University of Kentucky

UKnowledge

Educational, School, and Counseling Psychology Faculty Publications

Educational, School, and Counseling Psychology

4-2019

Development and Validation of the Awareness of Privilege and Oppression Scale-2

Michael James McClellan
Eastern Kentucky University

Lori Patricia Montross-Thomas
University of California - San Diego

Pamela Remer
University of Kentucky, premer@email.uky.edu

Yoshie Nakai
Eastern Kentucky University

Addison D. Monroe
Eastern Kentucky University

Follow this and additional works at: https://uknowledge.uky.edu/edp_facpub

🟡 Part of the Educational Psychology Commons, and the Student Counseling and Personnel Services Commons

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Repository Citation

Development and Validation of the Awareness of Privilege and Oppression Scale-2

Notes/Citation Information

Digital Object Identifier (DOI)

Original Research

Development and Validation of the Awareness of Privilege and Oppression Scale–2

Michael James McClellan[1] iD, Lori Patricia Montross-Thomas[2], Pamela Remer[3], Yoshie Nakai[1], and Addison D. Monroe[1]

Abstract

The two studies presented describe the revision process that led to the development of the Awareness of Privilege and Oppression Scale–2 (APOS-2) and efforts to evaluate the new measure's reliability and construct validity. In Study 1, a 26-item measure was developed from data gathered from a sample of 484 undergraduate students. An exploratory factor analysis suggested a four-factor solution made up of awareness of racism, sexism, heterosexism, and classism was appropriate. In Study 2, confirmatory factor analysis suggested the proposed hierarchical four-factor solution was the best available fit of the data using a second sample of 520 undergraduate students. The observed Cronbach alpha reliability estimates for the final 26-item total score and subscale scores in the two presented studies were as follows: Total score (.89, .88), Awareness of Heterosexism (.82, .82), Awareness of Sexism (.76, .76), Awareness of Classism (.81, .82), and Awareness of Racism (.84, .80).

Multicultural education course offerings are expanding at many colleges and universities throughout the United States (Howard-Hamilton, Cuyjet, & Cooper, 2011; Mallinckrodt et al., 2014). These courses frequently cover a range of topics (e.g., race, gender, sexual orientation, socioeconomic status [SES]; Flammer, 2001; Montross, 2003) to better prepare students for an increasingly diverse work environment (Bezrukova, Jehn, & Spell, 2012). One of the goals of multicultural education is to raise awareness of privilege and oppression (D. J. Goodman, 2001; Hays, 2005; Montross, 2003). Awareness of privilege and oppression refers to an individual's ability to recognize the social injustices that result from systemic privilege and oppression, and this construct has been identified as a foundational step in social identity development (McClellan, 2014; Montross, 2003; see Cass, 1979; Helms, 1990; Sue & Sue, 1990, 1999; Worell & Remer, 2003). The Awareness of Privilege and Oppression Scale (APOS; Montross, 2003) was developed to assess an individual's awareness of privilege and oppression related to race, gender, sexual orientation, and SES, and this measure has been utilized to evaluate the effectiveness of multicultural education (Remer, 2008). While the APOS (Montross, 2003) captures a critical aspect of social identity development, especially within the multicultural education context, previous research findings suggest that the psychometric

properties of the scale can be improved (McClellan, 2014; Remer, 2008). The purpose of the present study is to address the psychometric issues of the original APOS (Montross, 2003) by completing a revision of the measure. According to social identity development theory, awareness of privilege and oppression must occur as an individual moves from initial to more advanced levels of social identity (Worell & Remer, 2003).

Identity Development and Multicultural Education

Social identity development models have been created to explain the developmental trajectory of many social identities including race (Helms, 1990), gender (Downing & Roush, 1985), sexual orientation (Cass, 1979), general social identity development (Worell & Remer, 2003), as well as

[1]Eastern Kentucky University, Richmond, USA
[2]University of California, San Diego, La Jolla, USA
[3]University of Kentucky, Lexington, USA

Corresponding Author:
Michael James McClellan, Eastern Kentucky University, 521 Lancaster Avenue, 127 Cammack Building, Richmond, KY 40475, USA.
Email: m.mcclellan@eku.edu

many other social identities. Worell and Remer (2003) identified four social identity developmental levels that both privileged and oppressed group members may traverse within a given social identity (e.g., race, gender, sexual orientation, social class). Those four levels are (a) *Pre-Awareness*, (b) *Encounter*, (c) *Immersion*, and (d) *Integration and Activism*.

In Pre-Awareness or Level 1, both privileged (e.g., White individuals, men) and oppressed (e.g., People of Color, women) group members support values and beliefs that are often created by privileged group members with no recognition of the disparate nature of the impact of privilege and oppression (Worell & Remer, 2003). In Encounter or Level 2, the worldview of privileged and oppressed group members starts to diverge as they begin to understand the nature and impact of social advantages (e.g., can access the benefits of White privilege) or disadvantages (e.g., experience acts of racism) that are associated with privilege or oppression. This newfound understanding of privilege and oppression leads to negative emotions for both privileged (e.g., shame, guilt) and oppressed (e.g., anger) group members (Worell & Remer, 2003). In Level 3 or Immersion, privileged group members seek to initiate contact with, become more open to engaging, and begin to empathize with the experience of oppressed group members (Worell & Remer, 2003). On the contrary, oppressed group members prefer contact with other oppressed group members and seek to learn more about their own unique cultural heritage. In the fourth level, Integration and Activism, both privileged and oppressed group members become more comfortable being around members of the other group, better understand their own social identities, and are more prepared to recognize social inequities around them. Also, members of both groups are willing to equitably distribute valued societal resources and actively participate in social advocacy work.

Outcome Assessment Approaches in Multicultural Education

The multicultural education literature offers various methodological approaches to outcome measurement including assessment of identity development, cultural competency, and critical consciousness. The identity development approach seeks to determine which stage or level an individual is within a given identity development framework (O'Meara, 2001). In this approach, successive measurement attempts before and after diversity training would determine whether an individual has progressed to higher stages or levels of development (O'Meara, 2001).

Other identity development approaches focus on specific components of identity development theory (i.e., awareness of privilege and oppression, openness to diversity, intercultural empathy or engagement). For example, the APOS (Montross, 2003) measures an individual's awareness of privilege and oppression which captures Worell and Remer's

(2003) encounter and immersion levels (Levels 2 and 3) of social identity development. The Everyday Multicultural Competencies/Revised Scale of Ethnocultural Empathy (EMC/RSEE; Mallinckrodt et al., 2014) looks at awareness of racism, openness, and empathy which captures not only encounter and immersion but also integration and activism (Level 4) of the social identity development model.

Multicultural competency approaches generally seek to measure an individual's level of cultural competency (i.e., awareness of attitudes toward minority group members, cultural knowledge, and intercultural communication skills; Kim, Cartwright, Asay, & D'Andrea, 2003). For example, Kim et al. (2003) developed the Multicultural Awareness, Knowledge, and Skills Survey–Revised (MAKSS-R) to measure the counseling-related development of cultural competency. The MAKSS-R (Kim et al., 2003) does not assess every aspect of identity development (e.g., advocacy) and focuses on self-awareness rather than awareness of privilege and oppression.

The critical consciousness measurement approach seeks to measure an individual's (a) critical reflection, (b) critical motivation, (c) critical action, (d) critical agency, (e) critical behavior, (f) sociopolitical development, and (g) social perspective-taking (Diemer & Hsieh, 2008; Diemer, McWhirter, Ozer, & Rapa, 2015; Diemer, Rapa, Voight, & McWhirter, 2016; Thomas et al., 2014). There is no single measure that comprehensively assesses all of these dimensions. For example, McWhirter and McWhirter's (2016) Measure of Adolescent Critical Consciousness (MACC) specifically measures adolescents' critical agency (perceptions of empowerment) and critical behavior (follow-through), whereas Shin, Ezeofor, Smith, Welch, and Goodrich's (2016) Contemporary Critical Consciousness Measure (CCCM) measures critical reflection of racism, classism, and heterosexism.

Each of these four measurement approaches has strengths and limitations. The identity development approaches offer the opportunity to explore whether an individual moves to a higher level of identity development through participation in a multicultural education course (Worell & Remer, 2003). However, few identity development models agree on the overall number of stages a person must traverse and models also disagree on what milestones need to be accomplished within each stage or level (O'Meara, 2001). Cultural competency approaches often focus on the development of an awareness of personal attitudes, intercultural knowledge, and specific skills needed to interact with diverse individuals (Kim et al., 2003). However, these approaches rarely address advocacy, a milestone of higher levels of identity development (see Worell & Remer, 2003). Critical consciousness approaches appear to cover a number of aspects of identity development (e.g., reflection, motivation, action, etc.; Diemer & Hsieh, 2008; Diemer et al., 2015; Diemer et al., 2016; Thomas et al., 2014); however, no individual measure that assesses this construct appears to cover the full range of

identity development. The identity development components approach allows researchers and educators to focus on growth in a specific phase of social identity development (e.g., awareness of privilege and oppression). This method, however, lacks the comprehensiveness of other approaches unless multiple measures are employed. Researchers and educators often must rely on a combination of approaches or measures to evaluate multicultural education outcomes.

Assessment of Awareness of Privilege and Oppression

Assessment of awareness of privilege and oppression is consistent with the components approach of social identity development measurement model in multicultural education assessment. Awareness of privilege and oppression has been compared with other similar constructs assessed in the multicultural education context. Critical reflection (a sub-construct of critical consciousness) generally refers to the process by which individuals begin to question how hierarchical social structures create oppression (Diemer et al., 2016) and is, perhaps, closest in nature to the construct of awareness of privilege and oppression (e.g., both look at perceptions of oppression). One important way in which critical reflection and awareness of privilege and oppression are different is that critical reflection examines the ways in which individuals begin to question social structures that create oppression (Diemer & Hsieh, 2008) whereas awareness of privilege and oppression (as measured by the APOS; Montross, 2003) measures individuals' recognition of both social privilege or oppression (Montross, 2003). Our suggestion is that these two constructs are distinct because one must become aware of systemic privilege and oppression before that individual will begin to question the hierarchical structures of the system.

The literature describes three scales that specifically measure an individual's awareness of privilege and oppression. First, the APOS (Montross, 2003) is a 50-item, self-report, Likert-type scale that measures an individual's awareness of privilege and oppression in four areas: (a) race, (b) gender, (c) sexual orientation, and (d) SES. Montross (2003) conceptualized that overall awareness of privilege and oppression was made up of sub-constructs of specific types of awareness (e.g., awareness of racism). Montross (2003) demonstrated the proposed dimensionality of the construct through exploratory factor analyses (EFAs) with undergraduate students and psychology professional samples. Awareness of racism, sexism, heterosexism, and classism were all generally distinct factors; however, Montross' (2003) use of orthogonal rotation in her EFA found that some items loaded heavily on unintended factors suggesting there may have been more overlap in the factors than the author had hypothesized. Undergraduate students who were administered the APOS scored significantly lower, $t(383) = 27.51$, $p < .000$, than a sample of psychology professionals. While the overall scale showed reasonable reliability ($\alpha = .83$), reliability estimates of the subscales varied ($\alpha = .46-.75$). Remer (2008) utilized the APOS in a pre–post, control versus treatment group design to evaluate the effectiveness of multicultural education training in a sample of undergraduate students and found that the APOS was able to detect increases in awareness for the diversity course participants as expected. Montross (2003) originally conceptualized a second-order factor structure of awareness of privilege and oppression that is made up of four, first-order factors; however, there are no follow-up studies to confirm the factor structure of the APOS.

The second measure of awareness of privilege and oppression noted in the literature is the Privilege and Oppression Inventory (POI; Hays, 2005; Hays, Chang, & Decker, 2007). The POI is a 39-item, Likert-type, self-report inventory that measures an individual's awareness of privilege and oppression based on four forms of privilege and oppression including (a) *White Privilege Awareness*, (b) *Heterosexism Awareness*, (c) *Christian Privilege Awareness*, and (d) *Sexism Awareness*. Cronbach alpha reliability estimates for the total score (.95) and subscale scores (range from .79 to .92) are based on a sample of 428 graduate counseling-related trainees (Hays, 2005; Hays et al., 2007). Convergent validity has been demonstrated based on moderate correlations between the POI and measures of comfort and acceptance with cultural similarities and differences as well as attitudes toward racial diversity and gender equality (Hays, 2005). EFA and confirmatory factor analysis (CFA) evidence supported the proposed theoretical framework of the construct (Hays et al., 2007). Furthermore, Byrd and Hays (2013) utilized the Sexism Awareness and Heterosexism Awareness subscales as a combined outcome measure in a randomized, control versus treatment group, pre- versus post-test design to evaluate an LGBTQ training for professional school counselors and graduate student trainees. Students who completed the training scored higher on the combined outcome measure when compared with the control group at post-test.

The third scale, the Social Privilege Measure (SPM; Black, Stone, Hutchinson, & Suarez, 2007), is a 25-item, self-report, Likert-type scale that measures an individual's awareness of racial privilege. The instrument generates five subscale scores including (a) *Environmental Predictability*, (b) *Penalty*, (c) *Personal Credibility*, (d) *Protection*, and (e) *Visibility* as well as a total score. Cronbach alpha reliability estimates for the subscale scores (range of .66-.88) and the total score (.92) were based on a sample of 312 graduate psychology and counseling students. The SPI's five-factor structure supported by the evidence from exploratory and confirmatory factor analyses (Black et al., 2007) suggests that awareness of privilege and oppression, particularly awareness of racism, may also be multidimensional.

While these three scales showed adequate psychometric properties, further development of the construct of awareness of privilege and oppression is needed for several

reasons. First, no follow-up studies utilizing the SPM (Black et al., 2007) were found in the literature, suggesting there is currently little to no empirical evidence from which to evaluate the efficacy of the SPM (Black et al., 2007) as an outcome measure for multicultural education courses. Second, SPM's (Black et al., 2007) focus solely on awareness of racial privilege and oppression may limit the instrument's utility in multicultural education outcome research where multiple forms of privilege and oppression are often covered during a semester-long course.

The POI (Hays, 2005; Hays et al., 2007) has several advantages over the SPM (Black et al., 2007). The POI (Hays, 2005; Hays et al., 2007) has been used in outcome research with graduate students, covers three of the four most frequently covered forms of privilege and oppression addressed in multicultural education courses (i.e., race, gender, and sexual orientation; Montross, 2003), and has stronger psychometric evidence supporting the measure's use in diversity training research. The POI, however, is not without limitations. First, the POI's focus on White privilege rather than a broader range of overall awareness of racial privilege and oppression may limit the utility of this measure when a researcher is seeking to study an individual's awareness of societal oppression experienced by people of color. Second, the POI does not cover SES-based awareness of privilege and oppression, which Montross (2003) identified as one of the four most common forms of privilege and oppression covered in multicultural education courses. Finally, the POI's utility as a diversity training measure for graduate, counseling, and psychology students may not generalize to a more heterogeneous group of undergraduate students, so more work needs to determine whether the POI is helpful for use with this growing multicultural education course population.

The original version of the APOS (Montross, 2003) is the only awareness of privilege and oppression measure noted above that has demonstrated an ability to measure expected growth in undergraduate students' overall awareness in a treatment versus control group study of a selection of multicultural education courses (see Remer, 2008). Also, the APOS includes subscales for the four most common types of privilege and oppression awareness covered in multicultural education courses: race, gender, sexual orientation, and social class. The APOS has also been used with a broad range of undergraduate students, which suggests the measure has demonstrated greater generalizability to a broader range of trainee populations than either the SPM (Black et al., 2007) or the POI (Hays, 2005; Hays et al., 2007) at this time.

The Current Studies

The current studies aim to contribute to the multicultural education outcome measurement literature by enhancing a measure used in the identity development components approach, the APOS (Montross, 2003). The assessment of an individual's awareness of privilege and oppression captures

multiple stages of the social identity development process, which is expected to occur during the multicultural training process. The APOS has shown several advantages over the other instruments of the awareness of privilege and oppression construct. The original APOS, however, had a few psychometric issues. First, many of the scale items loaded higher on unintended factors than intended factors. This overlap of items, however, was not consistent with Montross' (2003) use of orthogonal rotation during her EFA, which suggests she viewed the four hypothesized factors (e.g., awareness of racism, sexism, classism, and heterosexism) as separate and distinct domains. Utilization of oblique rotation during the EFA would have suggested Montross (2003) believed some overlap between the factors was appropriate. It is this research team's belief that there is overlap between these factors and that oblique rotation of the factors is more appropriate. Second, the original APOS subscales demonstrated low reliability estimates ($\alpha = .46$ for sexism, $\alpha = .56$ for classism) in Montross' (2003) study. In addition, no confirmatory studies were noted in the literature to confirm the factor Montross' (2003) proposed factor structure. Collectively, these challenges with the demonstrated psychometric properties of the APOS (Montross, 2003) suggest that a revision of the measure is warranted.

In Study 1, we review the APOS items in each dimension qualitatively, revise the items, and examine the dimensionality of the revised scale through EFA. In Study 2, we further examine the dimensionality of the revised APOS (the APOS-2) and the proposed second-order factor structure through CFA as well as provide additional evidence of convergent and discriminatory validity. Although awareness of privilege and oppression has been identified in the literature as a multidimensional construct with a hierarchical factor structure (Flammer, 2001; Hays, 2005; Hays et al., 2007; Montross, 2003), the empirical evidence is scarce. In addition to addressing the psychometric problems of the original APOS, the current studies extend the multicultural education outcome measurement literature by examining the second-order factor structure of the awareness of privilege and oppression construct through CFA.

Study 1

Method

APOS-2 construction. We revised the APOS through the following steps: elimination and retention of original APOS items, new item development (item revision), expert rater feedback, and participant administration (Clark & Watson, 1995).

Elimination and retention of the original APOS items. In Step 1, we used Montross' (2003) data collected from undergraduate students to identify items to retain for inclusion in the APOS-2. Inadequate items were defined as items that failed

to load on a factor at or above a factor loading coefficient of .30 as well as items that unexpectedly loaded higher on unintended factors. We retained 26 out of the 50 original APOS items after eliminating 24 items that did not meet these two criteria.

Item revision. In Step 2, the 26 items retained from the original APOS were then evaluated and revised by members of the research team. This team included one female, doctoral-level psychologist and multicultural education course instructor and three (two female and one male) graduate-level researchers enrolled in a doctoral program in psychology who were knowledgeable about the extant literature related to awareness of privilege and oppression and who had previous research experience with the original APOS (Montross, 2003) including collecting data on multicultural education courses. Before writing new items, the four researchers reviewed the literature on specific manifestations of each type of awareness of privilege and oppression (e.g., racism, sexism, classism, and heterosexism) included in the APOS-2 as well as common item-writing strategies (i.e., avoid double-barreled items). This item development group generated the initial list of APOS-2 items over three sessions. In the first session, researchers evaluated the 26 items retained from the original APOS and determined that the items were representative of the content provided in the literature review. A goal of this process was to ensure the content validity of the scale by linking each item to one of the four specific content domains of privilege and oppression included in the measure. In the second and third sessions, the group identified where the new items were needed and wrote items based on the literature for each of the four specific dimensions. The resulting item pool included a total of 107 items.

Expert rater feedback. In Step 3, subject matter experts (SMEs) evaluated the 107-item pool for the APOS-2. Eight experts were recruited to evaluate and provide feedback on the items via email. The SMEs had either a history of two or more publications relevant to content areas included in the APOS-2 (i.e., racism, sexism, heterosexism, and classism), practical experience teaching social justice–focused diversity training, or experience with social justice–focused advocacy work that included at least one of the specific content areas included in the APOS-2 (e.g., racism). The SMEs who agreed to participate in the review were sent an email containing a web link to the survey with 107 items as well as specific questions intended to solicit feedback on the individual items. The SMEs also provided feedback to the four proposed subscales and overall scale. Based on the SME feedback, we eliminated 28 items and revised 37 additional items. The item pool contained a total of 79 items including 26 items from the original APOS (Montross, 2003). The number of response categories was changed from four to six to increase variability in responses (Weng, 2004).

In Step 4, the items were administered to a group of research participants to assess the psychometric properties of the APOS-2. The resulting 79 items were designed to measure an individual's awareness of privilege and oppression in four areas: (a) racism, (b) sexism, (c) heterosexism, and (d) classism. These items consisted of 21 that represented awareness of racism, 20 represented awareness of classism, 20 related to sexism, and 18 related to heterosexism. The number of response categories was increased from four on the original APOS (Montross, 2003) to six on the APOS-2 and allowed participants to express their level of agreement using the following categories: 0 = *strongly disagree*, 1 = *disagree*, 2 = *slightly disagree*, 3 = *slightly agree*, 4 = *agree*, and 5 = *strongly agree*. Higher subscale and total scores represented higher levels of awareness in the area measured.

Participants. Collectively, 5,309 prospective participants were recruited and 484 participants completed at least 90% of the survey items (9.11% response rate). Three hundred nine participants (63.8%) from the sample were female, 171 (35.3%) were male, and four (0.8%) reported they were transgender. The mean age of the participant group was 20.72 years ($N = 483$, $SD = 4.08$ years, range = 18-59 years). The sample was also predominantly Caucasian ($n = 392$, 81%) with 6.2% representing international students ($n = 30$), 4.8% identified as multiracial ($n = 23$), 4.1% African American ($n = 20$), 2.9% Hispanic ($n = 14$), 0.4% Native American ($n = 2$), 0.2% Turkish American ($n = 1$), and 0.2% Pacific Islander ($n = 1$). The participant group was slightly more non-Christian ($n = 246$, 50.8%) than Christian ($n = 238$, 49.2%) and largely heterosexual ($n = 427$, 88.4%) compared with 6.5% who reported they were somewhat heterosexual/somewhat homosexual ($n = 36$) and 4.1% were exclusively homosexual ($n = 20$). The sample was slightly more upperclassmen and senior. The student classifications reported by participants in the combined group were as follows: Freshmen, 26.4 % ($n = 128$); sophomores, 18% ($n = 87$); juniors, 25.6% ($n = 124$); and seniors, 30% ($n = 145$).

Instruments. Two measures were administered to Study 1 participants: a demographic questionnaire and the 79-item version of the APOS-2. The demographic questionnaire was utilized to provide data on the sample characteristics of research participants. The 79-item draft of the APOS-2 was administered to provide data for the item evaluation, internal consistency, and exploratory factor analytic portions of this study (see the expert rater subsection above in this "Method" section for a detailed description of the measure).

Procedures. Data were collected from June to November of 2013 at a large, public university located in the Southeast. A list of potential participants were randomly selected by the Registrar's Office and provided to the research team. Individual participants were recruited via email to participate in

Table 1. Exploratory Factor Analysis Results of the Four-Factor APOS-2 Solution (*N* = 484).

Factor	Eigenvalue	% of variance after extraction	Cumulative % of variance after extraction
1. Heterosexism	7.52	26.60	26.60
2. Sexism	2.63	7.70	34.29
3. Classism	1.80	5.16	39.45
4. Racism	1.37	3.57	43.02

Note. APOS = Awareness of Privilege and Oppression Scale.

an Internet-based survey presented using Qualtrics, an Internet-based survey data collection tool. The survey included the demographic questionnaire and the 79-item draft of the APOS-2. Participants who completed the study were given the option of entering in a raffle where one US$15 gift card was randomly drawn and awarded for every 100 participants who volunteered to participate.

Results and Discussion

Item selection. The item retention and elimination decision-making process included the analysis of response distributions, estimates of internal consistency at the subscale and total scale levels, and EFA. The final, 26-item draft of the APOS-2 (see Supplemental material) was developed through an iterative process. For example, two items were removed from the classism subscale due to limited response variability (e.g., all participants selected *strongly disagree*), and other items were eliminated for failing to load higher on a predicted factor than nonpredicted factors. A factor loading cutoff score of .30 was utilized in the item-retention decision-making process with items below that value being eliminated from the scale. Consistent with Montross (2003), we hypothesized a four-factor solution made up of awareness of racism, sexism, classism, and heterosexism. A four-factor solution in which items generally loaded on the proposed factors (e.g., sexism-related content items loaded generally on a factor with other items that were constructed to measure awareness of sexism) emerged. The iterative process continued after a tenable factor solution was reached to reduce the number of items included in the final solution. In total, 53 of the 79 items administered to participants were eliminated through this iterative process before a final 26-item solution was reached.

Imputation. Little's (1988) chi-square test and a missing values analysis were conducted on the four subscales of the APOS-2 before analyzing the data to determine whether imputation techniques were appropriate in an effort to maximize data available for analysis. The Awareness of Racism ($\chi^2 = 454.042$, *df* = 450, *p* = .438), Awareness of Sexism ($\chi^2 = 305.757$, *df* = 302, *p* = .429), Awareness of Heterosexism ($\chi^2 = 52.841$, *df* = 42, *p* = .122), and Awareness of Classism ($\chi^2 = 30.118$, *df* = 20, *p* = .068) subscales were

not significant at the .05 level. Consequently, these data were determined to be MCAR (missing completely at random) and appropriate for imputation. None of the variables were missing more than 5% of the data prior to the imputation process. Values were substituted for the missing data using the expectation maximization (EM) method. This process resulted in 484 complete cases for data analysis purposes.

EFA. The four-factor solution that emerged from the data was evaluated to determine the acceptability of the solution using maximum likelihood estimation and oblique oblimin rotation. Our use of oblique rotation techniques differed from Montross' (2003) use of orthogonal rotation and was intended to address our hypothesis that there would be some overlap in the factors (oblique rotation) rather than four distinct factors with no overlap (orthogonal rotation) that was proposed by Montross. The Bartlett's test on the reduced 26-item APOS-2 data was significant ($\chi^2 = 4{,}580.159$, *df* = 325, *p* = .000 < .050), and the KMO (Kayser–Meyer–Olkin) value (.904) suggested the data were appropriate for EFA. The four-factor solution accounted for 43.02% of the total variance explained (see Table 1 for a list of the eigenvalues and total variance explained for the four-factor solution and Table 2 for a full list of the APOS-2 items and factor loading coefficients observed in the EFA organized by subscale). The item content of the four factors was consistent with the four theoretically derived subscales of the APOS-2.

Factor 1 contained seven awareness of heterosexism items and accounted for 26.60% of the variance using the extraction sum of squared loadings. A sample item includes "teenagers who identify as gay or lesbian in school are at a greater risk for being physically assaulted." Factor 2 contained six awareness of sexism items and accounted for 7.70% of the total variance. This factor included items such as "women are better-suited to stay at home to raise children than men." Seven items loaded onto Factor 3 representing awareness of classism. This factor accounted for 5.16% of the total variance and included awareness of classism-related items such as "being poor has no bearing on a person's opportunity to earn a college degree." Factor 4 accounted for 3.57% of the variance. The fourth factor consisted of six items that represented awareness of racism such as "people of color experience high levels of stress because of the discrimination they face."

Table 2. Exploratory Factor Analysis Factor Pattern Matrix for the Four-Factor APOS-2 Solution.

	Factor			
	1	2	3	4
Awareness of Heterosexism				
3. In many workplaces, some employees would have concerns about hiring a gay or lesbian employee rather than a heterosexual employee.	**.97**	−.04	−.09	.11
7. Teenagers who identify as gay or lesbian in school are at a greater risk for being physically assaulted than heterosexual teens.	**.30**	.18	.17	−.22
15. Gay men and lesbian women often have concerns about kissing their partners in public.	**.38**	.09	.04	−.16
18. Gay men are more at risk for being terminated from a job than heterosexual men based solely on sexual orientation.	**.46**	.06	.02	−.21
21. When meeting new people, gay men and lesbian women have to spend extra time trying to figure out if it is safe to reveal their sexual orientation.	**.58**	−.01	.11	−.04
23. Some hiring officials may not hire gay or lesbian workers to avoid negative reactions from customers.	**.84**	−.06	−.01	.07
24. For many gay men and lesbian women, the choice about where to vacation can depend on how open a city is to homosexuality.	**.37**	.10	.05	−.11
Awareness of Sexism				
1. Men should do less house cleaning than their female partners.[a]	−.01	**.68**	.01	−.01
5. Women are better suited to stay at home to raise children than men.[a]	−.04	**.73**	−.14	−.15
8. Women are better suited as entry-level employees when compared to men.	.01	**.44**	.11	.16
12. Women often mean "yes" when they say "no" to a man's advances.[a]	−.03	**.43**	.17	.11
19. Women who dress provocatively want men to approach them for sex.[a]	.07	**.53**	.02	−.05
26. Men are better leaders than women.[a]	.02	**.71**	−.16	−.12
Awareness of Classism				
2. People who have money are more likely to live longer than people who do not have much money.	.09	.03	**.62**	.11
6. The stress associated with being poor can cause health problems.	.06	.20	**.32**	−.21
10. People who live on the "good" side of town are less likely to become ill from industrial plants than other people.	.05	−.03	**.42**	−.22
14. Being poor has no bearing on a person's opportunity to earn a college degree.[a]	−.03	.08	**.67**	−.01
16. A person from an affluent family has a greater chance to earn a college degree than an individual from a poor family.	.03	−.04	**.74**	−.05
22. Poor individuals are more likely to suffer from mental illness because of the way society treats them.	−.04	−.03	**.40**	−.32
25. Growing up in a low-income family hurts a person's chances for obtaining a job that will make them happy.	.04	−.02	**.61**	−.17
Awareness of Racism				
4. African American political candidates are generally less likely to be accepted by White constituents in their districts.	.26	−.07	.13	**−.45**
9. People of Color experience high levels of stress because of the discrimination they face.	.02	.08	.02	**−.74**
11. Racism continues to play a prominent role in society.	.07	.04	−.06	**−.63**
13. Most history books don't accurately show how People of Color helped America become the country it is.	.11	.08	.03	**−.56**
17. African Americans with lighter skin color are more likely to be promoted within corporations than African Americans with darker skin color.	.17	−.08	.18	**−.47**
20. People of Color receive less medical information from their physicians when compared to White individuals.	.10	−.08	.24	**−.47**

Note. Bold values indicate the highest factor loading for each item.
[a]Reverse-scored item.

The average inter-factor correlation for the four-factor solution was .33, and the average subscale to total score correlation was .74. These findings suggest that individual factors were more strongly related to global awareness of privilege and oppression rather than to more specific types of awareness (e.g., the other factors). These findings support

Table 3. Comparison of Original APOS and APOS-2 Reliability Estimates (Coefficient Alpha).

Sample	Total	APOS/APOS 2 subscales			
		Heterosexism	Sexism	Classism	Racism
APOS (Montross, 2003)	.83	.75	.46	.56	.71
APOS-2					
Study 1	.89	.82	.76	.81	.84
Study 2	.88	.82	.76	.82	.80

Note. APOS *N* = 258; APOS-2 Study 1 *N* = 484; APOS-2 Study 2 *N* = 520. APOS = Awareness of Privilege and Oppression Scale.

the previous work of Flammer (2001), Hays (2005), and Hays et al. (2007) and provide additional empirical support for a two-tiered solution in which global awareness of privilege and oppression is made up of more specific types of awareness (e.g., racism or sexism). Furthermore, these findings were also consistent with McClellan's (2014) work which also suggested that specific types of awareness (e.g., heterosexism) are intercorrelated and, hence, require oblique factor rotation.

Three items, numbers 6, 7, and 22, are noteworthy because these three items loaded higher on their theoretically derived factor; however, they demonstrated higher than anticipated cross-loadings on other factors as well (see Table 2). Items 6 and 7 were both retained because the cross-loading values were below our .30 factor loading cutoff value, their strongest factor loading was on their theoretically derived factor, because our hypothesis was that these factors overlapped conceptually (hence our use of oblique factor rotation), after reviewing the feedback from our expert rates, and after confirmation that the item content was consistent with the literature. Item 22 was more challenging. This item's strongest factor loading was on its theoretically derived subscale; however, the cross-loadings for a second subscale were higher than our .30 cutoff value. In the end, we elected to retain Item 22 because its highest factor loading was on its theoretically derived subscale and because we had hypothesized this subscale overlapped conceptually with the other measures of awareness and to fully incorporate item content noted in our literature review. Furthermore, the theoretical overlap was consistent with previous findings in the awareness of privilege and oppression literature (see Hays, 2005; Hays et al., 2007), which employed oblique rotation during their EFA.

Reliability analysis. Reliability analysis of the final 26-item APOS-2 and each of the four subscales was performed on the sample of 484 participants (see Table 3 for a comparison of the Cronbach alpha reliability estimates for the original APOS and both studies of the APOS-2). The Cronbach alpha reliability estimate for the 26-item total score was .89. Item-total correlations ranged from $r = .20$ to $r = .62$ with a mean item-total correlation of $r = .46$. The four APOS-2 subscales demonstrated the following satisfactory internal consistency

estimates: Awareness of Heterosexism ($\alpha = .82$), Awareness of Sexism ($\alpha = .76$), Awareness of Classism ($\alpha = .81$), and Awareness of Racism ($\alpha = .84$). The mean inter-item total correlations for each of the four subscales were as follows: Awareness of Heterosexism ($r = .55$), Awareness of Sexism ($r = .47$), Awareness of Classism ($r = .55$), and Awareness of Racism ($r = .61$).

The Study 1 results showed that the 26-item APOS-2 has four dimensions that are consistent with the literature of awareness of privilege and oppression. Compared with the original APOS (Montross, 2003), the APOS-2 demonstrated stronger factor loading of the item to its theoretically derived dimension than the other factors and improved subscale internal consistency.

Scoring. All of the items on the APOS-2 were scored from 1 to 6. The means and scoring ranges were calculated for the 484 participants after applicable items were reverse-scored. The mean scores and standard deviations for the 26-item APOS-2 total and subscale scores are included in Table 4.

Study 2

The purpose of Study 2 was to confirm the dimensionality of the APOS-2 and examine the second-order factor structure using CFA. In addition, the nomological network of the APOS-2 was examined. Specifically, we evaluated the relationship between the APOS-2 and another scale, the EMC/RSEE (Mallinckrodt et al., 2014) scale. Three EMC/RSEE subscales (Cultural Openness and Desire to Learn, Awareness of Contemporary Racism and Privilege, and Empathic Feeling and Acting as an Ally) were identified as measuring constructs that were conceptually similar to the APOS-2. We expected moderate and positive correlations with between the APOS-2 and these EMC/RSEE subscales and the weak correlations between the APOS-2 and the rest of EMC/RSEE subscales.

Method

Participants. The Study 2 sample consisted of 520 undergraduate students at a second-large public university (compared

Table 4. APOS-2 Means and Standard Deviations for Study 1 and Study 2.

APOS-2 scale	Study 1		Study 2	
	M	SD	M	SD
Total Score	105.93	17.04	105.36	16.15
Heterosexism	29.99	5.49	28.51	5.93
Sexism	27.30	5.18	28.85	4.89
Classism	27.79	6.33	26.48	6.45
Racism	20.85	5.49	21.48	5.14

Note. Study 1 N = 484; Study 2 N = 520. APOS = Awareness of Privilege and Oppression Scale.

to Study 1) located in the Southeast. The participants were predominantly female. Four hundred sixteen participants (80.0%) were female, 102 (19.6%) were male, and two (.4%) reported they were transgender. The mean age of participants was 21.71 years ($N = 520$, $SD = 6.57$ years, range = 18-58 years). The combined sample was also predominantly Caucasian ($n = 473$, 91.0%) with 4.0% identifying as African American ($n = 21$), 1.7% multiracial ($n = 9$), 1.3% Hispanic ($n = 7$), 1.0% Asian American ($n = 5$), 0.4% Native American ($n = 2$), 0.4% International ($n = 2$), and 0.2% as Other ($n = 1$). The sample was more Christian ($n = 343$, 66.0%) than non-Christian ($n = 177$, 34.0%) and largely heterosexual ($n = 445$, 85.6%) compared with 11.0% who reported they were somewhat heterosexual/somewhat homosexual ($n = 57$) and 3.5% were exclusively homosexual ($n = 18$). The student classifications reported by participants were as follows: Freshmen, 44.8 % ($n = 233$); sophomores, 18.3% ($n = 95$); juniors, 19.2% ($n = 100$); and seniors 17.7% ($n = 92$).

Instruments

APOS-2. The 26-item APOS-2 (see Supplemental material) is designed to measure an individual's awareness of privilege and oppression with a total score and the following four subscales: Awareness of Racism, Awareness of Sexism, Awareness of Classism, and Awareness of Heterosexism. Participants indicated their agreement with each statement using six response categories: 1 (*strongly disagree*), 2 (*disagree*), 3 (*slightly disagree*), 4 (*slightly agree*), 5 (*agree*), and 6 (*strongly agree*). Higher subscale and total scores represent higher levels of awareness in the area measured.

EMC/RSEE. The EMC/RSEE (Mallinckrodt et al., 2014) is a 48-item instrument designed as an outcome measure for multicultural programming. Similar to the APOS, the EMC/RSEE measures the specific components of social identity development. The six subscales include (a) Cultural Openness and Desire to Learn, (b) Resentment and Cultural Dominance, (c) Anxiety and Lack of Multicultural Self-Efficacy, (d) Empathic Perspective-Taking, (e) Awareness of Contemporary Racism and Privilege, and (f) Empathic Feeling and Acting as an Ally. The six response categories and scoring values for each

response are as follows: 1 (*strongly disagree*), 2 (*disagree*), 3 (*slightly disagree*), 4 (*slightly agree*), 5 (*agree*), and 6 (*strongly agree*). Higher subscale scores represent higher levels of the competency being measured. The reliability estimates for the six subscales of the EMC/SEE were as follows: (a) Cultural Openness and Desire to Learn ($\alpha = .89$), (b) Resentment and Cultural Dominance ($\alpha = .87$), (c) Anxiety and Lack of Multicultural Self-Efficacy ($\alpha = .77$), (d) Empathic Perspective-Taking ($\alpha = .64$), (e) Awareness of Contemporary Racism and Privilege ($\alpha = .85$), and (f) Empathic Feeling and Acting as an Ally ($\alpha = .78$).

Procedure. Data were collected during the fall 2015 and spring 2016 terms at a second-large public university (compared to Study 1) located in the Southeast. The participants were recruited from various undergraduate psychology courses and received research credits for their courses and represented a convenience sample. Participants completed an online survey that contains a demographic questionnaire, the 26-item APOS-2, and the EMC/RSEE (Mallinckrodt et al., 2014).

Results and Discussion

CFA. CFAs were conducted using maximum likelihood estimation as implemented in Mplus 6.0 (Muthén & Muthén, 2010; see Table 5 for the CFA model comparisons). In the first model, all 26 items were allowed to load on a single factor (i.e., Awareness of Privilege and Oppression). The second model is a first-order orthogonal model where all items loaded on the designated dimension (i.e., Awareness of Heterosexism, Awareness of Sexism, Awareness of Classism, and Awareness of Racism) and these dimensions were not correlated. The third model is the same as the second model except for the four factors were allowed to correlate (i.e., oblique). The fourth model is a second-order model in which all items were allowed to load on their designated dimensions and all dimensions loaded on a second-order factor of Awareness of Privilege and Oppression.

The fit indices for each model are presented in Table 5. The smaller model chi-square (χ^2) indicates the model has a better fit to the data (Kline, 2005). Comparative fit index

Table 5. Fit Indices for Confirmatory Factor Analysis of the APOS-2.

Model	χ^2	df	CFI	RMSEA (90% CI)	SRMR
1. One-factor model	1,829.09	299	.66	.10 [.095, .104]	.09
2. First-order orthogonal four-factor model	1,582.08	299	.71	.09 [.086, .095]	.19
3. First-order oblique four-factor model	1,037.74	293	.83	.07 [.065, .075]	.08
4. Second-order factor model	1,039.91	295	.83	.07 [.065, .074]	.08

Note. APOS-2 = Awareness of Privilege and Oppression Scale–2; CFI = comparative fit index; RMSEA = root mean square error of approximation; 90% CI = 90% Confidence Interval for RMSEA (lower limit, upper limit); SRMR = standardized root mean square residual.

(CFI) ratings above .90 suggest a reasonably good fit of the model (Hu & Bentler, 1999). Standardized root mean square residual (SRMR) suggests the difference between observed and predicted covariances, and values below .10 are preferred (Kline, 2005). Root mean square error of approximation (RMSEA) indicates the amount error of approximation per degree of freedom, and a smaller value suggests a better fit of the model. General guideline for interpretation suggests an approximate fit when RMSEA is below .05, a reasonable fit when RMSEA is between .05 and .08, and a poor model fit with RMSEA above .10 (Browne & Cudeck, 1993).

Overall, the four-factor model—including the second-order model—showed better fit than the single-factor model. This finding supports the hypothesized four-factor structure of the APOS-2. The superior fit of the four-factor oblique model over the four-factor orthogonal model suggests that the first-order factors, specific domains of awareness of privilege and oppression, are interrelated. Finally, the results for the first-order oblique model and the second-order model suggested that both models fit the data equally well. While the RMSEA and SRMR values indicated a reasonable fit of both models, the CFI values of these models are below the recommended value for a good fit (.90). The possible explanation of poor fit will be explored later in the "General Discussion" section. Considering the comparable fit of these two models and the theoretical background, the second-order model ($\chi^2 = 1,039.91$, $df = 295$, CFI = .83, RMSEA = .07, SRMR = .08) was preferred. The hierarchical model in which overall awareness of privilege and oppression is made up of specific types of awareness (e.g., awareness of racism) is better aligned with the theory of the model and with the literature (Flammer, 2001; Hays, 2005; Hays et al., 2007; Montross, 2003). The factor loadings for the preferred second-order model are presented in Table 6.

Reliability analysis. Internal consistency estimates of the APOS-2 total score and subscale scores for Study 2 were adequate (see Table 3 for Studies 1 and 2 reliability estimates). The Cronbach's alpha for the APOS-2 total score was .88. The four APOS-2 subscales demonstrated satisfactory internal consistency: Awareness of Heterosexism ($\alpha = .82$), Awareness of Sexism ($\alpha = .76$), Awareness of Classism ($\alpha = .82$), and Awareness of Racism ($\alpha = .80$).

Scoring. All of the items on the APOS-2 were scored from 1 to 6. The means and scoring ranges were calculated for the 520 participants after applicable items were reverse-scored. The mean scores and standard deviations for the 26-item APOS-2 total and subscale scores for Study 2 are included in Table 4.

Convergent and discriminant validity. The observed Pearson's correlations between the APOS-2 and the EMC/RSEE subscales are summarized in Table 7. The APOS-2 total and four subscales scores showed a moderate to strong positive correlation to each of the three conceptually similar subscales of the EMC/RSEE (i.e., Cultural Openness and Desire to Learn, Awareness of Contemporary Racism and Privilege, and Empathic Feeling and Acting as an Ally). In particular, the strong positive relationship between the APOS-2 total and four subscale scores (especially the Awareness of Racism subscale) and the EMC/RSEE's Awareness of Contemporary Racism and Privilege subscale (ranging from $r = .32$ to .71) is important because the content between these measures are conceptually more closely related because these scales measure awareness of privilege and oppression. This result provides evidence of convergent validity for the APOS-2.

Not all of the EMC/RSEE subscales appeared to have content that was conceptually similar to the APOS-2. The Resentment and Cultural Dominance, Anxiety and Lack of Multicultural Self-Efficacy, and Empathic Perspective-Taking subscales all three appeared to represent content that was conceptually less similar when compared with the APOS-2. For example, it is possible for an individual to be more aware of social privilege and oppression, and this awareness does not necessarily equate to lower levels of anxiety when this same individual is around people who are different from them. The observed Pearson's correlations between the APOS-2 and these three subscales ranged from $-.05$ to $-.43$ (see Table 7). This result provides evidence of the APOS-2's discriminant validity.

Table 6. APOS-2 Items and Factor Loadings From the Second-Order CFA Model.

Items	First-order loadings	Second-order loadings
Awareness of Heterosexism		
3. In many workplaces, some employees would have concerns about hiring a gay or lesbian employee rather than a heterosexual employee.	.46	.89
7. Teenagers who identify as gay or lesbian in school are at a greater risk for being physically assaulted than heterosexual teens.	.61	
15. Gay men and lesbian women often have concerns about kissing their partners in public.	.61	
18. Gay men are more at risk for being terminated from a job than heterosexual men based solely on sexual orientation.	.76	
21. When meeting new people, gay men and lesbian women have to spend extra time trying to figure out if it is safe to reveal their sexual orientation.	.68	
23. Some hiring officials may not hire gay or lesbian workers to avoid negative reactions from customers.	.71	
24. For many gay men and lesbian women, the choice about where to vacation can depend on how open a city is to homosexuality.	.60	
Awareness of Sexism		
1. Men should do less house cleaning than their female partners.[a]	.13	.20
5. Women are better suited to stay at home to raise children than men.[a]	.69	
8. Women are better suited as entry-level employees when compared to men.	.56	
12. Women often mean "yes" when they say "no" to a man's advances.[a]	.54	
19. Women who dress provocatively want men to approach them for sex.[a]	.53	
26. Men are better leaders than women.[a]	.59	
Awareness of Classism		
2. People who have money are more likely to live longer than people who do not have much money.	.64	.70
6. The stress associated with being poor can cause health problems.	.56	
10. People who live on the "good" side of town are less likely to become ill from industrial plants than other people.	.61	
14. Being poor has no bearing on a person's opportunity to earn a college degree.[a]	.50	
16. A person from an affluent family has a greater chance to earn a college degree than an individual from a poor family.	.72	
22. Poor individuals are more likely to suffer from mental illness because of the way society treats them.	.67	
25. Growing up in a low-income family hurts a person's chances for obtaining a job that will make them happy.	.73	
Awareness of Racism		
4. African American political candidates are generally less likely to be accepted by White constituents in their districts.	.54	.93
9. People of Color experience high levels of stress because of the discrimination they face.	.70	
11. Racism continues to play a prominent role in society.	.55	
13. Most history books don't accurately show how People of Color helped America become the country it is.	.56	
17. African Americans with lighter skin color are more likely to be promoted within corporations than African Americans with darker skin color.	.79	
20. People of Color receive less medical information from their physicians when compared to White individuals.	.62	

Note. APOS–2 = Awareness of Privilege and Oppression Scale–2; CFA = confirmatory factor analysis.
[a]Reverse-scored item.

General Discussion

The purpose of the included studies was to improve the multicultural education outcome measurement literature by revising a scale (the APOS-2) that addresses the psychometric issues of the original APOS (Montross, 2003) and by testing the proposed hierarchical nature of the awareness of privilege and oppression construct. A primary proposition in this adapted model was the importance of creating knowledge-based test items that are tied to the extant theory and literature. A comprehensive literature review and a panel of expert reviewers with specific knowledge of the content

Table 7. APOS-2 Correlations With Comparison Measures.

| APOS-2 scale | EMC/RSEE | | | | | |
| | Factor | | | | | |
	1	2	3	4	5	6
Total Score	.41**	−.43**	−.19**	−.05	.71**	.40**
Heterosexism	.20**	−.23**	−.03**	−.13**	.58**	.21**
Sexism	.44**	−.49**	−.43**	.06	.32**	.47**
Classism	.26**	−.20**	−.09	−.01	.40**	.20**
Racism	.32**	−.36**	−.04	−.03	.73**	.30**

Note. APOS-2 = Awareness of Privilege and Oppression Scale–2; EMC/RSEE = Everyday Multicultural Competencies/Revised Scale of Ethnocultural Empathy; Factor 1 = Cultural Openness and Desire to Learn; Factor 2 = Resentment and Cultural Dominance; Factor 3 = Anxiety and Lack of Multicultural Self-Efficacy; Factor 4 = Empathetic Perspective-Taking; Factor 5 = Awareness of Contemporary Racism and Privilege; Factor 6 = Empathetic Feeling and Acting as an Ally.
$^*p < .05.$ $^{**}p < .01.$

areas established the content validity of the items in the APOS-2. Both EFA (Study 1) and CFA (Study 2) results supported the four dimensions of the scale. The APOS-2 also showed improvements in the reliability estimates of the four subscales compared to the original APOS (see Table 3). Furthermore, the Study 2 results provided evidence of convergent and discriminant validity for the revised scale. Finally, the hierarchical CFA provided the preliminary statistical support for our proposed hierarchical relationship between the broad construct of awareness of privilege and oppression and awareness in the four specific domains.

The reliability estimates of the APOS-2 total score and the four theoretically derived subscales represent an improvement over the original APOS (see Table 3). The reliability estimates for all APOS-2 scores improved over the data presented in Montross (2003) original APOS study. Although the total score reliability estimate (α) for the original APOS was .83 (Montross, 2003), the APOS-2 showed higher estimates in both Study 1 (α = .89) and Study 2 (α = .88). The most notable improvements in internal consistency were in the Awareness of Sexism (α = .46 in the original APOS vs. α = .76 in the APOS-2) and Awareness of Classism subscales (α = .56 in the original APOS vs. α = .81-.82 in the APOS-2). These improved reliability estimates will allow researchers and educators to utilize the subscale scores for research or evaluation purposes in a way that wasn't appropriate with the original APOS (Montross, 2003).

The reduction in the overall number of items presented to participants also represents an enhanced feature of the APOS-2 when compared with the original APOS. The number of items included in the total score has been reduced by 24 (50 items in the original APOS vs. 26 items in the APOS-2). In addition, the number of items from each subscale is more balanced than observed in the original APOS. For example, the four subscales of the original APOS ranged from seven to 15 items in length, while the four subscales of the APOS-2 range from six to seven items per subscale. This

shorter and more balanced second edition of the measure provides an assessment tool for multicultural education that is realistic for implementation and will reduce the amount of time participants spend for administration purposes.

An intentional effort was made to ensure the content validity of the APOS-2. An individual item was only retained from the original APOS, if the item was both psychometrically desirable and included content observed in an updated review of the literature. The new items constructed for the APOS-2 were based in knowledge and concepts observed within the extant literature and were created by a group of social justice–focused researchers with specific knowledge of the content areas and with research experience utilizing the original APOS. Both the original APOS and the APOS-2 utilized a panel of expert raters to review item content. The expert rater panel for the APOS-2 was more diverse in terms of numbers (the original APOS utilized three expert raters vs. eight in the APOS-2) and specificity (at least one expert in the subject matter for each theoretically derived subscale provided feedback for the APOS-2). The feedback provided by an expert panel with more specific knowledge of the content areas is an important distinction between the item development processes employed in the original and updated versions of the instrument.

Black et al. (2007), Flammer (2001), Hays (2005), Hays et al. (2007), and Montross (2003) provided empirical support for the multidimensionality of awareness of privilege and oppression. The proposed oblique, four-factor structure of the APOS-2, which was theoretically constructed to measure awareness of heterosexism, sexism, classism, and racism, was supported by the data through an EFA in Study 1. A series of CFAs showed that the oblique four-factor model and the hierarchical four-factor model demonstrated comparable fit and were the best available fit for the data. These findings add to the awareness of privilege and oppression research by providing additional support for the theory that awareness of privilege and oppression is best described by a

hierarchical factor structure in which global awareness of privilege and oppression is made up of more specific subtypes of awareness (e.g., sexism) that are intercorrelated. Allowing the four factors to correlate in the factor analysis process and the use of CFA to evaluate the proposed hierarchical factor structure of the APOS-2 represents an improvement over Montross' (2003) original version of the measure. The fact that the proposed model did not meet the desired criteria for model fit across all of the fit indices suggests more work can and should be done to better clarify this construct in future studies.

This discussion of the dimensionality of the APOS-2 has real-world implications for multicultural education researchers and instructors. A primary goal in many multicultural education courses (D. J. Goodman, 2001) is to teach students about awareness of privilege and oppression. The current studies as well as the work of Flammer (2001), Hays (2005), Hays et al. (2007), and Montross (2003) all suggest that awareness of privilege and oppression is best represented by a model that suggests overall awareness of privilege and oppression is made up of specific types of privilege and oppression that overlap to some extent. This means that multicultural education researchers and instructors should consider utilizing measures that focus on the target areas of instruction when selecting measures to measure student development. For example, administering an awareness of racism scale in a women and gender studies class may not be able to accurately measure student growth and advancement within social identity development models because the specific content of the class is not being measured by a scale that measures awareness of racism. Being mindful of this assessment approach could allow researchers and instructors to better assess student growth within a social identity development context.

Limitations

Three limitations of the presented studies are noteworthy. The results of the model fit indices noted in the CFA (see Table 5), the continually low reliability estimates of the Awareness of Sexism subscale, and the representativeness of the two samples to the larger student population in the United States are three important limitation of the current studies. The current 26-item APOS-2 and the hierarchical four-factor solution represent the best possible fit of the data and the theory when compared to other potential models. However, the fit index, specifically the CFI, suggested the hierarchical four-factor solution was not a great overall fit of the data. One possible source of poor fit is the Awareness of Sexism subscale. Compared to the other subscales, the Awareness of Sexism subscale was linked to overall awareness of privilege and oppression to a lesser degree. Hays et al. (2007) had previously provided strong support for the hierarchical four-factor model using some subscale content areas that overlap with subscales presented in the APOS-2. It is possible that

the poor model fit could be due to the strong correlations between the Awareness of Heterosexism and Racism subscales, or perhaps the theoretically derived hierarchical factor structure of awareness of privilege and oppression does not apply equally to all subtypes of awareness (e.g., racism). One other explanation for the lack of fit could be related to the sample characteristics. The Sample 1 participants were generally more diverse while the Study 2 participants were less diverse collectively, and it is possible that the lack of diverse representation in the Study 2 group restricted the available responses and had a negative impact on the data. Perhaps participant responses on the APOS-2 may vary by geographic region or university setting or location. In the present study, we collected data from two separate universities located within the same geographic region in an effort to diversify the undergraduate student pool. Given that this measure was administered at universities with large White student populations, there is no way to know students at, for example, a historically Black college or university might respond given their differing contexts.

Future Research

We offer three areas of suggestion for future research. These include follow-up revisions on the Awareness of Sexism subscale, utilizing the APOS-2 in actual multicultural education course outcome research, and essential periodic updates of the instrument. First, future researchers should focus on the Awareness of Sexism subscale. This subscale was Montross' (2003) lowest Cronbach alpha reliability estimate (.46) in the original APOS. In the current study, the Awareness of Sexism subscale reliability was better (.76 for the two samples evaluated in Studies 1 and 2 for the APOS-2 versus .46 for the original APOS subscale), but it proved the most challenging of the subscales on the APOS-2 to construct. Future revisions to this subscale should consider the multidimensional nature of sexism and continue to refine item-writing strategies to better include the full range of content observed in the literature. It is possible that improvements in the sexism subscale may lead to a better model fit during future factor analytic studies.

Future research should also utilize the APOS-2 in multicultural education course research and to examine the utility of the measure with both undergraduate and graduate student populations. Remer (2008) provided evidentiary support for utilizing the original APOS as an undergraduate multicultural education course outcome measure. This type of research is vital to providing the type of empirical support necessary for gatekeepers who may approve this type of training within their universities, organizations, and schools in the future. Remer's (2008) work focused on undergraduates, and the original APOS (Montross, 2003) was employed to measure progress in full-semester academic courses. These are likely the type of learning environments where change will be most significant and easier to evaluate with instruments such as the

APOS-2 because these types of courses often last for extensive periods of time and cover a number of topics that often overlap with the APOS-2 content.

Finally, it is important to note that periodic revisions of the APOS-2 will need to occur to update the item content. The items included in the APOS-2 were literature-driven and based on current research related to societal manifestations of privilege and oppression. It is likely that the literature describing these manifestations will change over time. For example, one of the items that was considered for inclusion in the APOS-2 involved a lack of health care for individuals from lower SES backgrounds. This item was removed during the development of the APOS-2 because of the passage of the Affordable Care Act which has made healthcare more widely available in the United States. This measure will need to be updated periodically as laws, social norms, and new manifestations of privilege and oppression change for the instrument to remain relevant. The APOS-2 has strong potential to serve as an effective tool for multicultural education course researchers and instructors who seek construct-relevant measurement tools with good psychometric data.

Authors' Note

Declaration of Conflicting Interests

Funding

ORCID iD

Michael James McClellan [iD] https://orcid.org/0000-0003-4094-0999

References

Bezrukova, K., Jehn, K. A., & Spell, C. S. (2012). Reviewing diversity training: Where we have been and where we should go. *Academy of Management Learning & Education*, *11*, 207-227. doi:10.5465/amle.2008.0090

Black, L. L., Stone, D. A., Hutchinson, S. R., & Suarez, E. C. (2007). The development and validation of the Social Privilege Measure. *Measurement and Evaluation in Counseling and Development*, *40*, 16-32.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Byrd, R., & Hays, D. G. (2013). Evaluating a safe space training for school counselors and trainees using a randomized control group design. *Professional School Counseling*, *17*, 20-31. doi:10.5330/PSC.n.2013-17.20

Cass, V. C. (1979). Homosexuality identity formation: A theoretical model. *Journal of Homosexuality*, *4*, 219-235. doi:10.1300/J082v04n03_01

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*, 309-319. doi:10.1037/1040-3590.7.3.309

Diemer, M. A., & Hsieh, C. (2008). Sociopolitical development and vocational expectations among lower socioeconomic status adolescents of color. *The Career Development Quarterly*, *56*, 257-267. doi:10.1002/j.2161-0045.2008.tb00040.x

Diemer, M. A., McWhirter, E. H., Ozer, E. J., & Rapa, L. J. (2015). Advances in the conceptualization and measurement of critical consciousness. *The Urban Review*, *47*, 809-823. doi:10.1007/s11256-015-0336-7

Diemer, M. A., Rapa, L. J., Voight, A. M., & McWhirter, E. H. (2016). Critical consciousness: A developmental approach to addressing marginalization and oppression. *Child Development Perspectives*, *10*, 216-221. doi:10.1111/cdep.12193

Downing, N. E., & Roush, K. L. (1985). From passive acceptance to active commitment: A model of feminist identity development for women. *The Counseling Psychologist*, *13*, 695-709. doi:10.1177/0011000085134013

Flammer, L. J. (2001). *The nature of prejudice: Dimensions and patterns of racism, sexism, classism, and heterosexism among social groups* (Doctoral dissertation). Available from ProQuest. (Document No. 3014431)

Goodman, D. J. (2001). *Promoting diversity and social justice: Educating people from privileged groups*. Thousand Oaks, CA: Sage.

Hays, D. G. (2005). *Initial development and validation of the Privilege and Oppression Inventory* (Doctoral dissertation). Available from ProQuest. (Document No. 888840921)

Hays, D. G., Chang, Y. C., & Decker, S. L. (2007). Initial development and psychometric data for the Privilege and Oppression Inventory. *Measurement and Evaluation in Counseling and Development*, *40*, 66-79. Available from http://mec.sagepub.com/

Helms, J. E. (1990). *Black and White racial identity: Theory, research, and practice*. Westport, CT: Greenwood Press.

Howard-Hamilton, M. F., Cuyjet, M. J., & Cooper, D. L. (2011). Understanding multiculturalism and multicultural competence among college students. In M. J. Cuyjet, M. F. Howard-Hamilton, & D. L. Cooper (Eds.), *Multiculturalism on campus: Theory, models, and practices for understanding diversity and creating inclusion* (pp. 11-18). Sterling, VA: Stylus.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

Kim, B. S. K., Cartwright, B. W., Asay, P. A., & D'Andrea, M. J. (2003). A revision of the Multicultural Awareness, Knowledge, and Skills Survey–Counselor Edition. *Measurement and Evaluation in Counseling and Development*, *36*, 161-180. Available from http://mec.sagepub.com/

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.

Little, R. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*, 1198-1202. doi:10.1080/01621459.1988.10478722

Mallinckrodt, B., Miles, J. R., Bhaskar, T., Chery, N., Choi, G., & Sung, M.-R. (2014). Development and psychometric evaluation of a comprehensive scale to assess college diversity programming. *Journal of Counseling Psychology*, *61*, 133-145. doi:10.1037/a0035214

McClellan, M. J. (2014). *Awareness of Privilege and Oppression Scale-2: Construction and initial validation* (Doctoral dissertation, University of Kentucky, Lexington). Retrieved from http://uknowledge.uky.edu/edp_etds/22

McWhirter, E. H., & McWhirter, B. T. (2016). Critical consciousness and vocational development among Latina/o high school youth: Initial development and testing of a measure. *Journal of Career Assessment*, *24*, 543-558. doi:10.1177/1069072715599535

Montross, L. P. (2003). *Awareness of Privilege and Oppression Scale: Construction and initial validation* (Doctoral dissertation). Available from ProQuest. (Document No. 765146431)

Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Author.

O'Meara, T. M. (2001). *The impact of undergraduate multicultural courses on White students' racial identity development* (Doctoral dissertation). Available from ProQuest. (Document No. 764909811)

Remer, R. R. (2008). *Influence of diversity courses on undergraduates' ethnocultural empathy, openness to diversity, and awareness of privilege and oppression within a mastery or performance classroom context* (Doctoral dissertation). Available from ProQuest. (Document No. 1475174891)

Shin, R. Q., Ezeofor, I., Smith, L. C., Welch, J. C., & Goodrich, K. M. (2016). The development and validation of the Contemporary Critical Consciousness Measure. *Journal of Counseling Psychology*, *63*, 210-223. doi:10.1037/cou0000137

Sue, W. S., & Sue, D. (1990). *Counseling the culturally different: Theory and practice* (2nd ed.). New York, NY: John Wiley.

Sue, W. S., & Sue, D. (1999). *Counseling the culturally different: Theory and practice* (3rd ed.). New York, NY: John Wiley.

Thomas, A. J., Barrie, Rabiatu, B., Brunner, J., Clawson, A., Hewitt, A., Jeremie-Brink, G., & Rowe-Johnson, M. (2014). Assessing critical consciousness in youth and young adults. *Journal of Research on Adolescence*, *24*(3), 485-496. https://doi.org/10.1111/jora.12132

Weng, J. J. (2004). Impact of the number of response categories and anchor labels on coefficient alpha and test-retest reliability. *Educational and Psychological Measurement*, *64*, 956-972. doi:10.1177/0013164404268674

Worell, J., & Remer, P. (2003). *Feminist perspectives in therapy: Empowering diverse women*. Hoboken, NJ: John Wiley.

Author Biographies

Michael James McClellan, PhD is an assistant professor and a licensed psychologist in the Department of Psychology at Eastern Kentucky University. He teaches courses in multicultural psychology, rural mental health, addictions, clinical supervision, practicum, and internship and his research interests include diversity education, telehealth, and addiction treatment. He received his PhD in counseling psychology from the University of Kentucky.

Lori Patricia Montross-Thomas, PhD is a licensed counseling psychologist and an assistant adjunct professor in Family Medicine and Public Health at the University of California, San Diego, La Jolla. She received her PhD in counseling psychology from the University of Kentucky.

Pamela Remer, PhD is an emeritus faculty member in the Counseling Psychology at the University of Kentucky. She is a council of Representatives Member for Division 35 (Society for the Psychology of Women) with the American Psychological Association.

Yoshie Nakai, PhD is an associate professor in the Department of Psychology at Eastern Kentucky University. She received her PhD in industrial/gerontological psychology from the University of Akron. Her research interests include an aging workforce, job search behaviors.

Addison D. Monroe, MS is a recent graduate of Eastern Kentucky University where he obtained a master's degree in general psychology. He manages a research lab in the Department of Psychology at the University of Kentucky and is planning to persue a doctoral degree in clinical/counseling psychology and to continue to persue multicultural education research.