

Development and Validation of the Ableist Microaggressions Scale

The Counseling Psychologist

2017, Vol. 45(4) 570–599

© The Author(s) 2017

Reprints and permissions:

sagepub.com/journalsPermissions.nav

DOI: [10.1177/0011000017715317](https://doi.org/10.1177/0011000017715317)

journals.sagepub.com/home/tcp



**Kristin J. Conover¹, Tania Israel²,
and Karen Nylund-Gibson²**

Abstract

Overt and covert experiences of prejudice and discrimination have been associated with increased stress and distress among people with disabilities. Although researchers have explored people with disabilities' experiences with microaggressions, no measure exists to assess disability microaggressions. The present research was conducted to develop and validate a scale that measures microaggressions experienced by people with physical disabilities, the Ableist Microaggressions Scale (AMS). Study 1 developed the preliminary AMS based on eight primary disability microaggressions domains outlined in Keller and Galgay's (2010) qualitative study, as well as expert feedback, cognitive interviews, and a pilot study. Study 2 consisted of exploratory factor analysis ($N = 559$) to determine factor structure, and a 4-factor model emerged. Study 3 confirmed the factor structure through confirmatory factor analysis ($N = 833$), and demonstrated adequate validity and reliability for the 20-item AMS. Additional findings, implications, and limitations are discussed.

Keywords

ableism, disability, microaggressions, measure, factor analysis

¹Alliant International University, Los Angeles, CA, USA

²University of California, Santa Barbara, Santa Barbara, CA, USA

Corresponding Author:

Kristin J. Conover, California School of Professional Psychology, Alliant International University, Los Angeles, 1000 South Fremont Avenue #5, Alhambra, CA 91803, USA.

Email: kristin.conover@alliant.edu

Minority stress theory posits that nondominant group members experience chronic stress related to social stigma, marginalization, objectification, and discrimination, which leads to health disparities (Brooks, 1981; Meyer, 1995, 2003). Physical disability status has been conceptualized as a form of chronic stress (Turner & McLean, 1989) and has been associated with higher rates of anxiety, depression, substance abuse, and suicidality (Meltzer et al., 2012; Scott et al., 2009; Turner, Lloyd, & Taylor, 2006). This finding was supported in Gayman, Brown, and Ciu's (2011) study of 1,459 people with disabilities, in which they found that social stress, including daily discrimination and recent life events, was associated with depression and pain. Higher rates of illicit drug use have also been explicitly linked to perceived discrimination in this population (Li & Moore, 2001). In addition to elevated risks for individuals with physical disabilities, negative disability-related events and perceived discrimination have been found to adversely influence the mental health of their family members as well (Green, 2003; Mazur, 2008). Physical disability-related minority stress, therefore, appears to have broad negative consequences for mental health and well-being.

The most insidious and harmful prejudices often take subtle or covert forms, called microaggressions (Sue et al., 2007). Microaggressions are brief verbal or behavioral indignities that communicate a slight or insult toward someone based on social group membership (Pierce, Carew, Pierce-Gonzalez, & Willis, 1978). Whereas overt or intentional prejudice and discrimination may be addressed in laws and policies, microaggressions are inadvertent expressions of bias embedded in daily interactions. The covert nature of microaggressions means they are often ambiguous and the recipient may be unsure whether his or her perception of the exchange was accurate (Sue, 2010; Sue, Capodilupo, & Holder, 2008). In cases where a microaggression is apparent, the targeted individual must then decide how to respond. Given that microaggressions may be well intentioned or believed by the perpetrator to be innocuous, a response or nonresponse could have negative consequences for the target (Sue et al., 2007). Public discourse about microaggressions is on the rise (e.g., Vega, 2014), providing greater attention to these experiences and their impact.

Numerous authors have suggested that microaggressions contribute to minority stress among a variety of nondominant groups (Keller & Galgay, 2010; Nadal, 2013; Nadal, Issa, Griffin, Hamit, & Lyons, 2010; Nadal, Skolnik, & Wong, 2012; Sue et al., 2007; Wright & Wegner, 2012), but the mental and physical health effects of microaggressions are just beginning to be empirically determined. Several scales have been developed and published to quantitatively measure microaggressions, including racial and ethnic microaggressions (Mercer, Zeigler-Hill, Wallace, & Hayes, 2011; Nadal,

2011; Torres-Harding, Andrade, & Romero Diaz, 2012), gendered racial microaggressions (Lewis & Neville, 2015), sexual orientation microaggressions (Wright & Wegner, 2012), and microaggressions experienced by lesbian, gay, bisexual, and transgender people of color (Balsam, Molina, Beadnell, Simoni, & Walters, 2011). Recent research using microaggression scales indicate that experiences with microaggressions are associated with increases in stress, depression, and alcohol abuse (Blume, Lovato, Thyken, & Denny, 2012; Huynh, 2012; Nadal, 2013; Torres, Driscoll, & Burrow, 2010).

Despite the known impact of microaggressions on mental health among other nondominant groups, only one study has explored experiences of microaggressions among people with disabilities. Keller and Galgay (2010) conducted focus groups and used the data to develop a taxonomy of disability microaggressions. The eight domains they outlined included (a) *denial of identity*, occurring when personal identities or experiences unrelated to disability, or those specific to disability status, are ignored, minimized, or denied (e.g., “I can’t believe you are married”); (b) *denial of privacy*, referring to others’ demands for personal disability-related information (e.g., “What happened to you?” in a crowded elevator); (c) *helplessness*, including others’ attempts to help a person with a disability when they do not need help and attempts to save a person with a disability from the disability; (d) *secondary gain*, referring to others’ expectations for recognition, praise, or enhanced self-worth through interacting with a person with a disability; (e) *spread effect*, occurring when assumptions about various abilities in unrelated areas of functioning are made about a person with a disability (e.g., “Your other senses must be better than mine”); (f) *patronization*, including infantilization or treating the person with a disability like they are a child and, on the other hand, admiration of a person with a disability for completing almost any task (e.g., “Let me do that for you”); (g) *second-class citizenship*, meaning when a person with a disability is denied equality “because they are considered to be bothersome, expensive, and a waste of time, effort, and resources” (Keller & Galgay, 2010, p. 250); and (h) *desexualization*, occurring when aspects of sexuality or sexual identity are ignored or denied (e.g., “I would never date someone who uses a wheelchair”). Two auxiliary domains, *exoticization*, the assigning of a hypersexualized status based on a disability, and *spiritual intervention*, praying over the person with a disability, were less consistent in their data.

The purpose of our study was to develop and validate a scale, the Ableist Microaggressions Scale (AMS), to measure people with disabilities’ experiences with microaggressions. As well as contributing to the growing empirical literature on microaggressions, the development and validation of a disability microaggressions scale can facilitate future research that explores how microaggressions influence mental and physical health disparities

among people with disabilities. Our analyses include a determination and confirmation of the scale's structure via exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), as well as a test of measurement invariance. Convergent validity was tested using correlations with measures of perceived stress and depression, and discriminant validity was assessed using measures of socially desirable responding. As microaggressions are suggested to contribute to minority stress, which has been associated with chronic stress and depression, we hypothesized that perceived stress and depression would be significantly positively correlated with ableist microaggressions. Previous studies have also found significant positive associations between these constructs and other types of microaggressions (see Huynh, 2012, and Torres, et al., 2010). As socially desirable responding has not been found to correlate significantly with participants' responses to a scale of sexual minority microaggressions (Wright & Wegner, 2012), we hypothesized there would be no observed relationship between this construct and ableist microaggressions.

Study 1: Measure Development

Method

Item development. The first and second authors produced an initial pool of 44 items for the preliminary Ableist Microaggressions Scale (pAMS), based largely on examples described in Keller and Galgay's (2010) study and published microaggressions scales. The authors then discussed each developed item to clarify its meaning and examine conceptual overlap with other items. Nineteen items were then removed due to conceptual overlap. For example, "Because I have a disability, people do not speak directly to me about issues that concern me" was removed due to redundancy with "People ignore me, avoid me, or provide poorer service to me because I have a disability," which was retained. Although several microaggression scales use 4- or 5-point Likert-type scales, the authors selected a 6-point Likert-type rating scale, ranging from 0 (*never*) to 5 (*very frequently*) to capture a fuller range of experience.

Expert feedback. A panel of three psychologists who were not involved with other aspects of this study, were selected based on their experience conducting research related to physical disability and/or microaggressions. They responded to an online questionnaire, developed by the first author, about the wording and content of the instructions and initial items. For example, for each item, the panel members were asked, "How clear is the wording of this

item?” and “To what extent do you think this item represents its domain?” They were also invited to provide suggestions for the removal of items and addition of new items. Twelve of the 25 items were revised based on their feedback, and one item was added. For example, regarding the item “People assume that I have a higher degree of functioning or special talent in some area because I have a disability,” two respondents suggested the phrase “higher degree of functioning” was unclear. The item was revised to read “People assume that I have a special gift or talent because I have a disability.” In addition, two items were written in two different forms for the cognitive interviews in order to determine the most appropriate wording of the item. The item pool was expanded to 28 items for use in the cognitive interviews.

Cognitive interviews. Cognitive interviewing refers to “administering draft survey questions while collecting additional verbal information about the survey responses, which is used to evaluate the quality of the response or to help determine whether the question is generating the information that its author intends” (Beatty & Willis, 2007, p. 288). This process often includes elaborating on how answers are decided, explaining interpretation of items, and describing difficulty in responding (Beatty & Willis, 2007). The first author conducted two 90-min individual cognitive interviews via Skype. Participants were recruited through social and professional networks of the investigators, and were not intended to be representative of the wider disability population. The participants were one White male and one White female in their mid-20s and mid-30s living in California. One participant endorsed attending “some college” and one held a graduate degree. They both reported having a neuromuscular disease, with one participant stating the disability is “readily apparent” to others due to use of a wheelchair, and the other reporting the disability is “sometimes apparent” to others.

Upon receiving each item via Skype message, participants thought out loud about the item and provided a numerical response corresponding to their lifetime experience, ranging from 0 (*never*) to 5 (*very frequently*). Participants were then asked one to three probe(s) per item to reveal potential cognitive errors in understanding the item, such as “Can you repeat the statement in your own words?” (Willis, 2005). In addition, emergent or unscripted probes were occasionally asked based on participant observation or responses, such as “I noticed that you hesitated. Tell me what you were thinking” (Willis, 2005). At the conclusion of the interview, participants provided general feedback about the items.

During each interview, the first author took detailed notes as participants thought out loud while responding to questions, providing a numerical response, and answering probes. The investigator used these notes, along

with the audio recordings, to make revisions that would improve item comprehension (Willis, 2005). Five items were revised, four new items were added, one item was divided into two items, the alternative forms of two items were removed, three items became reverse scored, and the instructions were revised. For example, participants were asked to respond to the item “I am aware of few or no role models with disabilities in areas of life that are important to me.” Although the participants understood the item’s meaning, both were unclear how to provide a scaled response score. The item was revised to be reverse scored, “I am exposed to role models with disabilities in areas of life that are important to me,” in order to improve clarity of responding. Based on the interviews, the item pool was expanded to 31 items.

Pilot study. The complete survey was piloted via SurveyMonkey.com with seven volunteer respondents with physical disabilities. Participants were recruited using targeted social media. Upon completion of the online survey, participants provided feedback on the wording and content of the 31 items, experiences that were not captured by the items, and on the survey overall. As a result, *not applicable* (N/A) was added as a response option for three of the items, the wording of two items was revised, one new item was added, and the demographic questions were revised. The final version of the pAMS to include in the second study contained 32 items.

Study 2: Exploratory Factor Analysis

Method

Participants. The survey was accessed by 1,856 people. Participants were excluded from both studies if they did not complete all demographic and study items ($n = 293$), or did not report having a disability ($n = 171$) based on the Americans with Disabilities Act Amendments Act of 2008 definition as “a physical or mental impairment that substantially limits one or more major life activities of such individual, a record of such an impairment, or being regarded as having such an impairment” (Section 4[a][1]). After these initial exclusion criteria, 1,392 participants were included in the total sample. As recommended by Fabrigar, Wegener, MacCallum, and Strahan (1999), participants were randomly split into two groups, approximately 40% for inclusion in Study 2 ($N = 559$) and approximately 60% in Study 3 ($N = 833$), using SPSS (version 20). A 40–60% split was used to allow for a larger sample size to confirm the factor structure in the CFA (Dowdy et al., 2011). See Table 1 for sample demographics. As participants could select more than one category for race or ethnicity (as reflected in Table 1), we calculated the percentage who selected

Table 1. Participant Demographics Across Study 2 and Study 3

Variable	Study 2 (N = 559)	Study 3 (N = 833)
Age		
Range	18–80	18–72
Mean	35.47	34.30
Standard Deviation	13.43	13.01
Gender identity ^a		
Female	57.1%	52.7%
Male	41.3%	45.9%
Genderqueer	1.6%	1.2%
Transgender FtM	0.0%	0.7%
Transgender MtF	0.4%	0.1%
Other	0.0%	0.4%
Race and ethnicity ^a		
White or European American	82.5%	79.8%
Black or African American	8.6%	7.9%
Asian, Asian American, or Pacific Islander	6.3%	5.5%
Chicano/a, Hispanic, or Latino/a	4.8%	6.6%
American Indian, Native American, or Alaska Native	3.6%	2.9%
Biracial or multiracial	2.7%	3%
Middle Eastern	0.5%	0.5%
Disability type ^a		
Medical condition or chronic illness	42.0%	40.0%
Mobility or physical impairment	34.0%	34.0%
Blind, low vision, or vision impairment	20.2%	23.5%
Deaf, hard-of-hearing, or hearing impairment	18.1%	17.3%
Brain injury	5.4%	4.9%
Other	5.0%	4.1%
Disability onset		
Acquired	80.7%	78.2%
Congenital	15.2%	15.7%
Both	4.1%	6.1%
Disability visibility		
Visible	14.1%	16.2%
Semivisible	36.9%	37.2%
Invisible	45.3%	42.5%
Unsure	2.0%	3.1%
Other	1.8%	1.0%
Disability severity		
Mild	14.8%	13.6%
Moderate	51.3%	52.2%
Severe	26.1%	28.0%
Very severe	7.7%	6.2%
Sexual orientation		
Heterosexual	84.1%	83.3%
Sexual minority (identity other than heterosexual)	15.9%	16.7%

(continued)

Table 1. (continued)

Variable	Study 2 (N = 559)	Study 3 (N = 833)
U.S. State		
Total represented	47 & District of Columbia	48 & District of Columbia
Living environment		
Urban	14.1%	18.7%
Medium-sized city or town	14.3%	13.1%
Small city or town	28.1%	22.2%
Suburb	24.5%	30.0%
Rural	19.0%	16.0%
Employment status		
Unemployed	34.2%	29.7%
Regular full-time	28.3%	32.5%
Regular part-time	13.8%	13.1%
Temporary	8.9%	9.8%
Student	9.7%	9.8%
Contingent or on-call	5.2%	5.0%
Employment satisfaction		
Not satisfied	44.0%	40.9%
"Sort of" satisfied	26.1%	27.1%
Satisfied	29.9%	31.9%
Annual individual income		
Below \$11,000	32.7%	32.7%
\$11,000-19,999	19.9%	18.7%
\$20,000-39,999	25.2%	22.4%
\$40,000+	22.1%	26.1%
Socioeconomic status		
Poor	17.4%	15.8%
Working class	27.2%	23.3%
Lower middle class	18.2%	19.0%
Middle class	30.2%	32.7%
Upper middle class	6.1%	8.3%
Upper class	0.5%	0.6%
Other	0.4%	0.4%
Education		
Some high school	1.3%	1.1%
High school diploma or equivalent	10.7%	11.9%
Some college	36.0%	34.8%
Associate's degree	12.5%	11.9%
Bachelor's degree	26.3%	28.7%
Graduate or professional training	4.8%	3.2%
Graduate or professional degree	8.4%	8.4%

Note. FtM = female to male; MtF = male to female.

*Participants had the option to indicate more than one response, therefore some totals are greater than 100%.

only “White or European American” without selecting any additional response category, which was 76% for Study 2 and 75.6% for Study 3.

Procedure. The survey was accessed by 1,856 people via Amazon Mechanical Turk (MTurk), a survey recruitment website where anonymous participants complete tasks for a small incentive (Buhrmester, Kwang, & Gosling, 2011). In order to access the study, participants were required to be 18 years of age or older, reside in the United States (U.S.), and have at least 95% of their previous tasks approved upon quality review (see Paolacci, Chandler, & Ipeirotis, 2010). MTurk users accessed the survey via the designated URL link on the task page, where they were provided a statement explaining the purpose of the study, the participation criteria, potential risks and benefits to participation, and informed consent. Participants were informed that they would be completing measures on disability-related stress and well-being. The term “microaggression” was not used in participant recruitment or data collection, as it may have been suggestive to participants (Wright & Wegner, 2012). The keywords listed for this study were “survey,” “research,” “demographics,” “psychology,” “disability,” “illness,” “impairment,” “health,” “stress,” “social,” and “sexual orientation” (this term was included due to simultaneous recruitment for another study). The survey was Section 508 compliant and accessible to participants with disabilities. Participants received a \$0.50 incentive through MTurk’s payment system.

Measures

Demographic questionnaire. Participants completed a demographic and background questionnaire including age, gender identity, race and ethnicity, sexual orientation identity, geographic region, employment, socioeconomic status, and education. Although there are commonalities among people with physical disabilities, there is also a great deal of diversity in this community. Therefore, we collected information on the type of disability, levels of functioning, and intersecting identities (Olkin, 2004), such as type of disability, age of onset, visibility, and severity.

pAMS. The pAMS included 32 items. Participants rated their lifetime experiences with microaggressions on a 6-point Likert-type scale, ranging from 0 (*never*) to 5 (*very frequently*). Three positively worded items were reverse scored for all analyses. Three items included an option to respond “N/A,” as these items may not be relevant for people with invisible disabilities. Due to their nonrandom nature, N/A responses were scored as 0 rather than as missing data. Keller and Galgay’s (2010) eight domains were reflected in pAMS

items as follows: Denial of Identity (8 items), Denial of Privacy (2 items), Helplessness (3 items), Secondary Gain (2 items), Spread Effect (3 items), Patronization (4 items), Second-Class Citizenship (7 items), and Desexualization (3 items). Cronbach's alpha for the 32-item pAMS was .92.

Results

Preliminary analyses. Study 2 included no missing data because participants were required to respond to every question. Descriptive statistics for the 32 measured variables were examined, including means, standard deviations, and intercorrelations. To test assumptions of univariate normality, visual inspections of histograms, q-q plots, box plots, and bivariate scatterplots were conducted for all variables. No variables demonstrated bimodal distributions or severe non-normality. Statistics for skewness and kurtosis were also within the normal range (skewness < 2, kurtosis < 7) for all variables (Fabrigar et al., 1999). One variable contained outliers, but responses were determined to be accurate and were not removed. To account for non-normality among variables included in the EFA, maximum likelihood estimation (MLR) with robust standard errors was used, as this method is robust to non-normality.

Prior to factor analyses, the complete dataset ($N = 1,856$) was used to compare respondents who discontinued participation in the study after completing the demographics ($n = 105$) to a similarly sized random sample of participants who completed the entire study ($n = 104$). Current age between the two groups was not found to differ significantly, $t(207) = -1.29, p = .20$. Chi-square tests yielded no significant differences between the two groups based on race and ethnicity, $\chi^2(1, N = 209) = 0.98, p = .32$, sexual orientation identity, $\chi^2(1, N = 209) = 1.06, p = .30$, or gender identity, $\chi^2(1, N = 209) = 0.01, p = .94$. There was a significant difference based on severity of disability, $\chi^2(3, N = 209) = 17.08, p = .001$, such that participants who reported their disability impacted their life mildly or moderately were less likely to complete the study than those who reported impact as severe or very severe. For two disability-related variables, acquired versus congenital status and visibility status, chi-square tests could not be conducted due to sparse cells. Proportions were, therefore, hand calculated using pooled sample proportions, standard errors, and a z-score test statistic. No significant differences were found between participants who completed the entire survey and those who did not in terms of disability status as acquired, $p = .99$, congenital, $p = .66$, and both, $p = .53$. In addition, no significant differences were found between the two groups in terms of disability status as visible, $p = .66$, semi-visible, $p = .99$, invisible, $p = .99$, and unsure, $p = .51$. Therefore, it appears

that, other than severity of disability, demographic variables were similar among participants who did and did not complete the survey.

Exploratory factor analysis. Factor extraction was performed using MLR, which allowed for statistical significance testing of factor loadings, correlations among factors, and confidence intervals (Fabrigar et al., 1999). Geomin oblique rotation was used to establish the best solution, and rotated structure coefficients were reviewed to identify problematic items. Oblique rotation has been recommended when factors are likely to be correlated (Worthington & Whittaker, 2006), as well as for its greater accuracy, simpler solutions, and ability to provide correlation estimates among factors (Fabrigar et al., 1999).

Multiple criteria were used to determine the appropriate number of factors, including parallel analysis, Kaiser's eigenvalues, Catell's scree plot, chi-square test of model fit, comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean square residual fit index (SRMR; Fabrigar et al., 1999). Regarding parallel analysis, the point at which the plot of eigenvalues and scree plot intersect is identified and the model just before the intersection is studied, along with neighboring models (e.g., model with 2-, 3-, and 4-factors; Brown, 2015; Wei, Alvarez, Ku, Russell, & Bonett, 2010). Criteria for Kaiser's eigenvalues, Catell's scree plot, chi-square test of model fit, CFI, RMSEA, and SRMR were used as described by Fabrigar et al. (1999) and Kahn (2006). At all steps, multiple model solutions were examined to ensure the best fit was adopted, statistically and theoretically.

Mplus (Version 7.4; Muthén & Muthén, 1998-2016) statistical software was used for all analyses. Based on 1,000 random data sets, parallel analysis indicated a 6-factor model with an eigenvalue of 1.09, which was just below the average eigenvalue of 1.26 for the sixth factor from the random data sets. Therefore, the 4-, 5-, 6-, and 7-factor solutions were investigated first. Each of these solutions, however, included at least one factor with only one or two items, which was too poorly defined to be retained (Kahn, 2006). Upon removal of two items that loaded on one factor in all models, the parallel analysis indicated a 5-factor model with an eigenvalue of 1.09 falling below the average eigenvalue of 1.27, which was explored along with neighboring solutions. The geomin-rotated loadings of the 4-factor model were found to be the most interpretable and the model demonstrated adequate fit indices. Factor loadings at .30 or larger were used to indicate that an item loads to a particular factor, with no cross-loadings on other factors above .30 (Brown, 2015; Tabachnick & Fidell, 2007). Eight items were removed based on these criteria, and one item was removed due to poor conceptual alignment. Modification indices were requested and expected parameter change values

were reviewed. Compared to other modification indices presented, items 4 and 5 demonstrated a high expected parameter change value of 53.89, which indicates some redundancy in the items. Instead of including the residual covariance of these two items, item 4 was removed (for a list of the 12 items that were removed, see Supplemental Table 1A, available online at <http://journals.sagepub.com/doi/suppl/10.1177/0011000017715317>). This resulted in a 20-item EFA with a 4-factor geomin rotated solution with high loadings on one factor only and interpretable factors. Fit indices for the 4-factor model were as follows: $\chi^2 (164, N = 559) = 271.90, p < .001$, RMSEA = .05 [.04 .06], CFI = .96, and SRMR = .02, demonstrating good model fit.

The four factors included (a) Helplessness (Factor 1; five items), in which items describe individuals with disabilities being treated as if they are incapable, useless, dependent, or broken, and imply they are unable to perform any activity without assistance; (b) Minimization (Factor 2; three items), in which items suggest respondents are overstating their impairment or needs, and imply individuals with a disability could be able-bodied if they wanted to be or that they are actually able-bodied; (c) Denial of Personhood (Factor 3; five items), in which items reflect being treated with the assumption that a physical disability indicates decreased mental capacity and therefore, being reduced to one's physicality; and (d) Otherization (Factor 4; seven items), in which items measure being treated as abnormal, an oddity, or nonhuman, and imply people with disabilities are or should be outside the natural order. Keller and Galgay's (2010) themes of helplessness, patronization, secondary gain, and denial of identity were represented by items within the Helplessness factor; denial of identity was represented in the Minimization factor; infantilization, spread effect, and denial of identity were represented by items within the Denial of Personhood factor; and desexualization, second-class citizen, and spread effect were represented in the Otherization factor. Factor loadings for each factor are shown in Table 2. Factor determinacy estimates for the four factors were determined using Mplus and were all above 0.80 (Gorsuch, 1983 as cited in Brown, 2015) for the four-factor model (Helplessness = .92, Minimization = .84, Denial of Personhood = .96, Otherization = .94). The 4-factor solution accounted for 61.33% of the total variance in the items before rotation. As expected, factor correlations indicated significant positive relationships between most factors, as follows: Helplessness with Denial of Personhood, $r = .52, p < .05$; Helplessness with Otherization, $r = .52, p < .05$; Helplessness with Minimization, $r = .11, p > .05$ (nonsignificant); Denial of Personhood with Otherization, $r = .68, p < .05$; Denial of Personhood with Minimization, $r = .31, p < .05$; and Otherization with Minimization, $r = .31, p < .05$. Internal consistency reliability for three of the four factors was within adequate range, with Cronbach's alpha at .83 for Helplessness, .90 for Denial of Personhood,

Table 2. Pattern Matrix of Geomin Rotated 4-Factor Solution

Item	F1	F2	F3	F4	M	SD	h ²
5. People feel they need to do something to help me because I have a disability.	.74*	-.03	.11	-.12	1.95	1.54	.58
17. People express admiration for me or describe me as inspirational simply because I live with a disability.	.75*	.00	-.07	.06	1.74	1.56	.58
18. People express pity for me because I have a disability.	.65*	.20*	-.03	.16	2.11	1.52	.49
3. People do not expect me to have a job or volunteer activities because I have a disability.	.58*	-.07	.20*	-.01	1.61	1.61	.35
10. People offer me unsolicited, unwanted, or unneeded help because I have a disability.	.62*	.15*	.03	.02	2.13	1.64	.41
32. People are unwilling to accept that I have a disability because I appear able-bodied. ^a	.01	.71*	.00	-.07	3.28	1.85	.51
1. People minimize my disability or suggest that it could be worse.	.05	.52*	.07	.05	3.02	1.54	.28
28. People act as if accommodations for my disability are unnecessary. ^a	-.08	.47*	.15	.20	2.09	1.82	.29
13. People don't see me as a whole person because I have a disability.	.00	.02	.91*	-.01	1.63	1.56	.83
12. People act as if I am nothing more than my disability.	.03	.04	.85*	.01	1.43	1.51	.73
14. People speak to me as if I am a child or do not take me seriously because I have a disability.	.01	.00	.64*	.19	1.48	1.60	.44
8. People assume I have low intelligence because I have a disability.	.04	-.01	.56*	.16	1.44	1.64	.34
16. Because I have a disability, people attempt to make decisions for me that I could make myself.	.25*	.01	.42*	.21*	1.71	1.55	.28
24. People think I should not date or pursue sexual relationships because I have a disability.	-.06	-.18	.04	.89*	0.89	1.34	.83
26. People indicate that they would not date a person with a disability.	-.02	.07	.04	.73*	1.35	1.46	.53
27. People suggest that I cannot or should not have children because I have a disability.	.11	.14	-.15*	.70*	1.06	1.50	.54
22. People stare at me because I have a disability. ^a	.18*	-.17*	.03	.46*	1.39	1.63	.27
31. Because I have a disability, people seem surprised to see me outside my home.	.25*	.05	.13	.41*	1.15	1.42	.25
9. Because I have a disability, people assume I have an extraordinary gift or talent.	.26*	-.12	.07	.40*	0.91	1.29	.24
19. People suggest that living with a disability would not be a worthwhile existence.	.14*	.16*	.14	.32*	1.42	1.52	.17

Note. Pattern coefficient factor loadings > [.30] are in bold type. F = factor; F1 = Helplessness; F2 = Minimization; F3 = Denial of Personhood; F4 = Otherization; h² = communality estimate.

^aItem includes *not applicable* response option.

**p* < .05.

and .84 for Otherization. Minimization demonstrated relatively weak internal consistency reliability, however, with a Cronbach's alpha of .65. Cronbach's alpha for the 20-item AMS was .91.

Study 3: Confirmatory Factor Analysis, Reliability, and Validity

Method

Participants. Participants in this study ($n = 833$) met the same inclusion criteria and were recruited in the same manner as Study 2. See Table 1 for sample demographics. As participants were randomly selected for inclusion in Study 2 or 3, sample demographics are comparable.

Measures

Demographic questionnaire. Participants completed the same demographic and background questionnaire described in Study 2. Responses on the demographic questionnaire were used to describe the sample, explore within-group differences related to AMS scores, and position the findings in the literature.

AMS. Participants rated their lifetime experiences with microaggressions on a 6-point Likert-type scale, ranging from 0 (*never*) to 5 (*very frequently*). Scores are averaged across the 20 items, with scores ranging from 0 to 5. Examples of items include “People minimize my disability or suggest that it could be worse” and “People don’t see me as a whole person because I have a disability.” Cronbach’s alpha for the 20-item version of the AMS included in Study 3 was .92.

Perceived Stress Scale-4. The Perceived Stress Scale-4 (PSS-4; Cohen, Kamarck, & Mermelstein, 1983) is a 4-item self-report scale measuring the extent to which participants judge their lives to be unpredictable, uncontrollable, and overwhelming in the previous month. The PSS-4 is an abbreviated version of the widely used 14-item and 10-item versions of the Perceived Stress Scale (Cohen et al., 1983), which show a positive relationship with racial microaggressions (Torres et al., 2010). The four items selected were those that correlated most highly with the 14-item version. Responses are indicated on a 5-point Likert-type scale with 0 = *never*, 1 = *almost never*, 2 = *sometimes*, 3 = *fairly often*, and 4 = *very often*. Two of the items are reverse scored and then scores are summed across the four items. Higher scores reflect greater perceived stress. Cohen et al. (1983) reported an internal consistency reliability of .72 and 2-month test-retest reliability of .55. A study examining microaggressions against LGBT people of color reported an internal consistency reliability of .84 for the PSS-4 in their sample (Balsam et al.,

2011). Many studies have reported sound psychometric properties in the PSS-4 across a variety of populations (Goldstein, Dawson, Smith, & Grant, 2012; Karam et al., 2012; Lesage, Berjot, & Deschamps, 2012; Mitchell, Crane, & Kim, 2008). In the present study, the PSS-4 was used to examine convergent validity with the AMS. Cronbach's alpha for the PSS-4 was .78 for this sample.

Center for Epidemiologic Studies Depression Scale. The Center for Epidemiologic Studies Depression Scale (CESD-10; Andresen, Malmgren, Carter, & Patrick, 1994) is a 10-item self-report measure of depressive symptoms. The CESD-10 is a short-form version of the 20-item CES-D (Radloff, 1977), which was positively correlated with ethnic microaggressions in one study (Huynh, 2012). Participants are instructed to consider their experiences in the previous week. Responses are indicated on a 4-point Likert-type scale with 0 = *rarely or none of the time (less than one day)*, 1 = *some or a little of the time (1–2 days)*, 2 = *occasionally or a moderate amount of the time (3–4 days)*, and 3 = *most or all of the time (5–7 days)*. Two items are reverse scored and the scores are then summed across the 10 items, with total scores ranging from 0 to 30. Higher total scores indicate greater depressive symptoms (Huynh, 2012). Andresen et al. (1994) reported an overall test-retest item correlation of .71, and convergent validity with measures of stress ($r = .43, p < .005$), pain ($r = .30, p < .005$), and decreased health status ($r = .37, p < .005$). A study examining microaggressions against LGBT people of color reported an internal consistency reliability of .91 for the CESD-10 in their sample (Balsam et al., 2011). Sound psychometric properties have also been reported among older adults with chronic conditions (Zauszniewski, Morris, Preechawong, & Chang, 2004) and people with spinal cord injuries (Miller, 2009). In the present study, the CESD-10 was used to examine convergent validity with the AMS. Cronbach's alpha for the CESD-10 was .90 for this sample.

Socially Desirable Response Set Five-Item Survey. The Socially Desirable Response Set Five-Item Survey (SDRS-5; Hays, Hayashi, & Stewart, 1989) is a 5-item self-report measure of participants' tendency to provide socially desirable responses. Based on the Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1960), the SDRS-5 is a shorter alternative to lengthier social desirability measures. The abbreviated version was chosen for the current study to decrease respondent burden. Responses are indicated on a 5-point Likert-type scale, ranging from 1 (*definitely true*) to 5 (*definitely false*). After reverse scoring, any responses indicating an extreme score of 1 (*definitely true*) are scored 1 whereas any other response is scored 0. The scores are summed and higher scores reflect greater socially desirable

Table 3. Intercorrelations Among AMS Factors, AMS Score, and Study 3 Measures

Measure	F1	F2	F3	F4	AMS	PSS-4	CESD-10	SDRS-5
F1	—							
F2	.27**	—						
F3	.66**	.36**	—					
F4	.68**	.30**	.76**	—				
AMS	.84**	.52**	.89**	.90**	—			
PSS-4	.11**	.28**	.22**	.21**	.24**	—		
CESD-10	.14**	.34**	.25**	.24**	.29**	.75**	—	
SDRS-5	-.08*	-.09**	-.09*	-.15**	-.13**	-.16**	-.13**	—
<i>M</i>	1.96	2.76	1.51	1.17	1.69	1.93	1.26	.21
<i>SD</i>	1.18	1.34	1.28	1.03	0.96	1.08	.99	.40
Scale Range	0–5	0–5	0–5	0–5	0–5	0–4	0–3	0–1

Note. AMS = Ableist Microaggressions Scale; F = factor; F1 = Helplessness; F2 = Minimization; F3 = Denial of Personhood; F4 = Otherization; PSS-4 = Perceived Stress Scale-4; CESD-10 = Center for Epidemiologic Studies Depression Scale; SDRS-5 = Socially Desirable Response Set Five-Item Survey.

* $p < .05$. ** $p < .01$.

responding. Hays et al. (1989) reported internal consistency reliability ranging from .66 to .68, and 1-month test-retest reliability of .75. Heppner, Humphrey, Hillenbrand-Gunn, and DeBord (1995) reported a .79 correlation between the SDRS-5 and the commonly used 33-item Marlowe-Crowne scale. In the present study, the SDRS-5 was used to examine discriminant validity with the AMS; it was expected that socially desirable responding would not be significantly associated with ableist microaggressions. Cronbach's alpha for the SDRS-5 was .63 for this sample.

Results

Preliminary analyses. Study 3 included no missing data because participants were required to respond to every question. The same procedures outlined in Study 2 were followed regarding testing for normality. Although some items on the AMS, PSS-4, and CESD-10 demonstrated mild to moderate skewness and kurtosis, no variables demonstrated bimodal distributions or severe non-normality. The distribution of the SDRS-5 total score was normal. One item on the AMS and one item on the PSS-4 contained outliers, but responses were not removed, as they were determined to be accurate and part of the population. Intercorrelations among mean scores for AMS subscales, AMS total mean scores, validity measures, and descriptive statistics for these variables were also examined (see Table 3).

Table 4. Fit Indices for Alternative Models of the 20-item AMS

CFA Models	Scaled χ^2	$\Delta \chi^2$	df	RMSEA [95% CI]	CFI	SRMR	AIC	ECVI	
1-factor reduced	1350.59		170	.09	[.09, .10]	.79	.08	56097.46	1.77
4-factor reduced, correlated factors	810.21	540.38*†	164	.07	[.06, .07]	.89	.07	55403.98	1.13
Second-order, 4-factor reduced	825.07	—	166	.07	[.06, .07]	.88	.07	55417.62	1.14

Note. AMS = Ableist Microaggressions Scale; CFA = confirmatory factor analysis; RMSEA = root-mean-square error of approximation; CFI = comparative fit index; SRMR = standardized root-mean-square residual; AIC = Akaike information criterion; ECVI = expected cross-validation index. Scaled χ^2 are Satorra-Bentler adjusted chi-square values.

†Chi-Square differences were tested using the Satorra-Bentler scaled Chi-Square and were adjusted using the scaling correction factor for MLR. Non-nested models were compared using AIC and ECVI.

* $p < .001$.

CFA. CFA was used to confirm the 20-item, 4-factor structure (i.e., Helplessness, Minimization, Denial of Personhood, and Otherization) of the AMS identified by the EFA. The 20 AMS items were input into Mplus by latent factor, with reference variables input first for each factor. To set the scale of the CFA, unit-loading identification was used to fix the reference variable for each latent factor to a factor loading of 1, where variables with the highest loading were used as the reference variable. As in the EFA, MLR was used in CFA analyses because this method is robust to non-normality. Fit indices were examined to verify overall model fit of the 4-factor solution and were as follows: $\chi^2(164, N = 833) = 810.21, p < .001$, RMSEA = .07 [.06 .07], CFI = .89, and SRMR = .07, demonstrating fair model fit. The standardized factor loadings were all greater than .49, and all loadings were statistically significant. Alternative models were considered and compared. To compare the two nested models (i.e., the 1-factor model and 4-factor correlated model), we conducted a chi-square difference test. For the non-nested models (i.e., the 4-factor correlated model and the second-order model), we cannot use the chi-square difference test so we compared values of Akaike information criterion (AIC; Akaike, 1977) and expected cross-validation index (ECVI; Browne & Cudeck, 1993), where models with smaller values are favored. Based on these results, we concluded that the 4-factor correlated model was the best model (see Table 4). As the first-order and second-order model fit comparably with respect to fit indices, use of either AMS subscale or total scores may be appropriate. Figure 1 illustrates the final 4-factor correlated CFA model.

Measurement invariance across disability severity. As a significant difference was found for study completion based on disability severity, a multiple group CFA was conducted to test for measurement invariance across severity

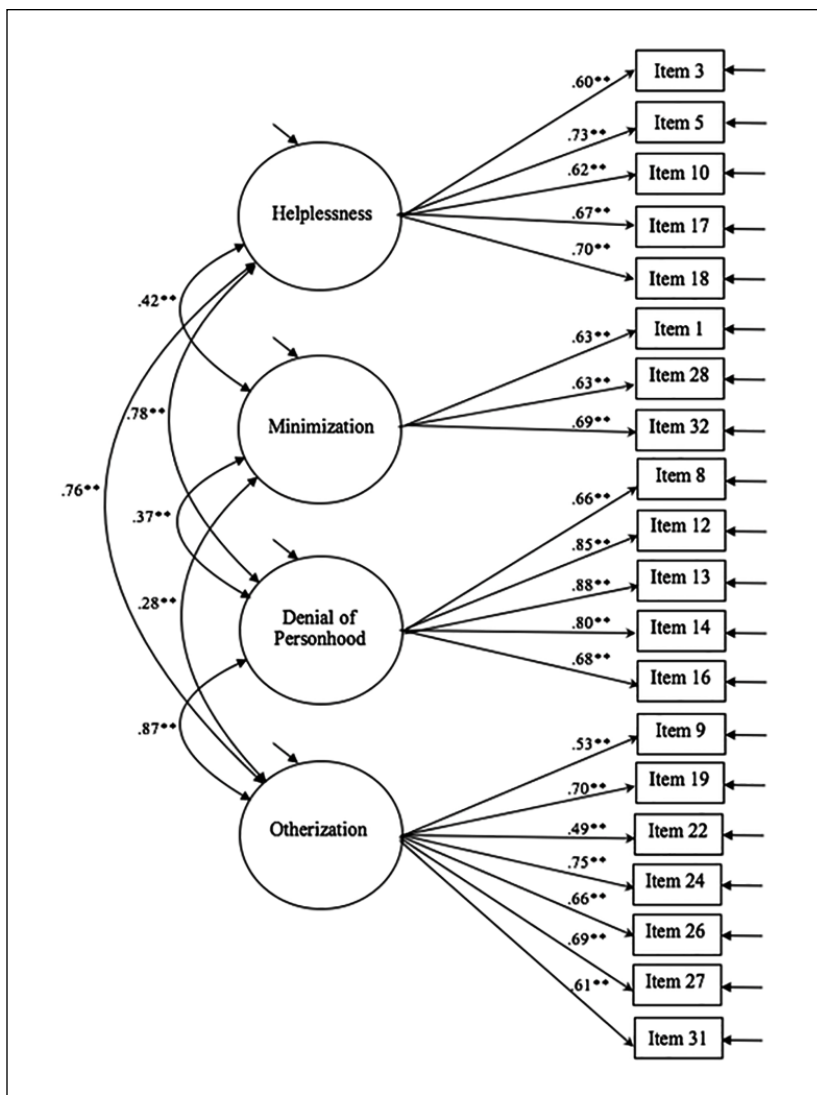


Figure 1. Standardized parameter estimates of confirmatory factor analysis, demonstrating a 4-factor solution for ableist microaggressions items.

** $p < .001$.

groups. Individuals who reported the impact of their physical disability as 1 = mild or 2 = moderate were grouped in the mild group, and individuals who

Table 5. Measurement Invariance Across Disability Severity

Model	Scaled χ^2	df	Model Comparison	$\Delta \chi^2$	Δdf	CFI	ΔCFI	RMSEA	SRMR
Mild ($N = 548$)	574.87	164				.89		.07	.07
Severe ($N = 285$)	547.87	164				.86		.07	.07
Configural (M1)	882.06	328				.91		.06	.06
Metric (M2)	915.71	348	M2 - M1	21.35 [†]	20	.91	-.002	.06	.06
Scalar (M3)	1005.41	368	M3 - M2	78.83 ^{†*}	20	.90	-.01	.06	.08

Note. CFI = comparative fit index; RMSEA = root-mean-square error of approximation;

SRMR = standardized root-mean-square residual; M1 = configural model (i.e., all parameters freely estimated across groups); M2 = metric model (i.e., loadings held equal across groups); M3 = scalar model (i.e., loadings and intercepts held equal across groups).

[†]Chi-Square differences were tested using the Satorra-Bentler scaled Chi-Square and were adjusted using the scaling correction factor for MLR.

* $p < .001$.

responded 3 = *severe* or 4 = *very severe* were grouped in the severe group. We began by fitting the 4-factor CFA in each model, which yielded adequate fit in each group as seen in Table 5. We then fit a series of models, constraining equality across the two groups and conducted a series of nested model tests using the MLR scaled chi-square difference test, as well as considered the change in the fit index CFI (denoted ΔCFI). A nonsignificant change in chi-square ($\Delta \chi^2$) when comparing the nested model indicates there is not a significant decrease in model fit by imposing the equality constraint (Wei, Wang, & Ku, 2012). As the $\Delta \chi^2$ has been shown to be oversensitive in larger samples (Meade, Johnson, & Braddy, 2008) and the sample size in the current study is large ($N = 833$), we also used the ΔCFI as way of assessing measurement invariance, where a change smaller than or equal to $-.01$ indicates that invariance should not be rejected (Cheung & Rensvold, 2002). Although Meade et al. (2008) suggested a cutoff of $.002$ for ΔCFI , with our large sample and the values presented in the power figure (see Figure 6 in cited article), we feel confident with the recommendation of ΔCFI of $-.01$.

We first fit the two severity groups together in one multiple group CFA model, which we labeled the configural model (M1) to be used as a baseline model as no parameters are constrained to be equal in this model. Next, we constrained the factor loadings to be equal across severity groups and compared model fit. The results indicated that constraining the loadings to be equal across severity groups in the metric model (M2) did not result in significant decrease in fit, $\Delta \chi^2 (20, N = 833) = 21.35, p = .37$. Next, we constrained the intercepts and loadings to be equal across severity groups in the scalar model (M3). Results indicated that this did significantly decrease model fit based on the chi-square difference test, $\Delta \chi^2 (20, N = 833) = 78.83, p < .01$. However, the ΔCFI for this model comparison is $-.010$, which is at

the threshold of indicating poor fit. Although the SRMR for the scalar model ($\text{SRMR} = .077$) is higher than those of the configural and metric models, it is still lower than the traditional cutoffs for CFA models, which is at or below .08 (Hu & Bentler, 1999). Taken together, the results of the measurement invariance testing indicate we can confidently assume measurement invariance of the AMS across severity groups.

Convergent and discriminant validity. Correlations between the AMS and the PSS-4 and between the AMS and the CESD-10 indicated a small, positive relationship between microaggressions and stress ($r = .24$) and between microaggressions and depression ($r = .29$), which offered some support for convergent validity. Discriminant validity was measured by testing the correlation between the AMS and the SDRS-5, and a small, negative correlation ($r = -.13$) was found. Although this small correlation was statistically significant, this was likely due to the large sample size.

Reliability. Cronbach's alpha for the AMS was .92. Internal consistency reliability for three of the four factors was within adequate range, with Cronbach's alpha at .85 for Helplessness, .90 for Denial of Personhood, and .84 for Otherization. Minimization demonstrated weak internal consistency reliability, with a Cronbach's alpha of .65. Item mean scores ranged from 0.90 to 3.24 ($SD = 1.25\text{--}1.87$), and the mean across all items was 1.70. Interitem correlations ranged from $-.07$ to $.80$. Item-total correlations ranged from $.21$ to $.77$, demonstrating strong correlations between individual items with the overall scale, with the exception of Items 1 and 32. Intercorrelations of factors appear in the path diagram shown in Figure 1, indicating a high correlation between Denial of Personhood and Otherization.

Supplemental analysis. One-way between-subjects Analyses of Variance were performed to explore disability visibility, disability type, disability severity, and age of impairment onset. Regarding visibility variables, a significant difference on AMS scores was found, $F(4, 828) = 10.41, p < .001, \eta^2 = .05$. A post-hoc Tukey HSD test ($\alpha = .05$) indicated that the visible group ($M = 2.09, SD = 0.99$) reported significantly higher levels of ableist microaggressions than the semi-visible ($M = 1.74, SD = 0.93; d = .37$) and invisible groups ($M = 1.49, SD = 0.93; d = .63$). Participants who were unsure of or indicated "other" regarding the visibility of their disability demonstrated no significant differences with any groups. A significant difference on AMS scores by type of disability was demonstrated, $F(5, 672) = 2.85, p < .05, \eta^2 = .02$; however, a post-hoc Tukey HSD test ($\alpha = .05$) yielded no significant differences. A significant difference on AMS scores was found regarding severity, $F(3, 829) = 24.23, p < .001, \eta^2 = .08$. A

post-hoc Tukey HSD test ($\alpha = .05$) revealed significant differences among all groups ($p < .001$), except between the severe group ($M = 2.01$, $SD = 0.94$) and the very severe group ($M = 2.11$, $SD = 0.93$; $p = .91$). As expected, the mild severity group reported significantly lower levels of ableist microaggressions ($M = 1.20$, $SD = 0.93$) than all other groups (moderate group, $d = -.43$; severe group, $d = -.87$; very severe group, $d = -.98$). The moderate severity group reported significantly lower levels of ableist microaggressions ($M = 1.60$, $SD = .92$) than the severe ($d = -.44$) and very severe groups ($d = -.55$). Finally, a significant violation on the Levene test, $F(3, 829) = 3.47$, $p < .05$, prompted the use of a nonparametric Kruskal-Wallis test regarding age of impairment onset, specifically comparing participants who reported an onset age of 0-10 ($n = 251$), 11-20 ($n = 228$), 21-40 ($n = 240$), 40 and above ($n = 114$), which revealed significant differences, $H(3) = 8.81$, $p = .03$. Post-hoc pairwise comparisons using the Dunn-Bonferroni method indicated that the youngest age of onset group reported significantly higher levels of ableist microaggressions ($M = 1.76$, $SD = 1.00$) than the oldest age of onset group ($M = 1.47$, $SD = 0.90$; $d = .31$).

Independent samples t -tests were then performed to assess whether AMS scores differed based on acquired or congenital disability status, race, and gender. Mean AMS scores between acquired and congenital disability status differed significantly, $t(780) = -2.29$, $p = .02$, with higher scores reported by participants with a congenital disability ($M = 1.85$, $SD = 1.03$) than an acquired disability ($M = 1.64$, $SD = .94$; $d = .21$). With regards to race, equal variances were not assumed due to a significant violation on the Levene test, $F(831, 301.731) = 14.462$, $p < .05$, and no significant difference between White participants and participants who identified as any other race or ethnicity was found, $t(301.73) = -0.88$, $p = .38$. There was also no significant difference between female and male participants, $t(819) = 1.96$, $p = .05$.

Discussion

The present study substantiates that people with physical disabilities do in fact report experiences of disability-specific microaggressions. Only 1% of the 833 participants in Study 3 did not endorse any of the microaggressions included in the AMS. This is an important contribution, as little research has been conducted to verify whether people with physical disabilities report microaggressions. In terms of the scale structure, the AMS's four underlying factors, including Helplessness, Minimization, Denial of Personhood, and Otherization, appear to be related, but distinct. The factors established in this study aligned well with those developed in Keller and Galgay's (2010) qualitative study, and represent seven of their eight domains, with "denial of privacy" the only theme not represented. Although the factors do not reflect all

of the domains, qualitative data are not intended to be representative, and the present data reflect a broader population. Results indicate the 20-item AMS appears to be a statistically sound measure of physical disability microaggressions and a good approximation of the everyday realities of microaggressions experienced by people with physical disabilities.

As physical disability is not a unidimensional category, within-group differences were found to influence reported rates of lifetime ableist microaggressions in this sample. Findings with regards to disability visibility suggest that people whose bodies are perceived as farthest from the able-bodied norm experience the most ableist microaggressions. Similarly, people with congenital disabilities and childhood onset of impairment reported more microaggressions than people with acquired disabilities and onset after age 40, respectively. Although this may reflect greater time to experience ableist microaggressions across the lifespan, it might also indicate important differences between these groups. As participants with mild or moderate disabilities were more likely to drop out of the study, it is important to note that AMS items may be less representative of the experiences of people with mild or moderate disabilities. It may also be helpful in future studies to explore how disability type relates to experiences with physical disability microaggressions.

In addition, whereas researchers have reported that 4.5 to 8.5% of the U.S. population is a sexual minority (Gates, 2011), 16.4% of the participants in the present study identified as lesbian, gay, bisexual, or queer (LGBQ). Other recent studies have also found that people with disabilities are more likely than able-bodied people to identify as LGBQ (Bernert, Ding, & Hoban, 2012; Fredriksen-Goldsen, Kim, & Barkan, 2012). Although the study's recruitment materials did not indicate any interest in or benefit from sexual minority identification, "sexual orientation" was included as a keyword on MTurk due to simultaneous recruitment for another study of LGBQ people with disabilities, and this may have increased sexual minority participation. LGBQ participants had the option to take an additional survey for a study of LGBQ people with disabilities that used the AMS data collected for this study, as well as other measures.

The results of this study have implications for the practice and training of mental health providers. Given that 15 to 20% of the U.S. population are people with physical disabilities (Kraus, Stoddard, & Gilmartin, 1996), clinicians are very likely to serve members of this community. Previous research suggests that the ambiguity of microaggressions may be linked to psychological consequences (Salvatore & Shelton, 2007), and therefore, treatment providers can support their clients by acknowledging and validating experiences of ableist microaggressions. Providers may want to explicitly inquire about experiences with microaggressions along with questions about blatant

harassment or discrimination. In fact, it may be helpful for mental health providers to employ the AMS in clinical settings to assess the amount and type of microaggressions to which their clients have been exposed, and to begin a dialogue about microaggressions.

Several limitations should be taken into account when interpreting the findings of the present study, and future research would benefit from addressing these limitations. Although there were many strengths in the development of the AMS, participant recruitment for the cognitive interviews and pilot study were conducted through the social and professional networks of the investigators, and included small samples. In addition, recruitment for Study 2 and 3 was conducted simultaneously, and although recommended in the literature, this method did not allow for item revision based on the results of Study 2. In addition, the use of MTurk as the primary recruitment method had many advantages, but the sample may not be comparable to others acquired through different methods. Although research has shown that MTurk workers make satisfactory social science participants, the use of MTurk as a source of income for participants may alter their primary motivations and conditions of participation. Mirroring MTurk demographics overall (Paolacci et al., 2010), the present sample demonstrated lower income and higher educational attainment than a very large sample of working-age adults with disabilities based on U.S. Census data (Erickson, Lee, & von Schrader, 2014). As the latter is presumed to be representative of the U.S. population, differences regarding income and education may influence the generalizability of the current findings. Finally, participation was limited to people with access to a computer and Internet, as well as to those interested in completing a psychological survey.

In terms of item development and scale structure, two of Keller and Galgay's (2010) domains only pertained to two items each in the pAMS, which made it unlikely for these domains to form separate factors. Of the four subscales that emerged, the Minimization subscale demonstrated lower internal consistency than the others. This subscale contained fewer items than the other factors with three items, including two that may not be applicable to certain groups of participants. In addition, the Denial of Personhood and Otherization factors correlated highly with one another, which should be considered in future studies. With regards to model fit, good model fit was demonstrated in the EFA but fair model fit was found in the CFA. Van Prooijen and Van Der Kloot (2001) provide possible explanations as to why a CFA model might not have good fit despite a well-fitting EFA, such as high model constraints, non-normal data, and sample size factors; future studies on different samples may add further insight into the CFA model for this measure.

Regarding validity and reliability, although relationships with convergent validity scales were positive, they were weaker than expected. Convergent

validity should, therefore, be explored using alternative scales in future studies. This finding may also reflect a moderated relationship, which should be investigated. For example, as social support has been posited to buffer against the negative effects of marginalization (Meyer, 2003). Perhaps the relationship between microaggressions and depression is moderated by perceived social support. Nonetheless, the findings demonstrating that greater disability visibility correlate with higher rates of reported ableist microaggressions lends additional support to the validity of the AMS. In terms of discriminant validity, the SDRS-5 demonstrated low internal consistency for this sample, which should be considered in future research. Furthermore, several variables across other validity scales demonstrated non-normality in this sample, which may have influenced the findings. Alternative forms of validity, such as predictive and incremental validity, were not tested and could be examined in future research using a measure of discrimination such as the Everyday Discrimination Scale (Williams, Yu, Jackson, & Anderson, 1997), which can be tailored to include physical disability discrimination.

Findings in this study inform several additional suggestions for future research on ableist microaggressions. First, the weak relationship between the AMS and perceived stress was unexpected, as previous research has connected other forms of microaggressions with perceived stress (Torres et al., 2010). Future studies should continue to examine this relationship and potential moderating variables, such as problem solving skills and coping strategies. Furthermore, future studies should explore how physical disability microaggressions influence mental health outcomes such as anxiety, depression, and internalized shame, as well as use of coping strategies. Second, participants who reported severe or very severe disabilities indicated the highest rates of ableist microaggressions, and therefore, it may be fruitful to examine correlates of microaggressions specific to this population. Third, the AMS is a measure of perceived microaggressions, and the extent to which perceived microaggressions are related to experiences of microaggressions should be clarified further.

Future scale development research should also consider the use of cognitive interviewing in the development of initial scale items and MTurk for participant recruitment. No previous microaggression scales have been developed using cognitive interviewing, which constituted an important methodological technique in this study. The use of MTurk allowed for greater oversight regarding participant demographics than traditional online surveys, and increased participation of diverse communities. In addition, the national community sample represents an important contribution to the literature on microaggressions, as many other studies have been limited to university student samples.

Authors' Note

Kristin J. Conover conducted the study as a doctoral student at the University of California, Santa Barbara.

Acknowledgments

A special thank you to Drs. Merith Cosden, Linda Mona, Susan Torres-Harding, and Richard Keller for your contributions to this study.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Akaike, H. (1977). On the entropy maximization principle. In P. R. Krishnaiah (Ed.), *Applications of statistics* (pp. 27–41). Amsterdam, The Netherlands: North-Holland.
- Americans with Disabilities Act Amendments Act, Pub. L. No. 110-325, § 4(a)(1), 122 Stat. 3553 (2008). Retrieved from <https://www.eeoc.gov/laws/statutes/adaaa.cfm>
- Andresen, E. B., Malmgren, J. A., Carter, W. B., & Patrick, D. L. (1994). Screening for depression in well older adults: Evaluation of a short form of the CES-D. *American Journal of Preventive Medicine, 10*, 77–84.
- Balsam, K. F., Molina, Y., Beadnell, B., Simoni, J., & Walters, K. (2011). Measuring multiple minority stress: The LGBT people of color microaggressions scale. *Cultural Diversity and Ethnic Minority Psychology, 17*, 163–174. doi:10.1037/a0023244
- Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly, 71*, 287–311. doi:10.1093/poq/nfm006
- Bernert, D. J., Ding, K., & Hoban, M. T. (2012). Sexual and substance use behaviors of college students with disabilities. *American Journal of Health Behaviors, 36*, 459–471. doi:10.5993/AJHB.36.4.3
- Blume, A. W., Lovato, L. V., Thyken, B. N., & Denny, N. (2012). The relationship of microaggressions with alcohol use and anxiety among ethnic minority college students in a historically White institution. *Cultural Diversity and Ethnic Minority Psychology, 18*, 45–54. doi:10.1037/a0025457
- Brooks, V. (1981). *Minority stress and lesbian women*. Lexington, MA: Lexington Books, D.C. Heath.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY: Guilford Press.

- 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). Beverly Hills, CA: Sage.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6, 3–5. doi:10.1177/1745691610393980
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233–255. doi:10.1207/S15328007SEM0902_5
- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24, 385–396.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349–354. doi:10.1037/h0047358
- Dowdy, E., Twyford, J. M., Chin, J. K., DiStefano, C. A., Kamphaus, R. W., & Mays, K. L. (2011). Factor structure of the BASC-2 Behavioral and Emotional Screening System Student Form. *Psychological Assessment*, 23, 379–387. doi:10.1037/a0021843
- Emerson, E. (2010). Self-reported exposure to disablism is associated with poorer self-reported health and well-being among adults with intellectual disabilities in England: A cross-sectional survey. *Public Health*, 124, 682–689. doi:10.1016/j.puhe.2010.08.020
- Erickson, W., Lee, C., & von Schrader, S. (2014). *Disability statistics from the 2012 American Community Survey (ACS)*. Ithaca, NY: Cornell University Employment and Disability Institute (EDI). Retrieved from <http://www.disabilitystatistics.org/>
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272–299. doi:10.1037/1082-989X.4.3.272
- Fredriksen-Goldsen, K. I., Kim, H.-J., & Barkan, S. E. (2012). Disability among lesbian, gay, and bisexual adults: Disparities in prevalence and risk. *American Journal of Public Health*, 102, 16–21. doi:10.2105/AJPH.2011.300379
- Gates, G. (2011). *How many people are lesbian, gay, bisexual, and transgender?* The Williams Institute. Retrieved from the William Institutes UCLA School of Law website: <http://williamsinstitute.law.ucla.edu/research/census-lgbt-demographics-studies/how-many-people-are-lesbian-gay-bisexual-and-transgender/>
- Gayman, M. D., Brown, R. L., & Cui, M. (2011). Depressive symptoms and bodily pain: The role of physical disability and social stress. *Stress and Health*, 27, 52–63. doi:10.1002/smi.1319
- Goldstein, R. B., Dawson, D. A., Smith, S. M., & Grant, B. F. (2012). Antisocial behavioral syndromes and 3-year quality-of-life outcomes in United States adults. *Acta Psychiatrica Scandinavica*, 126, 137–150. doi:10.1111/j.1600-0447.2012.01848.x
- Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

- Green, S. E. (2003). "What do you mean 'what's wrong with her?'": Stigma and the lives of families of children with disabilities. *Social Science & Medicine*, 57, 1361–1374. doi:10.1016/S0277-9536(02)00511-7
- Hays, R. D., Hayashi, T., & Stewart, A. L. (1989). A five-item measure of socially desirable response set. *Educational and Psychological Measurement*, 49, 629–636. doi:10.1177/001316448904900315
- Heppner, M. J., Humphrey, C. F., Hillenbrand-Gunn, T. L., & DeBord, K. A. (1995). The differential effects of rape prevention programming on attitudes, behavior, and knowledge. *Journal of Counseling Psychology*, 42, 508–518. doi:10.1037/0022-0167.42.4.508
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. doi:10.1080/10705519909540118
- Huynh, V. W. (2012). Ethnic microaggressions and the depressive and somatic symptoms of Latino and Asian American adolescents. *Journal of Youth and Adolescence*, 41, 831–846. doi:10.1007/s10964-012-9756-9
- Kahn, J. H. (2006). Factor analysis in counseling psychology research, training, and practice: Principles, advances, and applications. *The Counseling Psychologist*, 34, 684–718. doi:10.1177/0011000006286347
- Karam, F., Bérard, A., Sheehy, O., Huneau, M-C., Briggs, G., Briggs, G., ...the OTIS Research Committee. (2012). Reliability and validity of the 4-item Perceived Stress Scale among pregnant women: Results from the OTIS antidepressants study. *Research in Nursing & Health*, 35, 363–375. doi:10.1002/nur.2148
- Keller, R. M., & Galgay, C. E. (2010). Microaggressive experiences of people with disabilities. In D.W. Sue (Ed.), *Microaggressions and marginality: Manifestations, dynamics, and impact* (pp. 241–268). New York, NY: Wiley & Sons.
- Kraus, L., Stoddard, S., & Gilmartin, D. (1996). *Chartbook on disability in the United States: An InfoUse report*. Washington, DC: U.S. National Institute on Disability and Rehabilitation Research.
- Lesage, F. X., Berjot, S., & Deschamps, F. (2012). Clinical stress assessment using a visual analogue scale. *Occupational Medicine*, 62, 600–605. doi:10.1093/occ-med/kqs140
- Lewis, J. A., & Neville, H. A. (2015). Construction and initial validation of the Gendered Racial Microaggressions Scale for Black Women. *Journal of Counseling Psychology*, 62, 289–302. doi:10.1037/cou0000062
- Li, L., & Moore, D. (2001). Disability and illicit drug use: An application of labeling theory. *Deviant Behavior*, 22, 1–21. doi:10.1080/016396201750065784
- Mazur, E. (2008). Negative and positive disability-related events and adjustment of parents with acquired physical disabilities and of their adolescent children. *Journal of Children and Family Studies*, 17, 517–537. doi:10.1007/s10826-007-9171-0
- Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology*, 9, 568–592. doi:10.1037/0021-9010.93.3.568

- Meltzer, H., Brugha, T., Dennis, M. S., Hassiotis, R. J., McManus, S., Rai, D., & Bebbington, P. (2012). The influence of disability on suicidal behaviour. *European Journal of Disability Research*, 6, 1–12. doi:10.1016/j.alter.2011.11.004
- Mercer, S. H., Zeigler-Hill, V., Wallace, M., & Hayes, D. M. (2011). Development and initial validation of the Inventory of Microaggressions Against Black Individuals. *Journal of Counseling Psychology*, 58, 457–469. doi:10.1037/a0024937
- Meyer, I. H. (1995). Minority stress and mental health in gay men. *Journal of Health and Social Behavior*, 36, 38–56. doi:10.2307/2137286
- Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence. *Psychological Bulletin*, 129, 674–697. doi:10.1037/0033-2909.129.5.674
- Miller, S. M. (2009). The measurement of self-efficacy in persons with spinal cord injury: Psychometric validation of the Mooring Self-Efficacy Scale. *Disability and Rehabilitation*, 31, 988–993. doi:10.1080/09638280802378025
- Mitchell, A. M., Crane, P. A., & Kim, Y. (2008). Perceived stress in survivors of suicide: Psychometric properties of the Perceived Stress Scale. *Research in Nursing & Health*, 31, 576–585. doi:10.1002/nur.20284
- Muthén, L., & Muthén, B. (1998-2016). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.
- Nadal, K. L. (2011). The Racial and Ethnic Microaggressions Scale (REMS): Construction, reliability, and validity. *Journal of Counseling Psychology*, 58, 470–480. doi:10.1037/a0025193
- Nadal, K. L. (2013). *That's so gay! Microaggressions and the lesbian, gay, bisexual, and transgender community*. Washington DC: American Psychological Association.
- Nadal, K. L., Issa, M-A., Griffin, K., Hamit, S., & Lyons, O. (2010). Religious microaggressions in the United States: Mental health implications for religious minority groups. In D. W. Sue (Ed.), *Microaggressions and marginality: Manifestations, dynamics, and impact* (pp. 287–310). New York, NY: Wiley.
- Nadal, K. L., Skolnik, A., & Wong, Y. (2012). Interpersonal and systemic microaggressions toward transgender people: Implications for counseling. *Journal of LGBT Issues in Counseling*, 6, 55–82. doi:10.1080/15538605.2012.648583
- Olkin, R. (2004). Making research accessible to participants with disabilities. *Journal of Multicultural Counseling and Development*, 32, 332–343.
- Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, 5, 411–419.
- Pierce, C., Carew, J., Pierce-Gonzalez, D., & Willis, D. (1978). An experiment in racism: TV commercials. In C. Pierce (Ed.), *Television and education* (pp. 62–88). Beverly Hills, CA: Sage.
- Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1(3), 385–401. doi:10.1177/014662167700100306
- Salvatore, J., & Shelton, J. N. (2007). Cognitive costs of exposure to racial prejudice. *Psychological Science*, 18, 810–815. doi:10.1111/j.1467-9280.2007.01984.x

- Scott, K. M., Von Korff, M., Alonso, J., Angermeyer, M. C., Bromet, E., Fayyad, J., de Girolamo, G., et al. (2009). Mental-physical co-morbidity and its relationship with disability: Results from the World Mental Health Surveys. *Psychological Medicine*, 39, 33–43. doi:10.1017/S0033291708003188
- Sue, D. W. (Ed.). (2010). *Microaggressions and marginality: Manifestations, dynamics, and impact*. New York, NY: Wiley.
- Sue, D. W., Capodilupo, C. M., & Holder, A. M. B. (2008). Racial microaggressions in the life experience of Black Americans. *Professional Psychology: Research and Practice*, 39, 329–336. doi:10.1037/0735-7028.39.3.329
- Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A. M. B., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: Implications for clinical practice. *American Psychologist*, 62, 271–286. doi:10.1037/0003-066X.62.4.271
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon/Pearson Education.
- Torres, L., Driscoll, M. W., & Burrow, A. L. (2010). Racial microaggressions and psychological functioning among highly achieving African-Americans: A mixed-methods approach. *Journal of Social and Clinical Psychology*, 29, 1074–1099. doi:10.1521/jscp.2010.29.10.1074
- Torres-Harding, S. R., Andrade, A. R., & Romero Diaz, C. E. (2012). The Racial Microaggressions Scale (RMAS): A new scale to measure experiences of racial microaggressions in people of color. *Cultural Diversity and Ethnic Minority Psychology*, 18, 153–164. doi:10.1037/a0027658
- Turner, R. J., Lloyd, D. A., & Taylor, J. (2006). Physical disability and mental health: An epidemiology of psychiatric and substance disorders. *Rehabilitation Psychology*, 51, 214–223. doi:10.1037/0090-5550.51.3.214
- Turner, R. J., & McLean, P. D. (1989). Physical disability and psychological distress. *Rehabilitation Psychology*, 34, 225–242. doi:10.1037/h0091727
- Van Prooijen, J.-W., & Van Der Kloot, W. A. (2001). Confirmatory analysis of exploratively obtained factor structures. *Educational and Psychological Measurement*, 61, 777–792. doi:10.1177/00131640121971518
- Vega, T. (2014, March 21). Students see many slights as racial 'microaggressions.' *The New York Times*, p. A1. Retrieved from <http://www.nytimes.com/2014/03/22/us/as-diversity-increases-slights-get-subtler-but-still-sting>
- Wei, M., Alvarez, A. N., Ku, T.-Y., Russell, D. W., & Bonett, D. G. (2010). Development and validation of a Coping with Discrimination Scale: Factor structure, reliability, and validity. *Journal of Counseling Psychology*, 57, 328–344. doi:10.1037/a0019969
- Wei, M., Wang, K. T., & Ku, T.-Y. (2012). A development and validation of the Perceived Language Discrimination Scale. *Cultural Diversity and Ethnic Minority Psychology*, 18, 340–351. doi:10.1037/a0029453
- Williams, D. R., Yu, Y., Jackson, J. S., & Anderson, N. B. (1997). Racial differences in physical and mental health: Socioeconomic status, stress, and discrimination. *Journal of Health Psychology*, 2, 335–351. doi:10.1177/135910539700200305

- Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage.
- Wright, A. J., & Wegner, R. T. (2012). Homonegative microaggressions and their impact on LGB individuals: A measure validity study. *Journal of LGBT Issues in Counseling*, 6, 35–54. doi:10.1080/15538605.2012.648578
- Zauszniewski, J. A., Morris, D. L., Preechawong, S., & Chang, H-J. (2004). Reports on depressive symptoms in older adults with chronic conditions. *Research and Theory for Nursing Practice: An International Journal*, 18, 185–196. doi:10.1891/rtnp.18.2.185.61280

Author Biographies

Kristin J. Conover is an assistant professor of clinical psychology at the California School of Professional Psychology at Alliant International University, Los Angeles. She received her PhD in counseling, clinical, and school psychology from the University of California, Santa Barbara, with a doctoral emphasis in feminist studies. Her research focuses on physical disability and LGBTQ+ populations, and she provides psychotherapy in a private practice setting.

Tania Israel is a professor of counseling, clinical, and school psychology at the University of California, Santa Barbara. As Director of Research and Interventions for Sexual and Gender Minority Empowerment (project RISE), she conducts research relevant to scholarship, policy, and practice on interventions to support the mental health and well-being of LGBTQ individuals and communities. Dr. Israel is a Fellow of the American Psychological Association and a past president of the Society of Counseling Psychology.

Karen Nylund-Gibson is an associate professor in the Gevirtz School at the University of California, Santa Barbara. Her focus is on latent variable models, specifically mixture models, and she has published many articles and book chapters on developments, best practices, and applications of latent class analysis, latent transition analysis, and growth mixture modeling. Her research helps to bridge the gap between advanced statistical literature and application of these models to substantive issues.