Toward Effective Evaluation of Environmental Education: Validity of the Children's Environmental Attitudes and Knowledge Scale Using Data From a Sample of Irish Adolescents

Article in The Journal of Environmental Education · December 2006
DOI: 10.3200/JOEE.37.2.13-23

CITATIONS
35

READS
228

2 authors, including:

Malcolm Maclachlan
National University of Ireland, Maynooth

Some of the authors of this publication are also working on these related projects:

GATE related research View project

World Health Organization GATE Initiative View project
ABSTRACT: Evaluation of environmental education (EE) interventions is at an early stage of development. Within a quantitative paradigm, the authors examined the psychometric properties of the Children’s Environmental Attitudes and Knowledge Scale (CHEAKS) with data from a sample of Irish adolescents (N = 338) participating in a repeated measures design. Results lend support to published findings that recommend the instrument as reliable and valid. Comparative analysis between these data and published figures from the United States indicates differences in the constituent elements tapped by the scale. These differences are explored and recommendations for further study are made.

KEY WORDS: Children’s Environmental Attitudes and Knowledge Scale (CHEAKS), cultural comparison, ecological views, environmental education (EE)

Toward Effective Evaluation of Environmental Education: Validity of the Children’s Environmental Attitudes and Knowledge Scale Using Data From a Sample of Irish Adolescents

Anne Walsh-Daneshmandi and Malcolm MacLachlan
However, the idea that environmentally friendly behaviors can be increased through the medium of education is debated widely, with reports of mixed effects across a broad range of research designs (for a more detailed treatment, see, for example, Bell, Greene, Fisher, & Baum, 2001; Gifford, 2002). The question of evaluation of programs is increasingly important as EE reaches a level of maturity wherein allocation of resources is becoming more focused. Nonetheless, to date, there has been only modest development in the evaluation of EE. For instance, a search for the term “EE evaluation” yielded considerably fewer hits (128 and 88, respectively, on the two search engines). To contextualize this equation, it is useful to contrast the ratio of references for EE and EE evaluation (13,000:1 and 20,000:1 on the respective search engines) with hits for “education” and “education evaluation” (2:1 and 20:1, respectively). However, evaluation is a complex issue that brings to bear considerations of design and context.

In response to the lack of scales with robust psychometric properties (Gray, Borden, & Weigel, 1985) and in an effort to produce a scale to assess the effectiveness of environmental education, researchers developed the Children’s Environmental Attitude and Knowledge Scale (CHEAKS; Leeming, Dwyer, & Bracken, 1995) was developed. In this initial work, Leeming et al. administered the scale to a large sample of students (in the Memphis, Tennessee, metropolitan area) during the fall of 1991 (n = 1,219) and children (in the same classes) in the spring of 1992 (n = 1,241), with 1,041 students completing the questionnaire both times. On both occasions, 50 classes of students in 12 different schools participated in the study. The children ranged in age from 6 to 13 years (approximating Grade 1 students at 6 years old). Although Leeming et al. also classified students with reference to grade level, given the variation in school systems between the United States and Ireland, we confine our remarks to age only when comparing data across the Dublin and Memphis samples. Leeming et al. reported that the scale has impressive psychometric properties and appears to fill an important gap in evaluation of environmental education.

However, the validity of the measure has not been reported for students aged in their mid- to late-teens, a time when educational curricula in Ireland include instruction in EE. Furthermore, no data are available for samples outside the United States. We identified CHEAKS as having potential for developing an EE evaluation tool within the Irish context. In the present study, we analyzed the properties of the scale when administered to a sample of Irish teenagers. We also compared findings from this study with those of Leeming et al. (1995) to demonstrate the age-related changes in environmental attitudes and knowledge.

Gray et al. (1985) recommended that research techniques should “benefit from the prior measurement efforts of others” (p. 30); therefore, Maloney, Ward, and Braucht (1975) modeled CHEAKS on the structure and content of Maloney et al. highly acclaimed adult scale, which was intended to measure multiple attitudinal components. CHEAKS, in addition to assessing ecological knowledge, is constructed to encompass items that relate to major attitudinal components of commitment (verbal and actual), behavior, and affect.

Whereas CHEAKS was found to be highly reliable and valid for children up to Grade 7 (12- and 13-year-olds), the question of how reliable the scale would be with older children remains unexplored. In this present work, we present data from an older group of students aged 15–17 years living in a geographic location different from the participants in the original study (Leeming et al., 1995).

**Method**

**Participants**

In this study, we based the analysis on protocols from 338 students (218 boys and 120 girls), aged between 15 and 17 years (M = 16 years). These students attended transition year at 12 non-fee–paying
schools in the greater region of the capital city of Dublin. We also compared published figures from the Leeming et al. (1995) study and the original data from our study.

**Materials**

CHEAKS has been described in detail elsewhere (Leeming et al., 1995; Walsh-Daneshmandi, 2002), but for clarification it is useful to recall that the instrument measures, in addition to a global construct of environmentalism, two separate constructs: attitude and knowledge. The attitude subscale comprises 36 items that measure students’ attitudes toward environmental issues, reflecting verbal commitment, actual commitment, and affect (12 items for each dimension). The knowledge subscale comprises 30 items. The total scale score is derived from a summation of the subscale scores.

The 36-item attitude subscale items are presented in a 5-point Likert-type response format (ranging from 1 = *very false* to 5 = *very true*). Nine of the attitude items, 3 in each section, are negatively connotated and reverse scored to reduce the likelihood of response set. Therefore, the most proenvironmental response to each item is scored 5, whereas the least environmental response is scored 1. Possible scores on the attitude subscale range between 36 and 180, inclusive. For the 30 knowledge questions, correct responses are scored 6; therefore, possible scores range from 0 to 180. Possible total scores on CHEAKS range from 36 to 360, with higher scores indicating combined positive attitudes and increased knowledge.

This 66-item scale underwent minor alterations to adjust for cultural sensitivities after consultation with appropriate educational specialists. For example, “rubbish” replaced the more Americanized term “trash” (Item 6, verbal commitment subscale), “water tap” replaced “faucet” (Item 9, actual commitment subscale), and “Signets” replaced “The Boy Scouts of America” (Item 30, knowledge subscale). In addition, “perpetual” was replaced with “endless” (Item 12, knowledge subscale) following pilot testing.

Within the published guidelines, we computed subscale totals to measure attitudes (the sum of affect, verbal commitment, and actual commitment) and knowledge.

**Procedures**

In the present study, we administered CHEAKS as part of a larger protocol, titled the *Person and Environment Health Questionnaire* (Walsh-Daneshmandi, 2002). The same researcher systematically administered the protocols during regular school hours to groups of intact classes in the presence of a teacher. At the beginning of each session, the teacher introduced the researcher to the students. The researcher then introduced the survey using standardized statements. The duration of the survey was approximately 25 min. Participants completed CHEAKS on two occasions in the same year, between January and February on the first occasion and between March and May on the second. Following the second administration, an extra class period was allocated for feedback and debriefing. We used standard institutional procedures to receive ethical approval.

**Analyses**

To consolidate previous findings, we examined two forms of reliability (i.e., stability and internal consistency) and several forms of validity, including content, convergent and discriminant, contrasted groups, and factor analysis. In addition, we compared group scores across the Dublin and Leeming et al. (1995) studies to explore age-related differences.

**Results**

The means and standard deviations for both administrations of the attitude and knowledge subscales and total CHEAKS scores for both the Dublin and the Leeming et al. (1995) samples are pre-
presented in Table 1. To enhance ease of comparability across samples, we have approximated ages on 
the basis that Grade 1 students are 6 years old.

Trends in the data indicate that the total scores for the Dublin sample (15-, 16-, and 17-year-olds) 
were lower than those reported for the younger students in the Leeming et al. (1995) study. At the 
subscale level of analysis, a similar trend is apparent for the attitude scores but not for the knowledge 
scores. Scores for the knowledge scale indicate higher scores for the Dublin sample than either of the 
two Memphis groups (Leeming et al.).

When we employed analysis of variance (ANOVA) to compare the means of the present sample with 
those using summary data given in Leeming et al. (1995, Table 1), there was consistent evidence of sta-
tistically significant differences across group scores. We found significant differences on the total scores 
for the first administration, $F(2, 1525) = 122.81$, $p < .01$, but not for the second, $F(2, 1574) = 10.1$, 
$p > .01$; attitude scores for both the first, $F(2, 1553) = 327.83$, $p < .01$, and second administrations, 
$F(2, 1577) = 591.12$, $p < .01$; and knowledge scores for both the first, $F(2, 1529) = 196.21$, $p < .01$, 
and second administrations, $F(2, 1574) = 151.18$, $p < .01$.

**Stability**

To assess stability, we first administered CHEAKS during January and February and then between 
March and May of the same year. We calculated Pearson product–moment correlations for these pre-
and posttest CHEAKS subscale and total scale scores. 

Table 2 shows the correlations between total scale scores and attitude and knowledge subscale 
scores for the two administrations. The values for the test–retest correlations all reached statistical sig-
nificance ($p < .001$) and were substantial enough to support earlier claims for the stability of the envi-
ronmental attitudes and knowledge constructs.

**Convergent and Discriminant Validity**

The central tenet of this form of validity lies in the expectation that scales assessing different con-
structs produce lower correlations than do scales assessing similar constructs. We examined this 
aspect of validity by computing intercorrelations between the CHEAKS subscales and comparing the 
first and second administrations of the attitude and knowledge subscales. We predicted that correla-
tion coefficients for the pre- and posttest scores would be larger than those for the attitude and 
knowledge subscales. If low to moderate correlations exist between the subscales, we would interpret 
this as an indication that they measure independent constructs.

When we examined the correlations of the attitude and knowledge subscales (Table 2), we found 
values considerably lower than the test–retest correlations. As expected, these results support the pres-
ence of independent constructs within the two subscales.

**Internal Consistency**

We assessed this form of reliability by computing alpha coefficients for each of the CHEAKS sub-
scapes and total score for each administration. Alpha coefficients for the CHEAKS total score and the 
attitude and knowledge subscales were consistently high (ranging from .77 to .92) across adminis-
trations (Table 3). We observed the lowest values for the knowledge subscale at the first administra-
tion; however, improvement occurred at the time of the second administration. These analyses pro-
vide evidence of internal consistency on the CHEAKS scales that support its use for research 
(Bryman & Cramer, 1997) and screening purposes (Salvia & Ysseldyke, 1988).

When the ability of the CHEAKS to discriminate levels of knowledge was examined, via mean 
 scores, we saw that the respondents answered approximately 50% of the items correctly. Given the


<table>
<thead>
<tr>
<th>Scale</th>
<th>Current study</th>
<th>Leeming et al. (1995)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15–17 years old</td>
<td>10–13 years old</td>
</tr>
<tr>
<td>Time 1</td>
<td>Time 2</td>
<td>Time 1</td>
</tr>
<tr>
<td>M</td>
<td>SD</td>
<td>n</td>
</tr>
<tr>
<td>Attitude</td>
<td>100.88</td>
<td>25.73</td>
</tr>
<tr>
<td>Knowledge</td>
<td>93.92</td>
<td>31.17</td>
</tr>
<tr>
<td>Total</td>
<td>194.92</td>
<td>46.11</td>
</tr>
</tbody>
</table>

Note. Scores on the attitude subscale ranged from 36 to 180, the knowledge subscale from 0 to 180, and the total scale from 36 to 360.
chance expectation of 20% on this 5-option response format, this finding indicates responses that are more than twice the level expected by chance but reflects only half of the potential knowledge among this group of students. This is salient given the discrimination aim of the scale without the danger of a ceiling effect occurring. When we examined the distribution of mean scores, we saw that at least some of the students employed guessing as a strategy; however, this was not the general trend as alpha coefficients approach zero to the degree that responses reflect random guessing (Leeming et al., 1995). Given the consistently high reliabilities, we can surmise that random guessing was not a strategy generally adopted by the participants in the study.

Contrasted-Groups Validity

Given the nature of the study (students were in their mid-teens and the level of teacher involvement was minimal), we asked participants to rate themselves on a “Green-Self” rating (a combination of their knowledge, attitudes, and actions), on a scale of 0–10, with a score of 10 indicating very strong commitment to the environment. We then compared these data with scale scores as a basis for the contrasted-groups validity check. These Green-Self scores were recoded into above and below median scores.
to represent categories of “low environmentalists” and “high environmentalists” (Table 4). As we can see from Table 4, those students who rated themselves more environmentally committed scored consistently higher on both the attitude and knowledge subscales as well as the total scale of CHEAKS.

To explore the substance of this difference, we conducted multivariate analysis of variance with environmental commitment as the independent variable. This analysis indicated that the trends we saw in the mean scores were substantive.

A multivariate difference emerged, Wilks’s lambda = .922, F(2, 275) = 11.628, p < .001, and at a univariate level significant mean differences emerged on the total scale score, F(1, 276) = 18.32, p < .001, the attitude score, F(1, 276) = 22.08, p < .001, and the knowledge score, F(1, 276) = 5.412, p < .05.

**Factor Analysis**

Factor analysis was conducted in line with the criteria published by Leeming et al. (1995). That is, we applied principal component analysis with orthogonal rotation to the data. The scree analysis (Figure 1) indicated the utility of a two-factor solution for this dataset. The subsequent two-factor

**TABLE 4. Means and Standard Deviations for the Children’s Environmental Attitudes and Knowledge Scale Scores, by Level of Environmentalism**

<table>
<thead>
<tr>
<th>Scale</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Total</td>
<td>190.50</td>
<td>43.75</td>
</tr>
<tr>
<td>Attitude</td>
<td>96.94</td>
<td>24.26</td>
</tr>
<tr>
<td>Knowledge</td>
<td>93.56</td>
<td>31.58</td>
</tr>
</tbody>
</table>

**FIGURE 1. Scree plot for all 66 items on the Children’s Environmental Attitudes and Knowledge Scale.**
solution, after varimax rotation, resulted in the first factor with an eigenvalue of 9.9, close to the value of 8.3 reported by Leeming et al.

Once again, similar to published findings, the first factor explained 15.03% of the variance. This first factor predominantly comprised attitude questions. Three knowledge items (Items 13, 20, and 30) loaded onto this factor, but none had a loading greater than .185. The second factor explained a further 8.2% of the variance and predominantly comprised knowledge questions. The one attitude item loading on this factor (Item 1, verbal commitment) had a loading of −.134. In separate analysis for each of the two subscales, both the scree plot and distribution of variance (Table 5) supported the contention of a single dominant factor in each case.

Discussion

The benefits of having a psychometrically robust scale to measure children’s global environmental attitudes have been eloquently described previously (Leeming et al., 1995) and need only be concurred with here. The question of interest here was whether CHEAKS is applicable to a different age group and a different country.

A summary of the results indicates that data from this Irish sample support earlier claims for the psychometric credentials of CHEAKS. Although the findings in terms of reliability and validity are consistent and the two-factor structure is clear, differences in the essence of environmental awareness emerged when data from the two samples are juxtaposed. Comparison of mean scores on the CHEAKS scales (see Table 1), on each administration, shows that attitude subscale scores for the students in the United States were significantly higher ($p < .01$), but the Irish students recorded higher ($p < .01$) knowledge scale scores. This finding is interesting in light of the comparatively low level of knowledge reported in the European Union (EU) survey of environmental attitudes and knowledge for Irish participants (Faughnan & McCabe, 1998), on which Ireland ranked sixth on mean scores from seven countries. It must be remembered that the measures in the two studies were quite different, however; this study sampled responses to 36 items presented with six possible answers, one correct and five false; whereas the EU study reported on only 7 items. Some of the statements were worded so that the correct answer was true (e.g., “Some radioactive waste from nuclear power stations will be dangerous for thousands of years,” “Human beings are the main cause of plant and animal species dying out”) and others so that the correct answer was false (e.g., “If someone is exposed to any amount of radioactivity, they are certain to die as a result,” “Cars are not really an important cause of air pollution”). These statements were presented with the response options definitely true, probably true, probably not true, or definitely not

| TABLE 5. Factor Analysis Results for the Children’s Environmental Attitudes and Knowledge Scale Scores |
|--------------------------------------------------|--------------------------------------------------|--------------------------------------------------|--------------------------------------------------|
| Factor | Eigenvalue | Variance (%) | Factor | Eigenvalue | Variance (%) | Factor | Eigenvalue | Variance (%) | Factor | Eigenvalue | Variance (%) |
| Total | 2 | 9.92 | 15.03 | 2 | 8.30 | 12.6 | 2 | 8.30 | 12.6 |
| | 5.41 | 8.20 | 1 | 3.90 | 6.0 | 3 | 3.90 | 6.0 |
| Attitude | 1 | 9.55 | 26.52 | 1 | 8.09 | 22.5 | 2 | 8.09 | 22.5 |
| Knowledge | 1 | 4.89 | 16.30 | 1 | 3.73 | 12.4 | 3 | 3.73 | 12.4 |

THE JOURNAL OF ENVIRONMENTAL EDUCATION
true, and correct responses were interpreted from either of the appropriately directed definitely or probably response options. In addition, no between-country differentiation was made on the basis of age or education level in the EU study, despite a distinct variation in knowledge scores in the Irish data on this basis. Third, the present sample was 2 years younger than the youngest participant in the EU study. Nonetheless, this observation may have implications for cross-cultural application of educational materials or intervention strategies. The relative positions of the Irish and U.S. respondents switched when scores for the total scale were examined, indicating that, at a global level of environmentalism, the older group of U.S. students (aged 10–13 years) scored higher than the Dublin students (aged 15–17 years) at both administration times.

In their initial validation study, Leeming et al. (1995) found CHEAKS to be highly reliable and valid for children up to Grade 7 (aged 12–13 years). One of the issues for future research identified by the scale authors was the question of how reliable the scale would be with older children. In this article, we present data from an older group of students living in a geographically different location from the participants in the original study. We have employed, where possible, the same administration and analysis procedures to maximize comparability across the studies. Findings support the original validation study to the extent that the instrument retains its validity to a degree that is quite impressive.

We adopted three approaches to assess the reliability of each of the scales. The first line of investigation demonstrated the stability of scores between the first and second administration of protocols. We observed the significant ($p < .001$) test–retest correlation coefficients were observed on each of the subscales.

Second, results of internal consistency analysis corroborate the stability of results. Alpha coefficients for the CHEAKS total scale and attitude and knowledge subscales were consistently high (.77–.92) at both administration times. We observed the lowest values for the knowledge subscale at the first administration; however, an improvement had occurred at the time of the second administration. These findings concur with previous results and add to the body of evidence for internal consistency that support the scale’s use for research purposes (Bryman & Cramer, 1977) and screening purposes (Leeming et al., 1995).

The third approach adopted in examining the consistency of participants’ responses on CHEAKS was to examine the strength of discriminating levels of knowledge. By contrast to the average 44% correct responses previously recorded (Leeming et al., 1995), in our study, respondents answered approximately 50% of the items correctly. Given the chance expectation of 20% on this 5-option response format, this finding suggests a reasonable level of knowledge among this group of students. This is even more salient given the desire to provide a scale with strong discriminative ability without the danger of a ceiling effect occurring. When we examine the distribution of mean scores, we can see that some of the students did employ guessing as a strategy for answering the knowledge items; however, this was not the general trend.

Results from the contrasted-groups validity analysis lend further weight to the assertion that CHEAKS is an impressive instrument. The ability of CHEAKS to discriminate between self-rated environmental scores was observed not only in trends that emerged in the mean scores (see Table 4) but was also confirmed by the ANOVA scores from the group. Those students who rated themselves more environmentally committed scored consistently higher on both the attitude and knowledge subscales as well as on the total scale of CHEAKS.

The final analysis conducted involved replicated factor analysis. The model proposed in the Leeming et al. (1995) study (see Figure 1 and Table 5) adequately represented the Dublin data, and the results were quite similar. Once again, two factors emerged in the CHEAKS items, and the attitude and knowledge items loaded appropriately onto each of these factors.
Conclusion

The need to demonstrate psychometric credentials of a given instrument is well established in measuring environmental awareness, and there has been increasing interest in this endeavor, for example, environmental hazard perception (Walsh-Daneshmandi & MacLachlan, 2000), as well as for more global perspectives (Gifford, 2002). In this article, we have contributed to the area of EE evaluation by reporting results of validity study of CHEAKS conducted with a sample of Irish adolescents. The average age of the children in the present study, at 16 years, was higher than that of the original validation study, and findings support the confidence placed in the reliability of this measure. Contrasts between the findings in this study and that by Leeming and colleagues (1995) permit an interesting picture to emerge of possible cultural differences between U.S. and Irish samples, although we calculated inferential statistics using summary scores from two different studies. Notwithstanding the possible limitations associated with these procedures, the trends for difference in the data across the two countries remain. The source of such difference, however, awaits further exploration considering the myriad factors involved in any discussion of cultural effects (MacLachlan, 1997). For example, the data suggest age effects on knowledge scores but not on attitude scores. Given the level of difficulty in the knowledge items, it seems reasonable to suggest that CHEAKS may be just as appropriate for use with an older population. This expands the potential utility of the scale from both developmental and cross-cultural perspectives.

The conceptual space relating the person to the environment may be quite unique for the teenager. Levels of altruism and interest in philosophical perspectives are heightened at this stage of development. Another distinguishing characteristic at this stage of development is the greater influence attributable to one’s peers. Therefore, the motivating forces for environmental consciousness may be different in the adolescent from those salient for other age groups. Worldview may be more strongly rooted in the public persona at this time of life, rather than the integrated perspective in later stages of development or the egocentric perspective of earlier stages of development. The utility of one single measure to capture levels of environmental attitudes and knowledge across ages is clear. Our findings from these data suggest the suitability of CHEAKS as one such measure.

What remained unclear until now, however, was the cultural appropriateness of using such a measure in Ireland. Although there are many similarities within the U.S. and Irish cultures, the development of environmental education is quite different in the two countries. This article contributes to the evidence for the potential usefulness of using CHEAKS in the two cultural settings. Given the globalization of education and the transferability of educational material across geographic boundaries, it is increasingly important to match evaluation tools with educational material. When educators outside the United States draw on resources generated in the United States or vice versa, it also may be possible to evaluate effectiveness of such materials using a measure with demonstrated psychometric validity.

The sample used in this data was convenient in nature, and we recommend further study to collect data from a national representative sample across age groups. Such data would facilitate the development of CHEAKS norms, which in turn would greatly enhance the utility of such an instrument in the evaluation of EE programs in Ireland and elsewhere. The very structure of CHEAKS affords a unique way to measure self-reports of ecological activity alongside attitudes and knowledge. This may provide a tool to evaluate Gruenewald’s (2004) assertion that “the institutionalization of EE within general education has diluted it of its political purpose of acknowledging and transforming these crises” (p. 71).
NOTE
There are two state examinations during second-level education in Ireland, following the junior and senior cycles, lasting 3 and 2 years, respectively. The Transition Year Program, when available, is offered following the junior cycle. The curriculum for the program is developed in each school under broad guidelines issued by the Department of Education. In addition to a small number of core subjects, a variety of experiences, including work experience, are offered to the students.

REFERENCES