

The Children's Treatment of Animals Questionnaire: A Rasch Analysis

Shelby Elaine McDonald*, Courtney Vidacovich†, Frank R. Ascione*, James Herbert Williams* and Kathy E. Green†

*Graduate School of Social Work, University of Denver, Denver, USA

†Department of Research Methods and Statistics, Morgridge College of Education, University of Denver, Denver, USA

Address for correspondence:

Shelby McDonald,
Graduate School of Social
Work,
University of Denver,
Craig Hall, 2148 S. High St.,
Denver, CO 80208, USA.
E-mail: sem9883@gmail.com

ABSTRACT The study of human–animal interactions is limited by a paucity of empirically validated measures of humane treatment of companion animals. The current study reports findings from a psychometric analysis of the Children's Treatment of Animals Questionnaire (CTAQ; Thompson and Gullone 2003), an instrument that assesses children's humane interactions with non-human animals. Specifically, the current study extends what is known about the psychometric properties of the CTAQ by using traditional and item response theory analyses. The CTAQ was administered to a sample of 217 school-age children whose mothers were currently receiving residential or non-residential domestic violence services. Item-analysis, exploratory factor analysis, and parallel analysis were conducted to replicate previous psychometric evaluations of the CTAQ. Rasch analysis of the CTAQ was also conducted to provide a stringent test of unidimensionality and to identify potential invariance in item functioning across various demographic variables. The CTAQ showed adequate fit to the Rasch model; one modification, removal of item 5, was required. A Rasch principal components analysis of residuals indicated a single latent dimension among the remaining 12 items. Scale use was appropriate; Rasch-Andrich thresholds increased with category values and no disordering of categories was evident. Examination of item-person maps indicated the sample was also well-targeted. Notably, evidence of differential item function was found across Spanish and English translations. Overall, findings indicate that the CTAQ is an appropriate unidimensional measure of children's humane treatment of animals. The measure is particularly well-suited for children ages 7 to 12 years who are at risk for exposure to and perpetration of animal cruelty. We recommend use of a 12-item version of the CTAQ to enhance the utility of the total score as a latent measure of children's humane treatment of companion animals.

Keywords: children, companion animals, CTAQ



In the last two decades, human–animal interactions (HAI) research has garnered increased international attention while increasing in rigor (Ascione 2008; Mellor, Patterson-Kane and Stafford 2009). In particular, substantive literature has been published on the link between animal abuse and family violence (e.g., Ascione 1998; Walton-Moss et al. 2005; Ascione et al. 2007; Simmons and Lehman 2007). Among studies in this area, an emerging literature has suggested that children who are exposed to intimate partner violence (IPV) are more likely to witness and engage in abuse of animals than children from non-violent homes (e.g., Currie 2006; Ascione et al. 2007; Volant et al. 2008). While exposure to animal abuse, independently and in conjunction with IPV exposure, has been linked to increased risk of engaging in animal cruelty (e.g., Henry 2004; Ascione et al. 2007; Gullone and Robertson 2008), research has also suggested children who witness animal abuse are no less likely to engage in humane treatment of companion animals than children who have not experienced this type of violence exposure (Thompson and Gullone 2006).

Given recent scholarly attention to the developmental implications of human–animal interactions in childhood (see Esposito et al. 2011), positive engagement with companion animals among children at risk for exposure to and perpetration of animal abuse warrants attention. Unfortunately, research on human–animal interactions is challenged by a paucity of empirically validated and theoretically sound assessment measures (Thompson and Gullone 2003; Wilson and Barker 2003; Arbour, Signal and Taylor 2009). In the extant literature, the Children's Treatment of Animals Questionnaire (CTAQ) has received recognition as a psychometrically validated measure of children's humane treatment of animals and has been supported by several HAI researchers (i.e., Thompson and Gullone 2003, 2006, 2008; Maiuro et al. 2008; Arbour, Signal and Taylor 2009). Recently developed in 2003, the CTAQ was designed to measure behavioral outcomes targeted by humane education programs (HEPs) among normative samples of children without compromised levels of empathy.

A 13-item self-report measure, the CTAQ was first evaluated with data from a sample of 69 primary school children in Australia (Thompson and Gullone 2003). To our knowledge, psychometric properties of the CTAQ have been provided by only two published studies (Thompson and Gullone 2003, 2008). In the first evaluation of the measure, Thompson and Gullone (2003) reported adequate Cronbach's alpha coefficients for the sample as a whole (0.81) and by gender (males = 0.85; females = 0.74). Test-retest reliability was examined by calculating Pearson's correlation coefficients between time 1 and time 2 administrations, yielding an adequate coefficient of 0.64 ($n = 61$, $p < 0.001$). Convergent validity was also tested in this study. Pearson's analyses were calculated to determine correlations between total scores on the CTAQ and two measures of empathy also administered to participants. Specifically, the Index of Empathy for Children (IECA; Bryant 1982) and the Empathy subscale of the Social Skills Rating System (SSRS; Gresham and Elliott 1990) were utilized. This analysis was carried out for the entire sample at both time points and yielded small to moderate statistically significant correlations between the 13-item CTAQ total score and the IECA ($r = 0.25$, $p < 0.05$) and SSRS ($r = 0.37$, $p < 0.01$). However, when convergent validity analyses were conducted separately for each gender, the correlation for boys' CTAQ scores with the SSRS was small in magnitude and non-significant.

Comparable internal consistency ($\alpha = 0.82$) of the CTAQ was also found among older students (ages 12 to 18 years) in subsequent research by Thompson and Gullone (2008). In this study of 281 Australian youth, evidence of convergent validity with scores on the Cruelty to Animals Inventory (CAI; Dadds et al. 2004) was also reported ($r = -0.25$). Although the CTAQ

was utilized in studies by Thompson and Gullone (2006) and Arbour, Signal and Taylor (2009), psychometric properties of the measure were not provided in these publications.

The CTAQ offers a unique contribution to human–animal interaction assessment, yet a salient limitation of the measure is that one of the two published psychometric evaluations of the questionnaire was conducted on a small sample with insufficient power for rigorous testing of reliability and validity. Based on recommendations by Guadagnoli and Velicer (1988), Hinkin (1995), and Schwab (1980), Thompson and Gullone's (2003) sample size of 69 children is not adequate for identifying an accurate solution in exploratory factor analysis for a 13-item measure. Therefore, we cannot assume measure structure or reliability and validity of scores on the CTAQ among primary school children. A psychometric investigation of the measure utilizing a larger sample of children in primary school is needed to determine the psychometric quality of the CTAQ.

In addition to the issue of sample size, four other points of concern about the CTAQ are evident. First, it is important to note Thompson and Gullone (2003) did not limit their sample to pet-owning children exclusively. Instead, children without companion animals in the home were told to imagine how they would act if they had a companion animal and respond to each item accordingly (Thompson and Gullone 2003, 2006). The percentage of students who reported on imagined interactions with companion animals was not reported. In addition, Thompson and Gullone (2008) did not state whether their sample was limited to pet-owning youth or if the same instructional procedure present in the 2003 study was implemented. It is reasonable to assume the validity of the measure may be influenced by whether a child is reporting on perceptions of imagined behaviors in contrast to reporting on their actual behavior with a companion animal in their home. Therefore, the reliability and validity of the CTAQ among children with companion animals in the home warrants additional examination.

In a summary of the proceedings of the National Institutes of Health's (NIH) "Health Benefits of Pets" conference in 1987, Esposito et al. (2011) state, "major developmental changes in how children interact with pets are generally parallel to the developmental changes in interaction patterns that children have with familiar humans" (p. 206). Therefore, a second area of concern is that Thompson and Gullone (2003) provided no statement of an intended population for the measure in terms of age range or cultural background. While Thompson and Gullone (2008) note their sampling pool encompassed a range of schools including rural communities, high socioeconomic suburbs, and schools where students from lower socioeconomic statuses were enrolled, developmental, cultural, and socio-demographic variations in children's treatment of animals are also plausible and should be considered.

Similarly, a third limitation of the measure is that several CTAQ items (see Table 1) may not be appropriate for children who have less common types of companion animals (e.g., reptiles or amphibians) as the nature of humane interactions with certain species may manifest in a manner that is qualitatively different than interactions with companion mammals. For example, items on the CTAQ such as "cuddling," "grooming," or "playing dress up" do not reflect socially appropriate humane ways to engage with a pet goldfish or snake; thus, the conceptual definition of humane behavior may not be reflected by all of the CTAQ items unless the intended examinee population is children with mammals such as cats or dogs in the home. Therefore, differences in scale use and item functioning should be examined by companion animal type.

Finally, Thompson and Gullone (2003, 2008) failed to provide a rationale for the selected 3-point response scale (Never = 0, Sometimes = 1, Often = 2). In order for the measure to be administered across different ages and socio-demographic groups, we must be confident that

Table 1. Items on the Children's Treatment of Animals Questionnaire (CTAQ, Thompson and Gullone 2003).

1.	Play with
2.	Give food or water to
3.	Take for a walk
4.	Pat
5.*	Yell at
6.	Cuddle
7.	Cry with when I am sad
8.	Talk to
9.	Allow to stay in my room
10.	Play dress up with
11.	Groom
12.	Tell my secrets to
13.	Spend time with

*Reverse-scored.

children of different ages and race/ethnicities respond to items using the rating scale in the manner intended by the scale developers without semantic variations in interpretation (Linacre 2002, 2010; Chao and Green 2013). Prior research has indicated gender differences related to reliability and test-retest on the CTAQ (Thompson and Gullone 2003, 2008), yet no previous research has investigated use of the rating scale by gender or across cultures. Therefore, it is important to distinguish whether the rating scale functions similarly for both males and females and across racial and ethnic groups.

Current Study

Arbour, Signal and Taylor (2009) suggested that, "Future investigations of HEP would benefit from using a measure like the CTAQ to benchmark normative levels of behavior toward animals in children. This would then enable comparison with children who engage in cruel acts toward animals and an assessment of the efficacy of HEP within this population" (p. 145). However, the CTAQ has not been evaluated using a sample large enough to allow for a rigorous evaluation of factor structure among primary school children. Furthermore, the psychometric properties of the measure have never been examined using data from children who may be at risk for compromised relationships with family companion animals. Given the paucity of measures attending to humane behaviors toward companion animals, the reliability and validity of the CTAQ among children at increased risk of exposure to and perpetration of animal cruelty is an important step in extending what is known about the psychometric properties of the measure.

In the current study, we report findings from a psychometric analysis of the CTAQ (Thompson and Gullone 2003) in a sample of 7- to 12-year-old children exposed to IPV. Specifically, our psychometric investigation was guided by six primary aims. First, we examined the factor structure, reliability, and convergent validity of scores on the CTAQ in a sample of IPV-exposed US children who may be at increased risk for exposure to and perpetration of animal cruelty. To our knowledge, the CTAQ has never been subjected to Rasch analysis or other item response theory analyses. Therefore, the second aim of our study was to extend our knowledge of the psychometric properties of the CTAQ using Rasch analysis techniques in order to develop a detailed understanding of the measure for this population.

Additionally, this study investigated the appropriateness of the measure using a diverse sample. Our third aim was to examine invariance of the measure across socio-demographic variables. Specifically, we assessed differential item functioning across children's race/ethnicity, gender, and age, as well as maternal income level. Our study also offers data on children's exposure to and perpetration of animal cruelty. In support of Arbour and colleagues' (2009) discussion of the CTAQ, we also investigated invariance of the measure (i.e., differential item functioning) by children's animal abuse exposure and animal cruelty perpetration in order to explore whether the CTAQ is an appropriate measure for children in these subgroups. Of notable significance, the CTAQ was also translated into Spanish for our own research purposes; approximately 27% of participating children completed the translated version of the measure. Therefore, our fifth aim was to examine differential item functioning across children who completed the measure in Spanish, as compared with children who completed the measure in English.

Lastly, in our attempt to replicate and extend the work of Thompson and Gullone (2003), our final aim was to examine the relationship between scores on the CTAQ and total scores on the Griffith Empathy Measure (Dadds et al. 2008) and Inventory of Callous/Unemotional (CU) Traits (Frick unpublished), which were also administered to children in our study. Previous investigations of the CTAQ have not tested construct validity of the measure with CU traits. Given the scholarly evidence for the negative relationship between empathy and this interpersonal style (e.g., Blair et al. 2006; Dadds, Whiting, and Hawes 2006; Viding et al. 2009), we hypothesized scores on the CTAQ would be negatively correlated with scores on the ICU, demonstrating convergent validity with this construct.

Methods

Participant Recruitment and Sample Description

Data were collected as part of a larger study examining exposure to IPV and concomitant animal cruelty among women and their 7- to 12-year old children who were receiving IPV services from one of 20 domestic violence agencies in the state of Colorado. Women were eligible to participate in the study if they: (a) reported experiencing IPV within the past year; (b) had at least one child between the ages of 7 and 12 years; and (c) reported having a companion animal within the past year. Consistent with the approved IRB protocol, designated advocates from each agency were trained to recruit participants, obtain consent/assent, and administer survey materials. Mothers and children were interviewed separately by the same staff member. After completion of mother and child interviews, the staff provided mother and child with compensation for their time. Data for the current analysis reflect 217 mother-child dyads. The sample of children was 46% female and 54% male, with a mean age of 9.30 years ($SD = 1.66$), and were identified racially/ethnically by mothers as follows: 54.4% Hispanic/Latino, 22.6% White, 17.1% more than one race, 3.7% Black, 1.8% American Indian or Alaska Native, and 0.5% Asian American. Seventy-one percent of mothers reported having a dog in the home, 34% indicated a cat, and 25% of mothers reported other companion animals (e.g., fish, rabbits, lizards), with some participants indicating multiple companion animals in the home. This information was obtained via maternal report on a demographic survey.

Measures

The Children's Treatment of Animals Questionnaire (CTAQ; Thompson and Gullone 2003): As stated previously, the CTAQ is a 13-item self-report questionnaire for youth and is designed to assess children's humane treatment of animals. Each item is rated on a 3-point scale, ranging from

“never” to “often.” For the purposes of our study, the scale was also adapted in the Spanish language using the translation–back-translation procedure (van de Vijver and Hambleton 1996). Twenty-seven percent of children in our sample completed the Spanish version of the measure.

Pet Treatment Survey (PTS; Ascione unpublished): Children's exposure to animal abuse was assessed using a revised version of the Battered Partner Treatment Survey—Pet Maltreatment Assessment (BPSS; Ascione and Weber 1995). The BPSS, a structured interview used in two separate studies examining animal abuse in the context of IPV (Ascione et al. 2007; Volant et al. 2008), was revised purposefully for our sample and administered to mothers. In addition to asking respondents about past pet ownership history, two items specifically address animal abuse. In cases where women report harm to animals by a partner, a woman is then asked a series of follow-up questions including whether her child has seen or heard pets hurt in the home and if the children have ever hurt or killed animals in the home. These items from the PTS were used as dichotomous measures of children's exposure to and perpetration of animal abuse.

Inventory of Callous and Unemotional Traits—Caregiver Report Form (ICU; Frick unpublished): The ICU is a parent-report instrument that measures the presence and intensity of callous-unemotional traits among children (Essau, Sasagawa and Frick 2006). This 24-item measure is scored on a 4-point scale (ranging from 0 = not at all true to 3 = definitely true). Previous research indicates internal consistency of the total scale is adequate (Cronbach's $\alpha = 0.74$ to 0.85). Item-analysis for the current sample yielded an alpha coefficient of 0.89. Prior research has also reported evidence of construct validity of the ICU total score with measures of aggression, delinquency, and empathy (e.g., Kimonis et al. 2008; Roose et al. 2010).

Griffith Empathy Measure (GEM; Dadds et al. 2008): The GEM is a refinement of a parent-report measure of empathy used in our previous research (Ascione 1992; Ascione and Weber 1996). The instrument comprises 23 items scored from “strongly disagree” (−4) to “strongly agree” (+4). Previous research has advocated for the utility of the measure in capturing an overall dimension of empathy among children ages 4 to 18 years (Dadds et al. 2008); for the current sample, reliability of the total scale was 0.81.

Analysis Strategy

Item-Analysis and EFA: In a replication of previous psychometric evaluations of the CTAQ, classical test theory was used to conduct item analysis and exploratory factor analysis (EFA) in order to examine scale functioning. Specifically, principal components analysis with varimax rotation was employed to investigate construct validity. Parallel analysis, which simulates a normal, random sample and uses the correlation matrix to explore the expected eigenvalues compared with the sample's eigenvalues, was also conducted in SPSS 21 (Chicago, IL, USA) to determine which factor(s) to retain.

Rasch Analysis: Rasch (1980) analysis was conducted to examine the structure of the CTAQ using the software Winsteps (Linacre 2012). This technique is supported by a sizeable growing literature among fields such as psychology and other social sciences (e.g., Fischer and Molenaar 1995; Wright and Stone 2004; Bond and Fox 2007). The underlying theoretical assumption of the Rasch model suggests that a unidimensional construct is necessary for useful measurement. Therefore, this technique allows for the evaluation of the extent to which a unidimensional scale is generated by the items on the CTAQ. Generally, Rasch analysis allows for the identification of the extent to which items on the measure have reliably measured a single

construct (Linacre 2002; Bond and Fox 2007). If data for the sample fit the Rasch model, item and person estimates can be interpreted in terms of equivalent interval units generated from natural log transformations of raw scores (Bond and Fox 2007; Linacre 2012).

To determine whether each item meaningfully contributes to the measurement of a one-dimensional construct, Rasch fit indices are used to assess the extent to which an item or person performs as expected. In addition, a principal components analysis of residuals (PCAR) is used to determine the presence of a second factor. The replicability of item placements and person ordering is also provided through estimates of item and person reliability indices. In addition, Rasch analysis produces person separation, which appraises the ability of the items to identify different levels of the measure on a less-to-more continuum; subgroups of persons discriminated by the instrument are also identified (Green and Frantom 2002; Bond and Fox 2007). Finally, another benefit of Rasch analysis is the ability to identify gaps in the construct continuum by distinguishing persons and items that are not well targeted by the measure (Chao and Green 2013). Here, an item that is poorly targeted to the measure means that there is an inadequate number of individuals at an ability level comparable to the item's difficulty, thereby preventing adequate assessment of the item's difficulty (Linacre 2012).

Convergent Validity: Pearson correlations were calculated between scores on the CTAQ and both the GEM and ICU. In order to replicate previous convergent validity analyses conducted by Thompson and Gullone (2003), correlations were also computed separately for each gender.

Results

Item-Analysis and EFA

The reliability coefficient for Cronbach's alpha was 0.84. Except for item 5, each item contributed to the scale. Results indicated deletion of item 5 would raise Cronbach's alpha to 0.86. Since the increase in the value of alpha was minimal, item 5 was retained in the analysis. Principle components analysis (PCA) was conducted to investigate if the measure was characterized by multiple factors. We used parallel analysis to identify the number of factors to retain. When conducting parallel analysis, factors considered to be useful are the ones whose eigenvalues are larger in the sample compared with randomly generated data (Horn 1965). In our sample, the eigenvalue for only one factor was larger than the eigenvalue for a random sample and so one factor was retained for the CTAQ.

Rasch Analysis

Dimensionality: One of the major assumptions of the Rasch model is unidimensionality. Linacre's (2012) suggestion for examining dimensionality is for the principle components analysis of residuals to explain 40% or more of the total raw variance, with the first contrast having an eigenvalue of less than 2.0 with less than 5% unexplained variance. For this sample, the measure explained 49.8% of the variance with the unexplained variance in the first contrast having an eigenvalue of 1.8 with 7.4% unexplained variance. The higher than desired unexplained variance is not unexpected with such a short measure. Therefore, scores on the CTAQ met the expectations set by Linacre for unidimensionality.

Overall fit was examined as an index of the CTAQ's fit to a unidimensional model. Specifically, mean square (MNSQ) infit and outfit statistics were examined. As recommended by Linacre (2007, 2010), these values should be close to 1.0. Based on these standards, our data fit the model very well with an average infit MNSQ of 1.03 ($SD = 0.52$) and an average

Table 2. Rasch model summary statistics.

	Item	Person
Number Measured	12	214
Location Mean (SD)	0.00 (1.20)	0.59 (1.27)
Infit MNSQ Mean (SD)	1.01 (0.17)	1.03 (0.52)
Outfit MNSQ Mean (SD)	1.01 (0.27)	1.02 (0.65)
Reliability	0.99	0.82
Separation	8.92	2.15

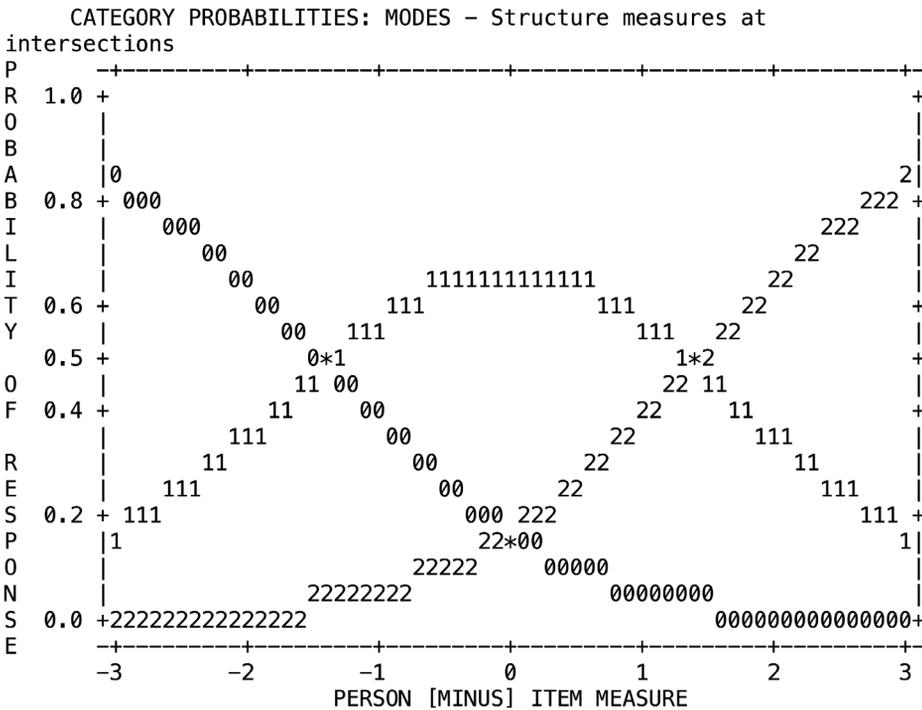


Figure 1. Step use by person-Item measure: Children's Treatment of Animals Questionnaire (CTAQ).

outfit MNSQ of 1.02 ($SD = 0.65$). These values indicate that there was little overall unexpected behavior. Consistent with results from the PCAR, these values lend support that the CTAQ is a unidimensional model.

Item and Person Fit: Item fit was examined to ensure that each item fit the Rasch model. According to Wright and Linacre (1994), the infit MNSQ values for rating a scale such as the CTAQ should be between 0.6 and 1.4. The MNSQ for items in our sample ranged from 0.79 to 1.70. Item 5 had an infit MNSQ of 1.70 and a point-measure correlation of 0.0. Therefore, the item was deleted; dimensionality results reported above reflect the final model with item 5 removed. Person fit was also examined to ensure that individuals were answering in a consistent, expected manner. There were no participants that misfit this measure with MNSQ infit values over 4.0.

Table 3. Step structure.

Category	Count	%	Observed Average	Sample Expected	Infit MNSQ	Outfit MNSQ	Structure Calibration	Category Measure
1	7	3	-1.64	-1.28	0.80	0.75	None	-4.10
2	66	31	-0.08	-0.12	1.01	1.07	-1.39	-1.56
3	141	66	1.07	1.07	0.96	0.96	1.39	0.98

Reliability: Reliability is measured by computing person and item spread across the measure. A separation of 2.0 is minimal with higher levels of separation indicating a larger range of items and persons. Person separation was 2.15, with reliability of person separation of 0.82, and a Cronbach's alpha of 0.86 (Table 2). The correlation coefficient between raw total scores and Rasch measure estimates for persons was calculated and, as anticipated, was high ($r = 0.93$, $p < 0.001$). Consistent with recommendations established by Gauggel et al. (2004), these values indicate the CTAQ is a reliable measure for this sample of children.

Scale Use: Results of the Rasch analysis indicated that participants in our study used the rating scale as intended. Category probability curves (Figure 1) demonstrated an even distribution of the three categories with clearly advancing steps. Rasch-Andrich thresholds increased with category values with no evidence of step misfit (Table 3). In particular, participants used the highest category more frequently, suggesting our sample was characterized by high levels of humane treatment of animals. This is demonstrated in the item-person map provided in Figure 2.

Targeting and Construct Coverage: The item-person map provided in Figure 2 presents items and persons on the same scale and demonstrates scale functioning for this sample. Participants represented near the top of the left-hand side of the item-person map are children who engage in higher levels of humane treatment of animals; participants represented near the bottom are children who score lower on humane treatment. Participants are spread fairly evenly throughout the item-person map, although the majority of the population appears near the top. Representation of items and participants along the map suggest this sample of children report frequent humane behaviors toward their companion animal. Approximately 5% of the children in our sample were characterized by extreme high scores on the CTAQ; however, this proportion of extreme scores is unlikely to be a threat to our assumption of unidimensionality (Christensen, Engelhard and Salzberger 2012).

Rasch analysis allows researchers to see how items act on a continuum of “easiest” to agree with to “hardest” to agree with, thus creating a kind of “ruler” of construct coverage. The items at the top of Figure 2 are questions that were hardest for participants to agree with, while those at the bottom were the easiest to agree with. The distribution of items, seen on the right-hand side of Figure 2, had item logit values between -3 and 4 , indicating the CTAQ is a good measure along this range (Linacre 2010). Examination of the map indicates two areas where the measure could be improved. Specifically, items 1 (“play with”) and 2 (“give food/water to”) are equal on the “ruler,” suggesting that only one of these two items is needed and the second is redundant as it is a measure of the construct at the same location; item 7 (“cry with when I am sad”) and item 12 (“tell my secrets to”) had equivalent item positions as well.

Invariance: Differential item functioning (DIF) was examined to ensure that the items were functioning in the same way across demographic groups. DIF was assessed using t -tests of the significance of differences in item logit position. At a significance level of 0.01, there was no DIF

type of companion animal (e.g., bird, lizard, and fish). The measure was also found to be invariant by companion animal type. Finally, no DIF was found when comparing children exposed to animal cruelty with those whose mothers reported no exposure. This finding was also consistent when comparing children who had perpetrated animal cruelty with those who had not engaged in animal cruelty. Therefore, the measure functioned as intended across the examined groups.

As previously mentioned, in our study the CTAQ was translated into Spanish for 58 participants. In order to verify that this scale is appropriate for Spanish-speaking children, we considered DIF for the English version and the Spanish version. Results indicated all items functioned equivalently in both versions with the exception of item 3, “take for a walk.” To examine whether this difference reflected an issue with the translation of the item and/or a culturally based variation, we also explored DIF for children who were identified by mothers as Hispanic or Latino (67%) compared with children who were not identified as Hispanic/Latino (27%). In the same manner, results evidenced significant DIF for item 3 among children of mothers from these two groups.

Convergent Validity

Evidence of convergent validity was not found between total scores on the 12-item version of the CTAQ and the GEM ($r = 0.13, p = 0.06$). This finding was consistent when we explored the relationship for boys and girls separately. A small but statistically significant correlation was found between scores on the CTAQ and ICU ($r = -0.26, p < 0.001$). As predicted, higher scores on callous/unemotional traits were related to less humane treatment of animals. When this analysis was conducted separately for boys and girls, only the relationship between boys' scores on the ICU and CTAQ remained significant ($r = -0.33, p < 0.001$).

Discussion and Implications

To our knowledge, our study is the first evaluation of the CTAQ in a US sample of children at-risk for compromised socio-emotional functioning. Results of the item analysis, EFA, and parallel analysis suggest the CTAQ demonstrates adequate reliability as a one-factor measure of children's humane treatment of animals. Results of the Rasch analysis, however, indicate that the CTAQ is an appropriate unidimensional measure of children's humane treatment of animals after removal of item 5 (“yell at”). Contrary to previous findings, our study did not support convergent validity of the CTAQ with our measure of empathy. Our results did suggest, however, convergent validity of the measure with callous and unemotional traits on the ICU; in particular, when this analysis was conducted separately by gender, convergent validity was only present for boys in our sample, with a small but significant correlation between higher levels of CU traits with lower scores on the CTAQ.

In light of our results, we recommend use of a 12-item version of the CTAQ that excludes item 5. Overall, our findings demonstrate the measure is a strong unidimensional assessment of children's humane treatment of animals and is well-suited for ages 7 to 12 years. Our results also suggest the measure exhibits invariance across gender, type of companion animal, and income level. Notably, no significant differences in item functioning were observed between children with or without exposure to animal abuse, as well as between children who had or had not perpetrated animal cruelty in the past.

Consistent with Thompson and Gullone's (2003) recommendations, research is still needed to determine the predictive validity of the CTAQ as well as its sensitivity to measuring change.

Future evaluations of the CTAQ would also benefit from investigating the level of agreement between children's reports on the CTAQ and observer-rated or parent-reported behaviors toward companion animals (Thompson and Gullone 2003). In addition, our findings of the measure's construct coverage suggest future revisions to the CTAQ should consider examining overlap between items 7 and 12 and 1 and 2 to determine if items could be deleted or revised to expand the range of the questionnaire. Similarly, given that nearly 5% of our sample received maximum scores, future revisions should incorporate items that are harder to endorse, in order to improve the scope and sensitivity of the measure and better discriminate individuals with high levels of humane treatment of animals.

An important finding of our study surrounds the differential item functioning of item 3 ("take for a walk") when comparing the original measure with the Spanish translation. This difference was also present when comparing children who were identified by mothers as Hispanic or Latino with children who were not identified as Hispanic/Latino. These findings suggest that our Spanish translation of the measure needs to be verified and/or modified for this item in a culturally considerate context if the translated version is to be implemented with Spanish-speaking and/or Latino/Hispanic children in future studies. Additionally, these results suggest future research is needed to determine the cultural significance of walking companion animals in the context of Latino and/or Hispanic culture.

An important consideration in human–animal interactions research is how culturally influenced attitudes toward animals shape relationships with animals. Scholars such as Kaufmann (1999) have argued that the ability of research in this area to guide effective research, prevention, and intervention rests on attention to the cultural context of what is considered to be socially acceptable interactions with nonhuman animals. Rigorous empirical studies of the developmental implications of child–animal interactions are scarce and plagued by methodological limitations inherent in human–animal interactions research. Our findings demonstrate that the application of Rasch techniques can add increased rigor to the study of human–animal interactions by extending what is known about the psychometric properties of existing and new measures, and providing detailed information about instruments when administered across various cultural and socio-demographic groups.

Acknowledgements

Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under award number R01HD066503. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This paper was presented in part at the 2014 annual meeting of the Society for Social Work and Research in San Antonio, Texas. The authors thank the community-based domestic violence advocates for their contribution to this work.

References

- Arbour, R., Signal, T. and Taylor, N. 2009. Teaching kindness: The promise of humane education. *Society & Animals* 17: 136–148.
- Ascione, F. R. 1992. Enhancing children's attitudes about the humane treatment of animals: Generalization to human-directed empathy. *Anthrozoös* 5: 176–191.
- Ascione, F. R. 1998. Battered women's reports of their partners' and their children's cruelty to animals. *Journal of Emotional Abuse* 1: 119–133.

- Ascione, F. R. ed. 2008. *International Handbook of Animal Abuse and Cruelty: Theory, Research, and Application*. West Lafayette, IN: Purdue University Press.
- Ascione, F. R. and Weber, C. V. 1995. Battered Partner Shelter Survey (BPSS). Logan: Utah State University.
- Ascione, F. R. and Weber, C. V. 1996. Children's attitudes about the humane treatment of animals and empathy: One-year follow up of a school-based intervention. *Anthrozoös* 9: 188–195.
- Ascione, F. R., Weber, C. V., Thompson, T., Maruyama, M. and Hayashi, K. 2007. Battered pets and domestic violence: Animal abuse reported by women experiencing intimate partner violence and by non-abused women. *Violence Against Women* 13: 354–373.
- Blair, K. S., Richell, R. A., Mitchell, D. G., Leonard, A., Morton, J. and Blair, R. J. 2006. They know the words, but not the music: Affective and semantic priming in individuals with psychopathy. *Biological Psychology* 73: 114–123.
- Bond, T. G. and Fox, C. M. 2007. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah: Erlbaum.
- Bryant, B. 1982. An index of empathy for children and adolescents. *Child Development* 53: 413–425.
- Chao, R. and Green, K. E. 2013. Rasch analysis of the Outcome Questionnaire with African Americans. *Psychological Assessment* 25(2): 568–582.
- Christensen, K.B., Engelhard Jr., G., and Salzberger, T. 2012. Ask the experts: Rasch vs. factor analysis. *Rasch Measurement Transactions* 26: 1373–1378.
- Currie, C. L. 2006. Animal cruelty by children exposed to domestic violence. *Child Abuse and Neglect* 30: 425–435.
- Dadds, M. R., Hunter, K., Hawes, D. J., Frost, A. D. J., Vassallo, S., Bunn, P., Merz, S. and El Masry, Y. 2008. A measure of cognitive and affective empathy in children using parent ratings. *Child Psychiatry and Human Development* 39: 111–112.
- Dadds, M. R., Whiting, C., Bunn, P., Fraser, J. A., Charlson, J. H. and Pirola-Merlo, A. 2004. Measurement of cruelty in children: The Cruelty to Animals Inventory. *Journal of Abnormal Child Psychology* 32: 321–334.
- Dadds, M. R., Whiting, C. and Hawes, D. J. 2006. Associations among cruelty to animals, family conflict, and psychopathic traits in childhood. *Journal of Interpersonal Violence* 21: 411–429.
- Eposito, L., McCune, S., Griffin, J. and Maholmes, V. 2011. Directions in human–animal interaction research: Child development, health and therapeutic interventions. *Child Development Perspectives* 5(3): 205–211.
- Essau, C. A., Sasagawa, S. and Frick, P. J. 2006. Callous–unemotional traits in a community sample of adolescents. *Assessment* 13: 454–469.
- Fischer, G. H. and Molenaar, I. W. 1995. *Rasch Models: Foundations, Recent Developments, and Applications*. New York: Springer-Verlag.
- Guggel, S., Böcker, M., Heinemann, A. W., Lämmle, G., Borchelt, M. and Steinhagen-Thiessen, E. 2004. Patient–staff agreement on Barthel Index scores at admission and discharge in a sample of elderly stroke patients. *Rehabilitation Psychology* 49: 21–27.
- Green, K. E. and Frantom, C. G. 2002. Survey development and validation with the Rasch model. Paper presented at the 2002 International Conference on Questionnaire Development, Evaluation, and Testing, Charleston, USA, November 14–17, 2002.
- Gresham, F. M. and Elliot, S. N. 1990. *Social Skills Rating System Manual*. Circle Pines, MN: American Guidance Service.
- Guadagnoli, E. and Velicer, W. F. 1988. Relation of sample size to the stability of component patterns. *Psychological Bulletin* 103: 265–275.
- Gullone, E. and Robertson, N. 2008. The relationship between bullying and animal abuse in adolescents: The importance of witnessing animal abuse. *Journal of Applied Developmental Psychology* 29: 371–379.
- Henry, B. C. 2004. The relationship between animal cruelty, delinquency, and attitudes toward the treatment of animals. *Society & Animals* 12: 185–207.
- Hinkin, T. R. 1995. A review of scale development practices in the study of organizations. *Journal of Management* 21(5): 967–988.
- Horn, J. L. 1965. A rationale and test for the number of factors in factor analysis. *Psychometrika* 30(2): 179–185.
- Kaufmann, M. E. 1999. The relevance of cultural competence to the link between violence to animals and people. In *Child Abuse, Domestic Violence, and Animal Abuse: Linking the Circles of Compassion for Prevention and Intervention*, 260–270, ed. F. Ascione and P. Arkow. West Lafayette, IN: Purdue University Press.

- Kimonis, E. R., Frick, P. J., Skeem, J. L., Marsee, M. A., Cruise, K., Munoz, L. C., Aucoin, K. J. and Morris, A. S. 2008. Assessing callous-unemotional traits in adolescent offenders: Validation of the Inventory of Callous-Unemotional Traits. *International Journal of Law and Psychiatry* 31: 241–252.
- Linacre, J. M. 2002. Understanding Rasch measurement: Optimizing rating scale category effectiveness. *Journal of Applied Measurement* 3: 85–106.
- Linacre, J. M. 2007. *A User's Guide to Winsteps/Ministep: Rasch-Model Computer Programs*. Chicago: Winsteps.
- Linacre, J. M. 2010. *A User's Guide to Winsteps Rasch Model Computer Program: Program Manual 3.70*. Chicago: Winsteps.
- Linacre, J. M. 2012. *A User's Guide to Winsteps: Program Manual 3.75.0*. Chicago: MESA Press.
- Maiuro, R. D., Eberle, J. A., Rastaman, P. and Snowflake, B. 2008. Pet abuse: Relationships to psychobiology, attachment processes, and domestic violence. In *The International Handbook of Animal Abuse and Cruelty: Theory, Research, and Application*, 133–153, ed. F. Ascione. West Lafayette, IN: Purdue University Press.
- Mellor, D. J., Patterson-Kane, E. and Stafford, K. J. 2009. *Sciences of Animal Welfare*. Oxford: Wiley-Blackwell.
- Rasch, G. 1980. *Probabilistic Models for Some Intelligence and Attainment Tests*. Chicago: University of Chicago.
- Roose, A., Bijttebier, P., Decoene, S., Claes, L. and Frick, P. J. 2010. Assessing the affective features of psychopathy in adolescence: A further validation of the Inventory of Callous and Unemotional Traits. *Assessment* 17: 44–57.
- Schwab, D. P. 1980. Construct validity in organizational behavior. In *Research in Organizational Behavior*, 3–43, ed. B. M. Staw and L. L. Cummings. Greenwich, CT: JAI Press.
- Simmons, C. and Lehmann, P. 2007. Exploring the link between pet abuse and controlling behaviors in violent relationships. *Journal of Interpersonal Violence* 22: 1211–1222.
- Thompson, K. L. and Gullone, E. 2003. The Children's Treatment of Animals Questionnaire (CTAQ): A psychometric investigation. *Society & Animals* 11: 1–15.
- Thompson, K. L. and Gullone, E. 2006. An investigation into the association between the witnessing of animal abuse and adolescents' behavior toward animals. *Society & Animals* 14: 223–243.
- Thompson, K. L. and Gullone, E. 2008. Prosocial and antisocial behaviors in adolescents: An investigation into associations with attachment and empathy. *Anthrozoös* 21: 123–137.
- Viding, E., Simmonds, E., Petrides, K. V. and Frederickson, N. 2009. The contribution of callous-unemotional traits and conduct problems to bullying in early adolescence. *Journal of Child Psychology and Psychiatry* 50(4): 471–481.
- van de Vijver, F. J. R. and Hambleton, R. K. 1996. Translating tests: Some practical guidelines. *European Psychologist* 1: 89–99.
- Volant, A. M., Johnson, J. A., Gullone, E. and Coleman, G. 2008. The relationship between domestic violence and animal abuse: An Australian study. *Journal of Interpersonal Violence* 23: 1277–1295.
- Walton-Moss, B. J., Manganello, J., Frye, J. and Campbell, J. C. 2005. Risk factors for intimate partner violence and associated injury among urban women. *Journal of Community Health* 30: 377–389.
- Wilson, C. C. and Barker, S. B. 2003. Challenges in designing human animal interaction research. *American Behavioral Scientist* 47: 16–28.
- Wright, B. D. and Linacre, J. M. 1994. Reasonable mean-square fit values. *Rasch Measurement Transactions* 8(3): 370.
- Wright, B. D. and Stone, M. H. 2004. *Making Measures*. Chicago: Phaneron Press.