EDUCATIONAL ASSESSMENT & EVALUATION | RESEARCH ARTICLE

The construct validity of the Arabic version of the Chedoke-McMaster attitudes towards children with handicaps scale

Ghaleb Hamad Alnahdi

Abstract: This study aimed to examine the construct validity of the Arabic version of the Chedoke-McMaster Attitudes towards Children with Handicaps (CATCH) scale in elementary school students. Data were collected from 672 elementary school students (357 boys and 315 girls). The included students ranged in age from 8 to 14 years. A statistical analysis was conducted using the following methods: confirmatory factor analysis, exploratory factor analysis, Rasch modeling, and internal consistency analysis. A good reliability coefficient was obtained for the scale. Confirmatory factor analysis indicated that the observed data failed to fit the hypothesized three-factor structure of the scale. However, the indices had a good fit after the negatively phrased items were excluded from the model. Confirmatory factor analysis did not support the construct validity of the 36-item Arabic version of the CATCH scale in elementary school students. The proposed 18-item Arabic version of the scale should be retested with different samples to confirm this study’s findings.

Subjects: Educational Research; Inclusion and Special Educational Needs; Child Development

Keywords: confirmatory factor analysis; attitudes; Arabic; Saudi Arabia; CATCH; children with disabilities; Rasch analysis

1. Introduction

In education-related fields, advocating for the inclusion, when possible, of students with disabilities among their peers is the prevailing practice. One of the main goals of inclusion is to increase the opportunities for children with disabilities to participate with their peers in school activities. The success of the inclusion process is not smooth or guaranteed (Bossaert & Petry, 2013). One issue
that may represent a challenge to the success of the inclusion process is students’ attitudes towards their peers with disabilities (McDougall et al., 2004).

Attitude is defined as “the affect for or against a psychological object” (Thurstone, 1931). Ensuring positive interactions between children with disabilities and their peers is important for instilling positive attitudes towards people with disabilities. Physical integration alone is not the goal of inclusion (Scior et al., 2013). Many studies have shown that attitudes towards people with disabilities can significantly predict behavior (Ajzen & Fishbein, 1980; Kraus, 1995).

Therefore, examining students’ attitudes is important, and a reliable scale is needed to assess children’s attitudes towards their peers with disabilities. An Arabic scale does not exist for this purpose; therefore, a scale in the Arabic language is needed to measure students’ attitudes towards their peers with disabilities.

Several scales were considered for adaptation in this study, including the Chedoke–McMaster Attitudes towards Children with Handicaps (CATCH) scale. The CATCH scale was recommended as one of the best scales for measuring children’s attitudes towards their peers with disabilities (Vignes et al., 2008). It is one of the few scales that measures three components of attitudes. The CATCH scale was developed by Rosenbaum et al. (1986) and it was administered to 304 children from grades 5 to 8 in four schools in Canada. In addition, the measure has been used in many countries with different languages and has shown acceptable psychometric properties (Rosenbaum et al., 1988; Vignes et al., 2008).

Theoretically, the CATCH scale is based on the three-component attitudes model proposed by Triandis in 1971 (Rosenbaum et al., 1986). These three components are: (A) affective attitudes (“I would like a handicapped child to live next-door to me”); (B) behavioral intention (“I would try to stay away from a handicapped child”); and (C) cognitive attitudes (“Handicapped children can make new friends”). The CATCH scale is consisting of 36 items divided into 3 subscales (3x12 = 36 items). These three dimensions include (a) an affective component that is related to feelings, (b) a behavioral component that is related to behavioral intention, and (c) a cognitive component that is related to beliefs. This scale has been used in studies for children as from 9 to 16 years (Bossaert & Petry, 2013; Rosenbaum et al., 1986).

Validation studies for CATCH have been performed in many countries (for example, Canada, United Kingdom, and Netherland). However, none of these studies have been conducted using an Arabic sample with an Arabic version of the scale. Rosenbaum et al. (1986) examined the properties of the scale with data collected from 304 students in elementary schools in Canada. The hypothesized model with three factors failed to fit the observed data. They concluded that by two subscales (factors), the affective and behavioral components being grouped into one subscale and the cognitive subscale, the model fit did improved.

A study conducted in the Netherlands reported different results (De Boer et al., 2012). They concluded that the items in the cognitive subscale did not fit within the scale and that the data fit a single-dimension scale with only 14 items. Six items were from affective subscale, 8 items were from the behavioral intent subscale, and no items were from the cognitive subscale. In England, a recent study used the Rasch model; data were collected from 416 students aged 7 to 16 to examine both the fit of the entire CATCH scale and the fit at the item level (Armstrong et al., 2016). The findings from that study did not support all three components of the scale. The authors found that the internal consistency of the cognitive component subscale was lower than that of the other two subscales. In addition, they were cautious regarding the cognitive subscale, as they found that children responded more negatively on this subscale. They suggested that the subscales would be more effective when used separately.
The present study aims to extend the available research on the factor structure of the CATCH by using a Saudi sample. The following research question will be addressed in this study: (1) Is the proposed three-factor structure of the CATCH preserved in the Arabic version of the scale?

2. Materials and methods

To test the data, a confirmatory factor analysis (CFA) was performed using Amos 20, and an exploratory factor analysis (EFA) was conducted using SPSS 21.0 (Statistical Package for Social Sciences, Chicago, IL, USA). EFA used “to examine the underlying dimensionality of the item set. Thus, they can group a large item set into meaningful subsets that measure different factors. The primary reason for using EFA is that it allows items to be related to any of the factors underlying examinee responses. As a result, [this study] can easily identify items that do not measure an intended factor” (Worthington & Whittaker, 2006, p. 807). “A fundamental feature of CFA is its hypothesis-driven nature” (Brown, 2015, p. 1). “EFA is a data-driven approach” (Brown, 2015, p. 14), therefore, it is useful to start from data when items do not measure the intended factor (Worthington & Whittaker, 2006). In addition to using the classic test theory approach to test the model fit, R software (eRm Package) was used to examine the fit of items from the item response theory perspective (Rasch modeling), and item fit statistics were checked and items with significant chi-square were considered to be potentially misfit items.

2.1. Participants

The study sample consisted of 672 students from nine elementary schools in Riyadh, Saudi Arabia (357 boys and 315 girls). The included students ranged in age from 8 to 14 years. The school system in Saudi Arabia is separated by gender. Therefore, the boys and girls in this study attend different schools. The study instrument was distributed to students in class during the school day, and an explanation of the instrument and necessary clarifications were provided to the students by the mediator and the teacher. A convenience sampling was implemented in this study, and the schools that were included were those where the mediators (teachers) worked and those nearby; however, the included schools were in different areas in the city of Riyadh. In addition, the sample size in this study fulfilled the recommendation of having 7 times number of items to examine the Structural validity, which means 36 * 7 = 252 (see the recently developed COSMIN study design checklist for patient reported outcome measurement instruments; Mokkink et al., 2019). Institutional review approval (IRB) to conduct this study was obtained from the Deanship of Scientific Research at the university. Parents’ written consent were obtained through school administration before we conduct this study.

2.2. Instruments

The CATCH is a 36-item measurement scale with three components. Twelve items each are included in the affective component (A), the behavioral component (B), and the cognitive component (C). The back-translation technique (Brislin, 1970) was used to ensure the accuracy of the translation of the scale into Arabic. First, the English version was translated into Arabic. Second, the Arabic version was given to a bilingual faculty member, who translated it back into English. Third, the new English version was compared to the original instrument to ensure that each item retained the same meaning. Scores were given from 4, strongly agree; 3, agree; 2, disagree; and 1, strongly disagree. The scores were reversed for all 18 negatively phrased items to 1, strongly agree; 2, agree; 3, disagree; and 4, strongly disagree (see Appendix A for the full scale). High total scores on the scale imply high positive attitudes towards peers with disabilities.

2.3. Analysis procedures

This study included four different types of analysis and eight total steps. First, the internal consistency (Cronbach’s alpha) was calculated for the three factors and for the scale as a whole. It was also calculated for the last two models after separating the negatively phrased items from the positively phrased items, as the EFA results suggested. Second, a CFA was conducted to examine the fit indices of the original three-factor model. Third, a CFA was conducted to examine the two-factor model created when the first and the second components were merged together.
into one component (A/B, C), which was deemed appropriate based on the results of Rosenbaum et al's factor analysis (Rosenbaum et al., 1986). Therefore, the third model has only two factors, the first factor (24 items) consist 12 item from the affective subscale and 12 items from the behavioral intent subscale. The second factor is the cognitive subscale with 12 items. Fourth, a CFA was conducted to examine the one-factor model, which was one of the solutions discussed by Bossaert and Petry (2013) after they conducted a factor structure analysis of the scale.

Fifth, in addition to the classic test theory, Rasch model statistics were computed to find the misfitting items. Sixth, a CFA was conducted to examine the fit indices after removing 6 items that did not fit according to the Rasch analysis. Seventh, an EFA was conducted to examine the item loadings and determine why the items did not load as expected in CFA based on the original structure of the scale. The results indicated that the negatively phrased items loaded differently. Eighth, a CFA was conducted to examine the three-factor model with the 18 positively phrased items, and another CFA was conducted to examine the three-factor model with the 18 negatively phrased items.

3. Results

3.1. Internal consistency

Table 1 shows that the internal consistency (Cronbach’s alpha) was 0.636 for the affective factor item scores, 0.651 for the behavioral factor items, and the lowest for the cognitive factor items, 0.542. The total scale had an internal consistency of 0.843 (see Table 1), which indicates a good level of internal consistency (George & Mallery, 2003). In addition, better internal consistency was found for the scale with only positively phrased items (0.861) compared with the scale with only negatively phrased items (0.772).

3.2. EFA loadings/parallel analysis

“Very often, those who conduct EFA follow up the analysis with CFA” (Brown, 2015, p. ix). One dimension (with 4 components) was suggested based on the results of Horn’s parallel analysis for component retention, which adjusts for finite sample bias in the retention of components and retains factors with eigenvalues larger than the 95th percentile estimated from the random data. Horn’s parallel analysis was recommended for determining the number of components to retain, for which it uses a Monte-Carlo simulation method (Verdugo et al., 2014).

As with the English version, the scree plot recommended the three-component solution. Three factors explained 31% of the variance. With reference to the factor structure, twenty of the 36 items had loadings >.30 for the individual factors. It was found that the items loaded on factor 2 (14, 18, 20, 22, 28, 30, 34, and 36) and factor 3 (10, 12, 16, 24, and 26) were all negatively phrased regardless of the proposed factor. Six items had loadings <.30 on all three factors (2, 3, 5, 7, 11, and 17). It was noted that participants perceived the negatively phrased items differently. Therefore, the author conducted another CFA to test whether the model fit improved after the negatively phrased items were excluded (Model 5). In addition, a CFA was conducted for the 18 negatively phrased items (Model 6).

| Table 1. Internal consistency statistics |
|-----------------------------------------|
| Cronbach’s Alpha | N of Items |
|------------------|-----------|
| Affective        | .636      | 12        |
| Behavioral intent| .651      | 12        |
| Cognitive        | .542      | 12        |
| Overall          | .843      | 36        |
| Overall (positively phrased items)     | .861      | 18        |
| Overall (negatively phrased items)     | .772      | 18        |
3.3. Rasch analysis

In addition to using the classic test theory approach to test the model fit, this study examined the fit of the items from the modern test theory perspective (Rasch modeling). The computation of the Rasch model statistics revealed 6 misfitting items. The p values were <.05 for these items, which means that these items are statistically unfit indices (see Table 2): item 2 (I wouldn’t introduce a handicapped child to my friend), item 3 (Handicapped children can do lots of things for themselves), item 10 (I would be afraid of a handicapped child), item 11 (I would talk to a handicapped child I didn’t know), item 12 (Handicapped children don’t like to make friends), and item 35 (I would miss recess to keep a handicapped child company). These items were

Table 2. Item fit statistics in the rasch modeling

| Item | $X^2$ | p value | Outfit MSQ | Infit MSQ | Outfit t | Infit t |
|------|-------|---------|------------|-----------|----------|---------|
| Item1 | 0.647 | 0.967   | 0.999      | -0.18     | 0.02     |
| Item2 | 0.006* |         | 1.197      | 1.054     | 2.45     | 1.3     |
| Item3 | 0.000* |         | 1.27       | 1.107     | 3.71     | 2.76    |
| Item4 | 0.182 | 1.065   | 1.004      | 1.04      | 0.12     |
| Item5 | 0.747 | 0.946   | 1.006      | -0.27     | 0.1      |
| Item6 | 0.541 | 0.988   | 0.999      | -0.17     | -0.03    |
| Item7 | 0.208 | 1.058   | 1.048      | 0.33      | 0.41     |
| Item8 | 0.913 | 0.897   | 0.92       | -1.87     | -2.15    |
| Item9 | 0.982 | 0.846   | 0.941      | -1.53     | -1.06    |
| Item10| 0.007*|         | 1.19       | 1.108     | 2.86     | 2.76    |
| Item11| 0.003*|         | 1.216      | 1.137     | 2.99     | 3.49    |
| Item12| 0.002*|         | 1.223      | 1.176     | 3.69     | 4.54    |
| Item13| 0.830 | 0.925   | 0.973      | -0.63     | -0.42    |
| Item14| 0.262 | 1.044   | 1.038      | 0.67      | 1.02     |
| Item15| 0.992 | 0.826   | 0.876      | -2.2      | -2.9     |
| Item16| 0.178 | 1.066   | 1.063      | 1.11      | 1.72     |
| Item17| 0.265 | 1.043   | 1.053      | 0.74      | 1.47     |
| Item18| 0.816 | 0.929   | 0.971      | -1.2      | -0.81    |
| Item19| 0.545 | 0.987   | 1.003      | -0.21     | 0.09     |
| Item20| 0.463 | 1.002   | 1.033      | 0.06      | 0.92     |
| Item21| 0.835 | 0.924   | 0.991      | -0.8      | -0.15    |
| Item22| 0.573 | 0.982   | 1.011      | -0.3      | 0.32     |
| Item23| 0.982 | 0.846   | 0.922      | -1.54     | -1.44    |
| Item24| 0.103 | 1.093   | 1.117      | 1.51      | 3.14     |
| Item25| 0.988 | 0.835   | 0.886      | -2.64     | -3.2     |
| Item26| 0.185 | 1.064   | 1.085      | 1.11      | 2.29     |
| Item27| 0.650 | 0.967   | 0.981      | -0.41     | -0.44    |
| Item28| 0.539 | 0.988   | 1.032      | -0.19     | 0.89     |
| Item29| 0.68  | 0.961   | 0.958      | -0.66     | -1.01    |
| Item30| 0.942 | 0.882   | 0.909      | -2.03     | -2.6     |
| Item31| 0.986 | 0.839   | 0.891      | -2.26     | -2.77    |
| Item32| 0.970 | 0.86    | 0.938      | -1.95     | -1.55    |
| Item33| 0.993 | 0.822   | 0.9       | -1.88     | -1.93    |
| Item34| 0.957 | 0.872   | 0.903      | -2.31     | -2.75    |
| Item35| 0.010*|         | 1.181      | 0.928     | 1.83     | -1.43   |
| Item36| 0.891 | 0.905   | 0.934      | -1.48     | -1.82    |

*Significant at p < .05, $X^2 = \text{Chi-square}$. 
excluded from Model 4 during the CFA. Five of these items (2, 3, 10, 11, and 12) loaded on different proposed factors during the EFA. In this analysis, only items fit statistics were considered the 36 items and used to find misfit items with significant chi-square were considered to be potentially misfit.

### 3.4. Confirmatory factor analysis

The construct validity of the three-factor structure of the scale was examined via CFA. The three-component model with 36 items was examined by CFA and failed to fit the data. Therefore, further rounds of CFA analysis were conducted to compare the different models (see Table 3).

CFA was utilized to examine the construct validity of the scale. In this analysis, the root-mean-square error of approximation (RMSEA) was reported, as recommended by Hu and Bentler (1998). Other indices were also reported, including the comparative fit index (CFI) and the Chi-square to degree of freedom ratio as recommended by Schermelleh-Engel et al. (2003). The originally proposed model with three factors had a poor fit with the data in most of the indices, as indicated by the goodness-of-fit indices: SBS-$\chi^2$ = (591, N = 672) = 2644.397 and p < .001; RMSEA = 0.070; CFI = 0.502; and chi-square/degrees of freedom ratio = 4.47 (see Figure 1).

The model fit did not improve in the next three models: Model 2 with only two factors, based on Rosenbaum et al. (1986) findings; Model 3 with one factor based on de Boer et al. (2012) findings; and Model 4 with three factors excluding 6 items based on the Rasch analysis. The CFI ranged from .493 to .736 for these three models, which indicates no improvement in fit in these three models.

Next, the items were separated by the type of phrasing (positive or negative), and CFA was conducted separately for both based on the EFA findings, in which negatively phrased items loaded differently than expected. Model 5, which excluded the negative items from the model, had a good fit with the data, as shown by the fit indices: SBS-$\chi^2$ = (132, N = 672) = 320.828 and p < .001; RMSEA = 0.045; CFI = 0.911; SRMR = 0.0493; and chi-square/degrees of freedom ratio = 2.43. Notably, there were improvements in all indices compared with the values from the first model.

Model 6, which included only the negative items, had a poor fit with the data, as shown by the fit indices: SBS-$\chi^2$ = (132, N = 672) = 395.163 and p < .001; RMSEA = 0.053; CFI = 0.832; and chi-square/degrees of freedom ratio = 2.99. The values for all indices dropped compared to those in the fifth model but were better than the value for the first one with 36 items.

With respect to the fit indices, Model 5 was the best model with acceptable fit indices, such as chi-square/degrees of freedom ratio < 3 (Schermelleh-Engel et al., 2003), CFI > .9, and RMSEA < .06 (Hu & Bentler, 1998; Schermelleh-Engel et al., 2003). To conclude, Model 5 with 18 positively phrased items (see Figure 2) was the best model that adequately explained the observed data with the hypothesized three-factor model.

### Table 3. Confirmatory factor analysis statistics

| Models | SBS-$\chi^2$ | p    | df | RMSEA | CFI   | $\chi^2$/df |
|--------|-------------|------|----|-------|-------|-------------|
| M1     | 2644.397    | .000 | 591| .070  | .502  | 4.47        |
| M2     | 2685.673    | .000 | 593| .071  | .493  | 4.52        |
| M3     | 831.232     | .000 | 230| .061  | .736  | 3.61        |
| M4     | 2004.231    | .000 | 431| .072  | .542  | 4.65        |
| M5     | 320.828     | .000 | 132| .045  | .911  | .243        |
| M6     | 395.163     | .000 | 132| .053  | .832  | 2.99        |

Note. SBS-$\chi^2$ = Satorra-Bentler scaled Chi-square; df = degrees of freedom; RMSEA = root-mean-square error of approximation; CFI = comparative fit index; and $\chi^2$/df = chi square to degree of freedom ratio (Arbuckle, 2010). M1 = 3 factors (36 items), M2 = 2 factors (A/B, C), M3 = 1 factor, M4 = 3 factors (~6 misfitted items), M5 = 3 factors (only positive items (18)), and M6 = 3 factors (only negative items (18))
4. Discussion
The main goal of the current study was to investigate the construct validity of the Arabic version of the Chedoke–McMaster Attitudes towards Children with Handicaps (CATCH; Rosenbaum et al., 1986) scale by examining its psychometric properties using a sample of Saudi students from nine elementary schools in Riyadh. The findings indicate the potential utility of the Arabic version when only the positively phrased items are included, since the proposed structure of the original version is preserved.

This study used the modern test theory approach to examine the fit of the scale as a whole and at the item level; items 2, 3, and 11 did not have good fit according to the Rasch analysis and did not load (>0.30) on any of the three factors, according to the EFA. Furthermore, items 10, 12, and 35 were reported as misfit items according to the Rasch modeling analysis.

Many of the negatively phrased items (such as items 24, 26, 28, and 30) were perceived differently by students, which may have caused the weak loading for these items or caused them to load on different components. Therefore, in the future, the negatively phrased items should be phrased differently, for example, an item like “I wouldn’t like a handicapped friend as much as my other friends” can be changed to be “I would like a handicapped friend as much as my other friends” . It is easier for participants to respond to positively worded items. If all items were
phrased positively, the participants would not have to switch between positively phrased items and negatively phrased items, which would prevent situations in which the positive and negative items are perceived differently by students (Roszkowski & Soven, 2010).

The participants in this study responded in way that fit the main structure of the 18-item scale. The scale with these 18 items showed much better fit indices with our students. Therefore, it is recommended that the model’s psychometric properties be examined with these 18 items and confirmed by different samples from different regions in Saudi Arabia or another Arabic country. The original scale has an adequate number of items (36), which allows the sample responses to be verified. Negatively phrased items may be necessary, to control for acquiescence bias (see Furr, 2017 for more details), if there are only a few items in the scale and there is a need to ensure that participants have read and understood the items.

5. Conclusions
This study is the first to examine the validity of the Arabic version of the CATCH scale in a sample of children and youths in Saudi Arabia. The findings supported the construct validity of the 36-item Arabic CATCH scale among elementary school students. However, evidence of the construct

---

**Figure 2. Standardized estimates for the CATCH model with 18 items (Model 5).** (a) Affective domain; (b) Behavioral domain; (c) Cognitive domain.
validity and adequate internal consistency of the 3-factor structure was obtained after removing negatively phrased items. This study is one step in the validation of the Arabic version CATCH scale, which is a useful tool for inclusive education. Therefore, the findings indicate the potential utility of the 18-item Arabic version of the scale.

In sum, this study found that the Arabic version of the CATCH scale with 18 items (out of 36 items) has acceptable psychometric properties for assessing children’s attitudes towards other children and youths with disabilities in Saudi Arabia and other Arab countries. However, further studies based on this proposed change to the scale with a larger sample are warranted to validate this study’s findings.

6. Implications
Some implications can be drawn from this study. First, the CATCH scale is a reliable instrument that could be useful to examine attitudes toward children and youth with disabilities in the Arabic region. However, the continued testing of this measure and its psychometric properties with different samples in Saudi Arabia and the Arabic region is important at the current stage. Second, this study was the first to examine the psychometric properties of the CATCH with a sample from Saudi Arabia; therefore, replications of this study with similar and different populations would be informative. Third, the scale with 18 items might have the potential to be a valid measure to examine the attitudes of children and youth towards peers with disabilities. However, further confirmation with different samples is essential. Fourth, future research would be recommended to use the Rasch analysis to examine whether the suggested 18 items scale is a unidimensional measure. As in this study the main focus was from the classical test theory approach, and misfit items analysis used to help to find the best fitting items for the CFA.

7. Limitations
There are a couple of limitations as regard this study findings. First, a convenience sampling was implemented in this study, and the schools that were included were those where the mediators (teachers) worked and those nearby; however, the included schools were in different areas in the city of Riyadh. Second, the participants were from Saudi Arabia only, a sample from other Arab countries will help to be more confident with the generalizability of this study finding.

Funding
This work was supported by the Prince Sattam bin Abdulaziz University.

Author details
Ghaleb Hamad Alnahdi
E-mail: ghalnahdi@gmail.com
ORCID ID: http://orcid.org/0000-0003-1692-9765

Special Education Department, College of Education, Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia.

Citation information
Cite this article as: The construct validity of the Arabic version of the Chedoke-McMaster attitudes towards children with handicaps scale, Ghaleb Hamad Alnahdi,Cogent Education (2020), 7: 1745540.

References
Ajzen, I., & Fishbein, M. (1980). Understanding attitudes and predicting social behaviour. Prentice hall.
Arbuckle, J. (2010). IBM SPSS amos 19 user’s guide. Amos Development Corporation.
Armstrong, M., Morris, C., Tarrant, M., Abraham, C., & Horton, M. C. (2016). Rasch analysis of the Chedoke-McMaster attitudes towards children with handicaps scale. Disability and Rehabilitation, 39(3), 281–290. https://doi.org/10.3109/09638288.2016.1140833

Bossonot, G., & Petr, K. (2013). Factorial validity of the Chedoke-McMaster attitudes towards children with handicaps scale (CATCH). Research in Developmental Disabilities, 34(4), 1336–1345. https://doi.org/10.1016/j.ridd.2013.01.007

Brislin, R. W. (1970). Back-translation for cross-cultural research. Journal of Cross-Cultural Psychology, 1(3), 185–216. https://doi.org/10.1177/135910457000100301

Brown, T. A. (2015). Confirmatory factor analysis for applied research. Guilford publications.

De Boer, A., Timmerman, M., Pijl, S. J., & Minnaert, A. (2012). The psychometric evaluation of a questionnaire to measure attitudes towards inclusive education. European Journal of Psychology of Education, 27(4), 573–589. https://doi.org/10.1007/s10212-011-0096-z

Furr, R. M. (2017). Psychometrics: An introduction. Sage Publications.

George, D., & Mallery, P. (2003). SPSS for windows step by step: A simple guide and reference, 11.0 update (4th ed.). Boston.

Hu, L.-T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychological Methods, 3(4), 424–453. https://doi.org/10.1037//1082-989x.3.4.424

Kraus, S. J. (1995). Attitudes and the prediction of behavior: A meta-analysis of the empirical literature.
## Appendix A.

### CATCH (36 items)

| Items                                                                 | Domain | Phrase Direction |
|-----------------------------------------------------------------------|--------|------------------|
| 1. I wouldn’t mind if a handicapped child sits next to me              | A      | P                |
| 2. I wouldn’t introduce a handicapped child to my friend               | B      | N                |
| 3. Handicapped children can do lots of things for themselves           | C      | P                |
| 4. I wouldn’t know what to say to a handicapped child                  | B      | N                |
| 5. Handicapped children like to play                                  | C      | P                |
| 6. I feel sorry for handicapped children                              | A      | N                |
| 7. I would stick up for a handicapped child who was being teased       | B      | P                |
| 8. Handicapped children want lots of attention from adults            | C      | N                |
| 9. I would invite a handicapped child to my birthday party             | B      | P                |
| 10. I would be afraid of a handicapped child                          | A      | N                |
| 11. I would talk to a handicapped child I didn’t know                  | B      | P                |
| 12. Handicapped children don’t like to make friends                    | C      | N                |
| 13. I would like a handicapped child to live next-door to me           | A      | P                |
| 14. Handicapped children feel sorry for themselves                    | C      | N                |
| 15. I would be happy to have a handicapped child for a special friend  | A      | P                |
| 16. I would try to stay away from a handicapped child                  | B      | N                |
| 17. Handicapped children are as happy as I am                          | C      | P                |
| 18. I wouldn’t like a handicapped friend as much as my other friends   | A      | N                |
| 19. Handicapped children know how to behave properly                   | C      | P                |
| 20. In class I wouldn’t sit next to a handicapped child                | B      | N                |
| 21. I would be pleased if a handicapped child invited me to his house. | A      | P                |
| 22. I try not to look at someone who is handicapped                    | B      | N                |
| 23. I would feel good doing a school project with a handicapped child  | A      | P                |
| 24. Handicapped children don’t have much fun                          | C      | N                |
| 25. I would invite a handicapped child to sleep over at my house       | B      | P                |
| 26. Being near someone who is handicapped scares me                    | A      | N                |
| 27. Handicapped children are interested in lots of things.             | C      | P                |
| 28. I would be embarrassed if a handicapped child invited me to his birthday | A  | N                |
| 29. I would tell my secret to a handicapped child                      | B      | P                |
| 30. Handicapped children are often sad                                 | C      | N                |
| 31. I would enjoy being with a handicapped child                       | A      | P                |
| 32. I would not go to a handicapped child’s house to play               | B      | N                |
| 33. Handicapped children can make new friends                          | C      | P                |
| 34. I feel upset when I see a handicapped child                        | A      | N                |
| 35. I would miss recess to keep a handicapped child                    | B      | P                |
| 36. Handicapped children need lots of help to do things                | C      | N                |

P = positively phrased, N = negatively phrased.
