-class discussion. Most of the feedback centered on word reduction and active voice to improve clarity. For example, one of our original items was: *I can describe how my identity is*

*influenced by belonging to different social groups.* We changed this to a simpler version: *I can explain how my social groups shape my identity.* The items presented in this article reflect the modifications based on feedback.

**Step 4: Survey administration and establishing the sample size.** This step includes collecting data from an adequate sample representing the intended population. Researchers should compile this data over multiple periods to ensure reliability. We met both requirements because we administered the SJEAS to our population over five semesters.

Regarding sample size, the literature presents several recommendations. We adopted the ratio approach, which suggests 10 participants for each item (Boateng et al. 2018). That means our 27-item scale would require 270 participants. From 2020–2022, we collected data from over 500 students, thus meeting the requirement. However, because they were under 18, IRB approval required parental consent and student assent. Our data includes students meeting that requirement ( $n=332$ ). We refer to *these* students as participants.

Boateng et al. note that when data are collected over a period of time, as was ours between 2020–2022, and from the same potentially idiosyncratic population, there is the risk of common error variance. Ideally, we would have collected data from multiple settings, for example, other schools. Doing so would be a recommendation for future research.

**Step 5: Item reduction.** Step 5 ensures a scale is straightforward, functional, and internally consistent. Techniques include item difficulty and item discrimination indices, distractor efficiency analysis, and inter-item and item-total correlations. The first three techniques imply “correct” responses. These were irrelevant because we participants to self-assess achievement of the LFJ 9th–12th grade learning outcomes for which there are no wrong and correct answers. However, we did want to know the inter-item and item-total correlations, which support deleting or modifying items.

Item-total correlations look at the relationship between items and the *total* of all items. We performed this analysis using Statistical Package for

Social Science (SPSS) Version 29 to calculate Cronbach's  $\alpha$  reliability coefficient. Ideally, the value should be above .70 (Hair et al. 2006). Ours was .934, suggesting a high internal consistency.

*Inter-item correlations* examine how items relate to each other. This relationship should be tight to ensure all items measure the same concept. For inter-item analysis, deletion is recommended for Cronbach's  $\alpha < .30$  due to low contribution and  $\geq .80$  due to duplicates (Chard et al. 2020). All  $\alpha$  scores were above .300 and below .800, suggesting cohesiveness. See the corrected item-total correlation column in Table 1. Although above .30, we eliminated one item, whose  $\alpha$  was .377, to improve the overall item-total correlation. That item was: *I know my family & cultural background*. Interestingly, we created this item in Step 1 when unpacking the double-barreled learning outcomes. The new Cronbach  $\alpha$  for the scale without this item was .935.

**Step 6: Extraction of factors.** In Step 6, researchers ensure internal consistency and extract non-salient (i.e. unnecessary) items by exploring

patterns (i.e. commonalities) among variables. Confirmatory factor analysis (CFA) and principal component analysis (PCA) are ways to do this (Kim 2008; Pallant 2016). PCA is used for data reduction, and CFA seeks explanations for the correlations between variables. However, PCA can be helpful as an initial step in CFA because it provides information regarding the maximum number and nature of factors (Tabachnick and Fidell 2001). For this latter reason, we chose PCA because we were more focused on extracting non-salient items.

Before PCA, researchers should determine the sampling adequacy for testing the scale's construct validity. Ideally, sample sizes should be over 300. Ours was  $n=332$ . We also examined the Kaiser–Meyer–Olkin (KMO) criterion and Bartlett's test of sphericity. Our KMO criterion (.930) exceeded the recommended value of .6, and Bartlett's test reached statistical significance,  $\chi^2=6157.05$ ,  $p < .001$ .

We set the loading standard at  $\geq .45$ . To determine the number of components to extract, we examined the scree plot and calculated

**Table 1.** Step 5: Item-total statistics.

	Mean if item deleted	Variance if item deleted	Corrected item-total correlation	Cronbach $\alpha$ if item deleted
Identity				
I1 – I know my family and cultural background.	89.22	108.72	0.34	0.94
I2 – I can explain how social groups shape my identity.	89.2	106.35	0.54	0.93
I3 – I know my social group identities define me.	89.27	106.41	0.45	0.93
I4 – I know social group identities define others.	89.19	107.41	0.43	0.93
I5 – I am aware of how my identity is interpreted by others.	89.06	107.69	0.46	0.93
I6 – I am aware of how my identity impacts others.	89.11	106.41	0.51	0.93
I7 – I am proud and confident about my identity.	89.07	107.45	0.48	0.93
I8 – I have a positive view of myself.	89.3	108.11	0.41	0.93
Diversity				
D1 – I don't think of or treat others as inferior.	88.81	108.34	0.48	0.93
D2 – I interact comfortably with others who are similar/different.	88.76	107.89	0.56	0.93
D3 – I respectfully describe how others are similar/different.	88.83	106.29	0.68	0.93
D4 – I respectfully express curiosity about others' experiences.	88.86	106.1	0.68	0.93
D5 – I exchange ideas and beliefs in an open-minded way.	88.95	105.71	0.66	0.93
D6 – I connect by showing empathy, respect, & understanding.	88.82	106.35	0.69	0.93
D7 – I understand diversity includes unequal power relations.	88.9	106.38	0.59	0.93
Justice				
J1 – I relate to all people as individuals, rather than their groups.	89.01	106.54	0.59	0.93
J2 – I can identify stereotypes when I see or hear them.	88.98	108.13	0.5	0.93
J3 – I can recognize and describe unfairness and injustice.	88.91	106.05	0.67	0.93
J4 – I can explain the impact of biased words and behaviors.	88.92	105.66	0.67	0.93
J5 – I can explain the impact of unfair practices and laws.	88.98	104.63	0.68	0.93
J6 – I am aware of my adv/disadvantages based on social groups.	88.89	105.6	0.66	0.93
Action				
A1 – I empathize when people are excluded or mistreated.	88.86	106	0.7	0.93
A2 – I stand up to exclusion, prejudice, and injustice.	89.14	103.93	0.69	0.93
A3 – I have the courage to speak against biased/hurtful words/actions.	89.11	105.38	0.57	0.93
A4 – I communicate respectfully, even when I disagree.	89.1	106.44	0.55	0.93
A5 – I stand up when I see someone who is excluded.	89.21	104.77	0.63	0.93
A6 – I stand up when I see someone shown prejudice or discrimination.	89.13	104.36	0.68	0.93



eigenvalues. A total of five components with eigenvalues  $\geq 1$  were extracted. Also, the scree plot revealed a break after the fifth component. Three items (A4, D1, and J2) had lower communality values (.427, .484, and .487) than the others, all above .55. We eliminated these items and reran the analysis with the remaining 23 items. Again, five components emerged. Considering the conceptual meaning of the scale, we identified these components as follows:

1. Social justice skills: Entry-level
2. Social justice skills: Advanced-level
3. Identity: Cognitive
4. Identify: Social
5. Identity: Emotional

We based the naming of these components on Derman-Sparks' anti-bias framework and its intended use (Derman-Sparks et al. 2019). The LFJ outcomes are based on this framework, as stated in the Introduction. One of the framework's goals is to support educators in developing

a curriculum that helps children explore and apply new understandings and behaviors. Cognitive, social, and emotional competency and related skills are a significant focus of those understandings and behaviors. We named our components as such so that educators and curriculum coordinators could use the data to tailor and scaffold instruction.

The explanatory power of each component was 25.9%, 14.0%, 10.4%, 8.6%, and 7.7%, respectively, with a total cumulative explanatory power of 66.7%. See Table 2. It should be noted that these findings suggest that LFJ's Identity learning outcomes might consist of three unique components. We elaborate on this and the implications in the Discussion section.

### Phase 3: Scale evaluation

**Step 7: Tests of dimensionality.** To determine whether a measurement will work with other samples and longitudinally, researchers perform tests of dimensionality. These tests might include

**Table 2.** Step 6: Factor analysis of the SJEAS.

Components and items	M $\pm$ SD	Component				
		1	2	3	4	5
Social justice skills: Entry-level						
D6 - I connect by showing empathy, respect, and understanding.	3.62 $\pm$ .55	.717				
D2 - I interact comfortably with others who are similar/different.	3.69 $\pm$ .54	.712				
D5 - I exchange ideas and beliefs in an open-minded way.	3.50 $\pm$ .61	.711				
D3 - I respectfully describe how others are similar/different.	3.61 $\pm$ .56	.699				
A1 - I empathize when people are excluded or mistreated.	3.58 $\pm$ .57	.689				
D4 - I respectfully express curiosity about others' experiences.	3.58 $\pm$ .57	.682				
J4 - I can explain the impact of biased words and behaviors.	3.52 $\pm$ .61	.665				
D7 - I understand diversity includes unequal power relations.	3.54 $\pm$ .64	.657				
J3 - I can recognize and describe unfairness and injustice.	3.54 $\pm$ .58	.650				
J1 - I relate to all people as individuals, rather than their groups.	3.44 $\pm$ .61	.615				
J6 - I am aware of the advantages/disadvantages I have because of my social groups.	3.55 $\pm$ .63	.582				
J5 - I can explain the impact of unfair practices and laws.	3.46 $\pm$ .68	.574				
Social justice skills: Advanced-level						
A6 - I stand up when I see someone shown prejudice or discrimination.	3.32 $\pm$ .70		.801			
A5 - I stand up when I see someone who is excluded.	3.24 $\pm$ .71		.793			
A3 - I have the courage to speak against biased/hurtful words/actions.	3.33 $\pm$ .73		.780			
A2 - I stand up to exclusion, prejudice, and injustice.	3.31 $\pm$ .72		.452			
Identity: Cognitive						
I4 - I know social group identities define others.	3.25 $\pm$ .73			.866		
I3 - I know my social group identities define me.	3.17 $\pm$ .79			.859		
I2 - I can explain the influence of social groups on my identity.	3.25 $\pm$ .69			.680		
Identity: Social						
I5 - I am aware of how my identity is interpreted by others.	3.38 $\pm$ .66				.814	
I8 - I am aware of how my identity might impact others.	3.38 $\pm$ .72				.817	
Identity: Emotional						
I6 - I have a positive view of myself.	3.34 $\pm$ .72					.801
I7 - I am proud of and confident about my identity.	3.14 $\pm$ .70					.786
Eigenvalue		9.53	1.92	1.49	1.29	1.14
Explained variance		25.91	14.04	10.4	8.66	7.66
Cumulative variance		25.91	39.94	50.38	59.03	66.69
KMO = .30, Bartlett's test = $\chi^2 = 6157.05$ , $p < .001$						
Total Cronbach's $\alpha = .925$	Cronbach's $\alpha$	.916	.884	.814	.812	.746

confirmatory factor analysis, bifactor modeling, and measurement invariance. While we did administer the SJEAS over multiple semesters with different students, we could not implement it longitudinally or in another setting, as required by these techniques. Therefore, we did not perform these tests. The information from these tests is critical at the school and district levels before generalizing results or using the data to make programmatic or policy decisions. At the classroom level, the data from these tests will help educators and curriculum coordinators make informed decisions about whether to use whole-scale or subscale scores when tailoring instruction and measuring impact.

**Step 8: Tests of reliability.** Reliability tests evaluate the degree of consistency an instrument exhibits under identical conditions (Boateng et al. 2018). We chose to perform the split-half test and test-retest using Pearson product-moment correlation. The split-half test assumes that two halves of a scale should yield similar reliability scores, and both halves still measure the larger construct independently (Pallant 2016). One of the statistics associated with this test is the Spearman-Brown coefficient. Reliability is associated with scores greater than .80. Using SPSS to perform this test, our score was .861, suggesting high internal consistency.

The test-retest examines the degree to which participants' performance is repeatable (Boateng et al., 2018). In the final semester of collecting data to evaluate the SJEAS, we administered it twice in the second author's classrooms as a pre-/post-test to measure the impact of social justice education instruction. We administered these tests six weeks apart, before and after a multi-week curriculum unit aligned with the LFJ learning outcomes (AUTHORS, 2024). Pearson product-moment correlation indicated a robust and positive relationship between the two tests,  $r = .722$ ,  $n = 77$ ,  $p = .001$ . While this sample size was small and, therefore, inadequate to make generalizations, the findings suggest high internal consistency. As for the impact results, we share those in the Discussion section instead of here, as this article focuses on developing and validating the scale.

**Step 9: Tests of validity.** Validity is the extent to which an instrument measures the construct(s) it was designed to measure. Although we focus on it in this final step, validation began with Step 1 and continued throughout. The most common validity tests are content, criterion, and construct (Boateng et al. 2018).

Researchers can perform *content validity testing* before administering a scale. We did this test in Step 2. We indicated that our tool was designed to reflect the LFJ learning outcomes and to be used with instruction aligned with those outcomes. It was also tested in classrooms where that content was being taught. Finally, teachers and students provided feedback on the tool at the beginning and throughout the various iterations to ensure it reflected what was instructed.

*Criterion validity* comes in two forms: predictive and concurrent. For scales like ours, *predictive validity* is the extent to which the scale predicts future behavior. We did not perform this test because we did not plan to follow the students into the future. *Concurrent validity* is the extent to which scores on the new scale compare to a "gold standard" version. We did not perform this test because there is no comparative. As stated in the introduction, there are other social justice assessment tools; however, they do not align with the LFJ outcomes. Our SJEAS attempts to fill that void.

*Construct validity* is the extent to which the scale reflects a hypothetical, theoretical construct and real-world criteria (Pallant 2016; Boateng et al. 2018). The SJEAS meets this form of validity at face value because of its rootedness in the LFJ learning outcomes. Those outcomes are based on the Derman-Sparks anti-bias framework, which promotes curriculum and competencies related to real-world behaviors related to culture, race, and other social issues (Derman-Sparks et al. 2019). However, to further validate the SJEAS, Boateng et al. point to four statistical tests of construct validity: convergent, discriminant, differentiation by "known groups," and correlation analysis. We did not perform these tests during our pilot testing in large part due to the absence of comparative scales. Future testing could include comparing the SJEAS and results to scales that measure similar or related concepts,



such as self-awareness, social awareness, and relationship skills, or comparing results to qualitatively rich assessments, such as student reflections and role-playing.

## Discussion

The LFJ learning outcomes provide educators with benchmarks to develop social justice educational strategies. However, there are no standardized tools to conduct an initial formative assessment or measure the impact of instruction. Without that feedback, educators and curriculum coordinators lack the information to appropriately tailor instruction to their student audience or justify the effects of social justice education to stakeholders and decision-makers. While some assessments capture social justice, diversity, and equity aspects, none align with the LFJ outcomes (Chang and Cochran-Smith 2022). To address that void, we created the SJEAS.

Self-assessments are a common way to capture “where” an audience is at a moment in time. A recent meta-analysis revealed that 43 of 45 social justice, diversity, and equity-related assessment tools were self-assessments. Participants responded to prompts about their understanding and mental states on these self-assessments, much like our SJEAS. For example, the Teacher Multicultural Attitude Survey presents statements about multicultural awareness of, comfort with, and sensitivity toward issues related to teaching diverse students in K–12 settings, much like our SJEAS, which prompts students to consider concepts related to social justice. On the latter survey, participants respond to these statements on a Likert scale of agreement, again much like SJEAS (Ponterotito et al. 1998). Another self-assessment, the Culturally Responsive Teaching Self-Efficacy Scale, measures teachers’ or teacher candidates’ beliefs about culturally responsive teaching (Siwatu 2007). For this assessment, participants self-rate their competency to perform specific tasks, much like some of our SJEAS statements prompt students to self-rate their competency to perform acts of social justice.

While self-assessments do not measure behavioral performance in the real world, they provide helpful information. They supply a glimpse into

how a person might act in a given context. This idea is supported by contemporary health behavior theories that demonstrate that how people think, including knowledge, attitudes, and beliefs, is a crucial influence on behavior (Sebastianelli et al. 2020; Vaffis et al. 2021; Jung and Yang, 2022). These theories include, but are not limited to, the health belief model, the theory of planned behavior, and the transtheoretical model (Becker 1974; Ajzen 1985; Prochaska and Diclemente 1986; Ajzen and Schmidt 2020). Scales associated with these theories inform the design of educational interventions and measure their impact, much like we intend our scale to be used by others.

## Implications for educators and curriculum coordinators

Our SJEAS captures students’ self-perceived cognitive and social-emotional skills in relation to the LFJ social justice learning outcomes. When administering the SJEAS, users must decide whether to use the whole-scale or individual components’ results. Our Step 6 results suggest the LFJ learning outcomes, while divided into four domains, might be made up of three, or as many as five, components. Two components could be regarded as levels of social justice skills, and another consists of interrelated concepts of identity. One *could* examine these components separately; however, the reliability of the whole scale ( $\alpha = .925$ ) is more significant than its parts. Further, clarity on the relationships between the subscales is still needed because we could not perform Step 7 – Tests of Dimensionality. Users who do choose to look at the components separately might want to supplement the SJEAS with open-ended prompts. Such prompts would also provide complementary data for users examining the whole-scale results. We discuss that in Limitations and Opportunities for Future Research.

The SJEIS can serve as a valuable tool for gathering insight into a student audience before delivering instruction aligned with those outcomes (i.e. formative assessment) and evaluating the impact of that instruction. For example, as an initial formative assessment, we compared

aggregate scores to previous semesters to gauge “where” the current students were in relation to LFJ outcomes. If we noted a lower-than-average aggregate score, we *did* look at the individual SJEAS item aggregate results. While we do not recommend reporting aggregate item scores for non-instructional use, they can provide helpful information. For example, the overall student aggregate score was lower in one of the semesters. We examined the aggregate item scores and discovered students might need focused instruction on the Identity: Social sub-component, which centers on the influence of social groups on one’s identity. Another semester, we noted that students self-assessed lower on the Identity: Emotional sub-component. Items making up that sub-component focus on viewing oneself positively and feeling proud and confident about one’s identity. In both instances, we adjusted instruction, particularly classroom discussions, to talk more deeply about those sub-component areas.

As noted in Step 8, in the final semester of gathering data for this article, we administered the SJEAS as a pre-/post-test to evaluate the impact of the 6-week-long LFJ-aligned instruction. Results suggest that the curriculum may have positively affected the students. A paired samples t-test revealed a statistically significant increase from the pretest ( $M=3.14$ ,  $SD = .44$ ) to the post-test ( $M=3.46$ ,  $SD = .41$ ),  $t(76) = 8.80$ ,  $p < .001$ . As with the initial formative assessment described above, we cannot generalize the results due to an inadequate sample size. However, we can celebrate the possibility that the instruction may have made a difference and that the SJEAS could potentially serve as an impact evaluation tool.

### **Limitations and opportunities for future research**

We could not perform all nine of the steps of the Boateng et al. framework. That said, continued testing and improvement opportunities exist for the SJEAS. For example, regarding reliability, there is value in conducting longitudinal research to study the long-term impact of social justice education or administering the scale to other groups and settings so that tests of dimensionality can be

performed (Step 7). As for validity, one could compare SJEAS results to more qualitatively rich data such as classroom assignments or focus groups. One might also compare the results from the SJEAS to those from other scales. We did not do the latter because no other scale was rooted in the LFJ learning outcomes. While the SJEAS is still being prepared for extensive survey research, it provides a subject-specific assessment tool that educators and curriculum coordinators can use to tailor instruction and, with more testing, could be utilized to demonstrate the impact of social justice education.

### **Conclusion**

*Social justice is both a goal and a process.* Bell (Adams et al., 2022)

Like social justice, developing a scale is a goal and a process. This article presented the preliminary evaluation of a new scale called the SJEAS. In introducing the process, we offer insight into the steps involved with developing and validating scale-based assessment tools. The SJEAS aims to provide educators and curriculum coordinators with an assessment tool to complement social justice education aligned with the LFJ 9th–12th-grade social justice learning outcomes. Such data is essential to ensuring social justice curricula that develop young people’s insight, skills, and strength to live in our complex society and to address social inequities (Derman-Sparks et al. 2019; Hammond 2021). Continued testing and complementary data sources will further enhance the value of the scale.

### **Human subjects approval statement**

The Northeastern Illinois University Institutional Review Board determined this study to be exempt.

### **Disclosure statement**

The authors declare no conflict of interest.

### **Funding**

We have no funding to acknowledge

## References

- Adams, M., L. A. Bell, D. J. Goodman, D. Shlasko, R. R. Briggs, R. Pacheco, eds. 2022. *Teaching for diversity and social justice*. 4th ed. New York, NY: Routledge.
- Ajzen, I. 1985. From intentions to actions: A theory of planned behavior. In *Action-control: From cognition to behavior*, ed. J. Kuhl and J. Beckman, 11–39. Heidelberg, Germany: Springer.
- Ajzen, I., and P. Schmidt. 2020. Changing behavior using the theory of planned behavior. In *The Handbook of behavior change*, eds. K. Hamilton, L. D. Cameron, M. S. Hagger, N. Hankonen, T. Lintunen, 17–31. Cambridge, MA: Cambridge University Press.
- Becker, M. H. 1974. *The health belief model and personal health behavior*. N.J.: Slack Thorofare.
- Bell, L. A. 2022. Theoretical foundations for social justice education. In *Teaching for diversity and social justice*, eds. M. Adams, L. A. Bell, D. J. Goodman, D. Shlasko, R. R. Briggs, R. Pacheco, 1–24, 4th ed. New York, NY: Routledge.
- Boateng, G. O., T. B. Neilands, E. A. Frongillo, H. R. Melgar-Quinonez, and S. L. Young. 2018. Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health* 6:149. doi: 10.3389/fpubh.2018.00149.
- Chang, W.-C., and M. Cochran-Smith. 2022. Learning to teach for equity, social justice, and/or diversity: Do the measures measure up? *Journal of Teacher Education* 0 (0):002248712210752. doi: 10.1177/00224871221075284.
- Chard, C. A., D. S. Nelson, K. A. Walters, et al. 2020. An inclusive approach to exploring perceptions of body image, self-esteem, and physical activity among Black and African-American girls: Smart Fit Girls Melanin Magic. *Journal of Park & Recreation Administration* 38 (3):133–51. doi: 10.18666/JPra-2019-9710.
- Davis, E. 2020. Digging deep into the Social Justice Standards: Justice. *Learning for Justice* (64):15–7.
- Derman-Sparks, L. & ABC Task Force. 1989. *Anti-bias curriculum: Tools for empowering young children*. Washington, DC: NAEYC.
- Derman-Sparks, L., J. Olsen Edwards, and C. M. Goins. 2019. Understanding anti-bias education: Bringing the four core goals to every facet of your curriculum. *Young Children* 74 (5):6–13.
- Hair, J. F. J., W. C. Black, B. J. Babin, R. E. Anderson, and R. L. Tatham. 2006. *Multivariate data analysis*. 6th ed. Englewood Cliffs, NJ: Prentice-Hall.
- Hammond, Z. 2021. Liberatory education: Integrating the science of learning and culturally responsive practice. *American Educator* 45 (2):4–11. 39.
- Hicks, E. T., M. Alvarez, and M. M. Domenech Rodríguez. 2023. Impact of difficult dialogues on social justice attitudes during a multicultural psychology course. *Teaching of Psychology* 50 (2):175–83. doi: 10.1177/00986283221104057.
- Jung, H., and Y. Yang. 2022. Reliability and validity of the Korean version of the social justice scale in nursing students. *International Journal of Environmental Research and Public Health* 19 (21):14443. doi: 10.3390/ijerph192114443.
- Kim, H.-J. 2008. Common factor analysis versus principal component analysis: Choice for symptom cluster research. *Asian Nursing Association* 2 (1):18–24.
- Learning for Justice. 2019. *Let's talk: A guide to facilitating critical conversations with students*. <https://www.learningforjustice.org/sites/default/files/2019-12/TT-Lets-Talk-December-2019.pdf>
- Learning for Justice. 2022. *Social justice standards: The learning for justice anti-bias framework*. 2nd ed. Montgomery, AL: Southern Poverty Law Center.
- Liv, J. 2021. Digging deep into the social justice standards: Action. *Learning for Justice* (66):15–7.
- Lli, Y. 2019. Digging deep into the Social Justice Standards: Diversity. *Learning for Justice* (63):15–7.
- Pallant, J. 2016. *SPSS survival manual*. 6th ed. New York, NY: Open University Press.
- Ponterotito, J. G., S. Baluch, T. Greig, and L. Rivera. 1998. Development and initial score validation of the Teacher Multicultural Attitude Survey. *Educational and Psychological Measurement* 58 (6):1002–16. doi: 10.1177/0013164498058006009.
- Prochaska, J. O., and C. C. Diclemente. 1986. Toward a comprehensive model of change. In *Treating addictive behaviors*, eds. W. R. Miller, N. Heather, 3–27. New York, NY: Plenum Press.
- Sebastianelli, R., O. Isil, and Y. Li. 2020. Infusing social justice into decision sciences core courses. *Decision Sciences Journal of Innovative Education* 18 (4):522–48. doi: 10.1111/dsji.12220.
- Siwatu, K. O. 2007. Preservice teachers' culturally responsive teaching self-efficacy and outcome expectancy beliefs. *Teaching and Teacher Education* 23 (7):1086–101. doi: 10.1016/j.tate.2006.07.011.
- Tabachnick, B. G., and L. S. Fidell. 2001. *Using multivariate statistics*. 4th ed. Boston, MA: Allyn and Bacon.
- Torres, C. 2015. Why teaching about social justice matters. *Learning for Justice Magazine*. <https://www.learningforjustice.org/magazine/why-teaching-about-social-justice-matters>
- Vaffis, S., E. J. Anderson, R. M. Fettkeather, T. Warholak, and E. Hall-Lipsy. 2021. Validation of a pre-/post-survey to measure change in student social justice attitudes and empathy. *Currents in Pharmacy Teaching & Learning* 13 (10):1270–7. doi: 10.1016/j.cptl.2021.07.016.
- Vissoi, J., J. Gallis, and R. D'Anna. 2022. *Measure development*. Institute DGH [https://sites.globalhealth.duke.edu/rdac/wp-content/uploads/sites/27/2022/06/Core-Guide\\_Measure\\_Development\\_05\\_02\\_22.pdf](https://sites.globalhealth.duke.edu/rdac/wp-content/uploads/sites/27/2022/06/Core-Guide_Measure_Development_05_02_22.pdf)
- Yang, J. 2019. Digging deep into the Social Justice Standards: Identity. *Learning for Justice* (61):15–7.

## Appendix – *final scale*

Below appears the final scale. The original Learning for Justice domain is provided as a reference.

Scale item	Original LFJ Domain
1. I connect by showing empathy, respect, and understanding.	D
2. I interact comfortably with others who are similar/different.	D
3. I exchange ideas and beliefs in an open-minded way.	D
4. I respectfully describe how others are similar/different.	D
5. I empathize when people are excluded or mistreated.	A
6. I respectfully express curiosity about others' experiences.	D
7. I can explain the impact of biased words and behaviors.	J
8. I understand diversity includes unequal power relations.	D
9. I can recognize and describe unfairness and injustice.	J
10. I relate to all people as individuals, rather than their groups.	J
11. I am aware of the advantages/disadvantages I have because of my social groups.	J
12. I can explain the impact of unfair practices and laws.	J
13. I stand up when I see someone shown prejudice or discrimination.	A
14. I stand up when I see someone who is excluded.	A
15. I have the courage to speak against biased/hurtful words/actions.	A
16. I stand up to exclusion, prejudice, and injustice.	A
17. I know social group identities define others.	I
18. I know my social group identities define me.	I
19. I can explain the influence of social groups on my identity.	I
20. I am aware of how my identity is interpreted by others.	I
21. I am aware of how my identity might impact others.	I
22. I have a positive view of myself.	I
23. I am proud of and confident about my identity.	I