SOCIAL BEHAVIOR AND PERSONALITY, 2001, 29 (1), 71-80

MISSING VALUES IN PEER ASSESSMENT OF SOCIAL BEHAVIOR USING THE RCP

PAUL VEDDER
Leiden University, The Netherlands

MINA O'DOWD
Stockholm University, Sweden

This article deals with probable causes for missing values when using the Revised Class Play instrument for peer assessment of social behavior. The study was conducted with 10 - 13 year old primary school children in Sweden. The findings reveal that missing values may be caused by the cognitive and linguistic complexity of the instrument (RCP), as well as by gender stereotypes.

The results indicate that investigating the causes of missing values provides valuable information about the validity of the RCP. This information can be used to improve the instructions on how and when to use the RCP.

The social competence of school children is important for their cognitive and academic functioning (cf. Barnes & Sternberg, 1989; Patrick, 1997; Wentzel & Asher, 1995) as well as for their social and emotional wellbeing (Morison & Masten, 1991; Van Lieshout & Ferguson, 1991). Children who are aggressive or who disturb other children's activities — children who, in other words, are not socially competent — run a higher risk of drop-out and juvenile delinquency (cf. Parker & Asher, 1987). In addition to the social and educational consequences of problem-

Paul Vedder, Leiden University, Center for the Study of Education and Instruction, The Netherlands, and Mina O'Dowd, Stockholm University, Institute of International Education, Sweden.

This study was made possible by the Nils-Eric Svensson stipend of the Bank of Sweden Tercentenary Foundation awarded to the first author in 1995.

Appreciation is due to reviewers including: Dr. Gerard H. Maassen, Department of Methodology and Statistics, Utrecht University, The Netherlands, and Dr. Francisco Braza, Estacion Bioligica de Donana, Seville, Spain.

Please address correspondence and reprint requests to: Paul Vedder, Leiden University, Center for the Study of Education and Instruction, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands. Phone: (31) 71-5273401; Fax: (31) 71-5273619; Email: <vedder@fsw.leidenuniv.nl >

atic behavior, social competence has generally become more important, not least due to developments in the labor market: Ekstrand (1994) maintains that children's ability to "enjoy being in a group... is important for adults in a society where almost everybody works outside the home" (p. 10), such as is the case in the majority of countries in the Western world.

The importance of social competence in the school setting poses new and challenging tasks for assessment and training. This paper focuses on the task of assessment. Given the fact that there are numerous ways of assessing social competence and that these methods all have particular functions, benefits and shortcomings, the authors wish to stress that their contribution is limited to the use and validity of one particular measure of social competence, namely The Revised Class Play (henceforth referred to as RCP; Masten, Morison & Pellegrini, 1985). For reviews and evaluations of other measures of social competence, reference is made to Demaray, et al. (1995) and Haager and Vaughn (1995), among others.

The RCP is widely used as a measure for the peer assessment of social behavior. All children in a class are asked to decide which of their classmates would be best suited to play a particular role in an imaginary play. Each of the roles reflects either positive or negative behavior. Using a sample from the USA, Masten and her colleagues established that the RCP yields three-factor analytically derived scores: one for sociability/leadership, one for aggressive/disruptive behavior, and one for sensitivity and isolation. The first score is considered to reflect positive aspects of social competence, whereas the other scores are considered to represent negative aspects of social competence.

The same factor structure has been replicated in a number of studies in a variety of countries: The USA (Luthar & McMahon, 1996; Realmuto, August, Sieler & Pessoa-Brandao, 1997), Canada (Chen, Rubin & Sun, 1992), China (Chen, Rubin & Li, 1997), Israel (Krispin, Sternberg & Lamb 1992), Italy (Casiglia, LoCoco, & Zapulla, 1998), Sweden (Vedder & O'Dowd, 1999), the Netherlands (Aleva, 1992; Vedder, 1999), and the Netherlands Antilles (in the Caribbean)(Kromhout & Vedder, 1996; Vedder, 1999).

The RCP has met with certain difficulties when it has been used in diverse settings. For example, the instrument met a certain amount of reluctance or resistance when it was used in the Netherlands Antilles (Vedder, 1999), because the children who participated in the study found it difficult to complete all of the items. Their reluctance resulted in 20 percent missing values. Unfortunately, researchers who use the RCP do not usually report on the frequency of missing values. The authors postulate that, when analyzed, missing values can provide insight into the validity of the RCP, albeit in terms of its limitations. The following example from the Antillean study can illustrate this point. In this study the 10 - 13 year olds were asked to propose two casts for an imaginary play: one cast consisting of boys and one cast consisting of girls. There were two reasons for this request: the first reason

was that the reliability of scores for the chosen children would increase, if all of the children were asked to include both girls and boys, and the second reason was that the authors wished to study gender-specific patterns of choice, in order to gain better understanding of the consequences of gender-specific socialization practices in the Netherlands Antilles (Vedder, 1995; Vedder & Kook, 1993). Missing values were anticipated, although the authors were not prepared for the 20 percent missing values they received. These missing values are the result of the children's unwillingness to choose children of the opposite sex. In the Antilles, the RCP clearly indicates children's willingness to bridge the gender gap. In Sweden, where the authors conducted another study on social competence among primary school pupils, missing values were again encountered - unexpectedly - which provided an opportunity for closer study of the nature of missing values in the RCP.

With reference to gender-specific socialization, Sweden differs considerably from the Netherlands Antilles. Much has been done, and continues to be done, to bridge the gender gap in Sweden in all of the arenas of social interaction. The authors assumed that Swedish children would be less reluctant to complete the items and to choose two casts, one for boys and one for girls. Furthermore, they assumed that fewer missing values would be found. As will be shortly shown, this assumption was ungrounded.

In addition to efforts made to analyze the effect of sex stereotypes on the completion of the RCP, an attempt was also made to analyze the cognitive and linguistic complexity of the role descriptions as a possible cause of missing values.

METHOD

As already indicated, the authors did not plan to do a study on missing values. Had such a study been planned they would, at the outset, have included measures for the children's level of cognitive and language proficiency, in order to study the relationship between non-response and the cognitive and linguistic complexity of the items. Instead it was decided to let experts rate the complexity of the items. As a consequence, the optimal design, with which the combined and separate influence of subject and item characteristics on missing values would have been measured on the same subjects, could not be used. The authors were necessarily obliged to do separate analyses for the influence of gender and for the influence of cognitive and linguistic complexity. The unit of analysis in this study is the item.

Participants

In two schools in northern Stockholm, fourth, fifth and sixth grade children agreed to participate in a pilot study, the purpose of which was to test the usefulness of a translated and adapted version of the RCP study. A total of 156 children in seven classes, 51 percent of which were girls, completed the RCP-list. Their

ages ranged from 10 - 13. The pupils were predominantly Swedish middle-class children (95 percent).

For the second part of the study, five Swedish Ph.D. students in education at Stockholm University, 32 students in communication and social work at the University of Gävle, and six primary school teachers in Södertalje rated the linguistic and cognitive complexity of the 32 role descriptions that comprise the RCP.

Instrument

The RCP consists of one-sentence descriptions of behavior. Most of these describe behavior that is socially competent, according to Western norms, and are labeled 'sociability/leadership'. The other items describe behavior that is socially incompetent, according to Western norms, and are labeled either 'aggressive/disruptive' or 'sensitive/isolated'. Examples of the items are 'Makes new friends easily', 'Picks on other kids' and 'Rather plays alone than with others'. The Swedish version that was used is a translation of the version used by Kromhout and Vedder (1996) in the Netherlands, which includes 24 items from the original American version (Masten et al., 1985), and eight adapted items. The 32 Dutch role descriptions were translated into Swedish and back-translated to ensure similarity with the Dutch version.

Procedure

In both schools teachers agreed to collect the data at a time and on a day that suited them. The RCP-forms were to be collected two weeks after delivery. In one school the headmaster was instructed how to use the forms, and he promised to instruct the teachers in the school. In the other school, the researchers instructed the teachers. Written instructions were given to all of the teachers as well, which they could read aloud to their pupils, should they decide to do so.

The instrument is administered in school classes. Each of the children receives a paper on which the items are listed as well as a list with the names of all of their classmates, to which numbers are also assigned. All of the children are asked to pretend to be the directors of a play. Each child is asked to write the number of one boy and one girl for every role, choosing the classmates who most closely resemble each role description. Children are not allowed to choose either themselves or anyone who does not belong to the class. The list with names has two functions. Firstly, the list serves as a reminder, so that the children will not forget to include the names of classmates who do not happen to be present on the day in question. Secondly, each child has a number on the list. The children are instructed to use this number when they fill in the list. This measure was taken in order to guarantee confidentiality for all those involved in the study. Furthermore, the instructions state that children are encouraged to skip as few items as possible.

In order to analyze the possible influence of the cognitive and linguistic complexity of the role descriptions, the experts were asked to read all of the role descriptions and to indicate eight roles which they assessed to be the most cognitively and linguistically complex for 10-11 year old children.

RESULTS

The overall response rates (average of the percentage of items completed, divided by 100) in the seven classes were .47, .51, .40, .47, .60, .36, and .48. When the forms were collected, some teachers said that many children found it difficult to choose children of the opposite sex. They also commented that some children did not understand certain role descriptions and that some could not choose a child for particular roles, as they felt that none of the children in their class fitted the role description.

In school A, two teachers said that they had not received the instructions specifying that children should be encouraged to complete as many items as possible - despite the fact that they had received the written instructions in which this fact was clearly stated. Teachers in school B had decided not to push children to complete the items that they had skipped. These children had expressed either a lack of understanding of the descriptions by asking questions and by getting annoyed - or an unwillingness to complete all of the items. Teachers had allowed these children to skip most of the role descriptions for which they could not easily choose a classmate.

In order to determine the children's response scores, the absolute number of choices made by 156 children for each of the items was calculated. This resulted in 32 scores divided in three clusters, which correspond to the three factors represented in the RCP: 13 scores representing social competence/leadership, 11 scores representing aggressive/disruptive behavior, and eight scores for sensitivity/isolation. A low score means a low response rate, while a high score means a high response rate for a particular item. In order to analyze gender influences, the authors distinguished scores given to boys (RCP-b), scores given to girls (RCP-g), and the sum of the scores given to boys and girls (RCP). In Table 1, the average response scores per distinguished cluster of items are presented. This average score is the sum of response scores for the items of a cluster, divided by the number of items in a cluster. The Modified Least Significant Difference test (ModLSD) was used to test the differences between all average scores. This is a conservative procedure for pair-wise comparisons of means. The *t*-tests conducted are two-sided and the observed significance level (p) is adjusted for the fact that multiple comparisons are made. The uncorrected significance level is .05.

TABLE 1
MEAN RESPONSE SCORES BY RCP-FACTOR SCALE (STANDARD DEVIATIONS BETWEEN BRACKETS)

	RCP	RCP-B	RCP-G
1) social competence	175.8 (35.0)	85.6 (18.0)	90.2 (19.9)
2) aggressiveness	153.7 (47.4)	97.0 (24.5)	56.7 (24.3)
3) sensitivity	112.3 (44.3)	47.3 (21.0)	65.0 (24.7)
F, *df* = 31	5.7 (p=.008)	13.6(p=.001)	7.0(p=.003)
ModLSD	1>3	1, 2 > 3	1 > 2, 3

Table 1, RCP column, shows that the children deem items measuring sensitivity to be more difficult to assign than the social competence and aggressiveness items. This picture is, however, somewhat complicated when the gender of the chosen children is taken into consideration. Both boys and girls find it is more difficult to attribute roles dealing with sensitivity to boys than it is to attribute roles that reflect either aggressiveness or social competence. When it comes to choosing girls for roles, it is equally difficult for children to attribute roles that reflect sensitivity as it is for them to choose roles that reflect aggressiveness. It is easiest for children to choose a girl for role descriptions of positive, socially competent behavior.

When the differences between the choices given to girls and to boys (T-test, paired samples, $p<.05$) were tested (the rows in Table 1), this measure revealed that overall, irrespective of subscale, it is equally easy to choose a boy or a girl ($T = 1.5$; $p = .14$); the non-response, when boys are chosen, is comparable to the non-response, when girls are chosen. The same holds true for the subscale "social competence" ($T = -1.1$; $p = .28$). With regard to items from the subscale "aggressiveness", however, it is much easier to choose a boy than a girl ($T = 10.7$; $p = .00$), whereas with items that reflect sensitivity, it is easier to choose girls ($T = -4.2$; $p = .004$). The picture that emerges is one of traditional gender stereotypes: Boys are aggressive, girls are not. Girls are sensitive, boys are not. The possible correspondence with differences in actual behavior between boys and girls was not explored in the present study.

Whether or not the cognitive and linguistic complexity of the descriptions plays a role in children's non-response was also explored. Table 2 shows the mean difficulty score for each of the RCP-factor scales. The difficulty score is the total number of times a particular item was chosen by one of the 43 adult experts. Each item has a difficulty score. A low score means that adults evaluate a particular item to be rather easy, while a high score indicates that the adults evaluate a particular item to be linguistically or cognitively complex.

TABLE 2
MEAN DIFFICULTY SCORES BY RCP-FACTOR SCALE (STANDARD DEVIATIONS BETWEEN BRACKETS)

Social competence	aggressiveness	sensitivity	*F,df*=31	*ModLSD*
6.9 (3.9)	10.4 (4.4)	17.1 (5.5)	12.7 (p=.001)	sen>soc,agg.

Table 2 shows that the adults deem items measuring sensitivity to be more difficult than the items measuring social competence and aggressiveness. Their evaluation corresponds, at this factor scale level, with the differences in average response scores for boys and girls together (see Table 1).

A correlation (Pearson p.m.) between the children's response scores and the adults' difficulty scores was calculated. Since the items are treated as cases, there are 32 cases. The correlation amounted to -.56 ($p < .001$), indicating that, when adults evaluate an item as cognitively and linguistically more demanding than other items, the children's response score for that item is lower. It seems that part of the disappointing response rates can be explained by the difficulty children have in responding, due to the cognitive and linguistic complexity of the items. The design did not facilitate the discernment of the relative influence of gender and cognitive/ linguistic complexity on children's missing values.

DISCUSSION

This non-response study provides the following findings:

- the linguistic and cognitive complexity of some items probably explains why the children had difficulty completing these items;
- items for the subscale sensitivity are slightly more difficult than the other items;
- the occurrence of missing values is probably affected also by gender stereotypes: For items referring to aggressiveness, it is easier to mention boys than girls, whereas for items referring to sensitivity, it is easier to choose girls than boys.

As stated before, this study was actually meant to be a pilot study, conducted prior to a large study, on social competence in Swedish primary schools (see Vedder & O'Dowd, 1996, 1999). As a consequence of the findings from the pilot study, some items were adapted, and it was decided that the authors should be present in the classroom to instruct the children and to assist teachers. While the children completed the RCP, the authors walked around in the classroom and checked whether or not children had skipped roles or items. The children who had skipped roles were encouraged to try one more time. Five of the 671 participating children became noticeably upset when they were asked to complete the RCP. Three of the

children were upset due to the fact that it took them such a long time to make their choices. The other two children had difficulty coping with the idea that they were asked to assign roles that referred to "bad behavior" to one or more of their classmates. The first three children were all coached to complete the task, while their classmates were working on another task. Upon completion of the task, the children expressed the opinion that they were happy that they had completed the RCP. The last two children did not complete the task. Afterwards, the authors had a short conversation with these children, in which they stressed that they understood their reluctance. It was emphasized that they should not worry about not completing the task. The remaining 666 children all completed the RCP with a maximum of two roles for boys and two roles for girls not being cast. It seems as if neither the cognitive and linguistic complexity nor sex role stereotypes affected these children. Nevertheless, there is reason to believe that the children in this pilot study were not very different from the children in the main study. The children in the main study appear to have managed to resolve their problems with the cognitive and linguistic complexity of the role descriptions.

Cognitive and linguistic complexity is a common problem in research instruments. Ensuring that the formulations used are able to be understood by the subjects is the only solution to this problem. The difference in the number of missing values between the pilot study and the main study suggests that the lack of missing values in the latter may be partly the consequence of the children's increased effort, facilitating their understanding of most of the items, and partly attributable to the instructions encouraging them to complete as many items as possible. What does this indicate about the validity of the RCP scores?

If a subject - despite initial reluctance - nevertheless chooses a peer for a role, then the score received by the peer in question is not necessarily invalid: The subject may have reconsidered the prior evaluation, or may just have made a second effort to accomplish the task at hand. If, however, a child decides to skip an item, then this also may be the "best score" possible, assuming that neither re-consideration nor added effort would have led to a valid alternative.

The aspects of social competence that are measured with the RCP are not evenly distributed in, and between, groups of children. Aggressive children, for instance, may be absent from particular groups, whereas they may be abundantly present in other groups. Both situations make it difficult for children to choose one child for a particular RCP item representing aggressive behavior. Re-consideration and extra effort will not lead to improved validity for the choice that is eventually made by the pupil under these circumstances. In fact, such circumstances threaten the validity of the RCP and can be seen more generally as the consequence of the fact that subjects have to choose one child, or one girl and one boy, for each role. In these situations a missing value can be considered a "best score", when the alternative is that positive or negative qualities are assigned inaccurately.

It is proposed - for these reasons - that researchers encourage students to complete as many items as possible. However, children should explicitly be allowed to decide whether or not it is possible for them to make a fair choice, even if this decision entails their skipping one or more items.

The authors have suggested that missing values, as well as regular scores linked with cross gender evaluations, may stem from gender stereotypes. Other studies - for an overview see Hymel, Wagner, and Butler (1990) — suggest that peer evaluations of social competence reflect both actual social behavior and stereotypes. RCP scores do not allow for a distinction between these two sources of differences in social competence between children. This is especially troublesome when the RCP is used in educational settings. Although it is evident that it is important for an educator to know whether efforts should be made to try to change actual behavior, to do something about stereotypes or to attempt to rectify both, educators will need more information than that provided by RCP-scores alone to be able to make appropriate decisions with regard to the corrective procedures.

REFERENCES

Aleva, L. (1992). Assessment methods for social competence in elementary school. In P. Vedder (Ed.) *Measuring the Quality of Education (79-91)*. Amsterdam: Swets & Zeitlinger.

Barnes, M. L. & Sternberg, R. (1989). Social intelligence and decoding of non-verbal cues. *Intelligence*, 13(3), 263-287.

Casiglia, A., LoCoco, A. & Zappulla, C. (1998). Aspects of social reputation and peer relationships in Italian children: A cross-cultural perspective. *Developmental Psychology*, 34, 723-730.

Chen, X., Rubin, K. & Sun, Y. (1992). Social reputation and peer relationships in Chinese and Canadian children: A cross-cultural study. *Child Development*, 63, 1336-1343.

Chen, X., Rubin, K.H. & Li, D. (1997). Relation between academic achievement and social adjustment: Evidence from Chinese children. *Developmental Psychology*, 33, 518-525.

Demaray, M. K., Ruffalo, S. L., Carlson, J., Busse, R. T., Olson, A. E., McManus, S. M. & Leventhal, A. (1995). Social skills assessment: A comparative evaluation of six published rating scales. *School Psychology Review*, 24, 648-671.

Ekstrand, G. (1994). *Children, norms and values in three countries on three continents*. Malmö: Malmö School of Education, University of Lund.

Haager, D. & Vaughn, Sh. (1995). Parent, teacher, peer, and self-reports of the social competence of students with learning disabilities. *Journal of Learning Disabilities*, 28(4), 205-215.

Hymel, S., Wagner, E. & Butler, L. (1990). Reputational bias: View from the peer group. In S.R. Asher & J.D. Coie (Eds.) *Peer rejection in childhood* (156-186). Cambridge: Cambridge University Press.

Krispin, O., Sternberg, K. & Lamb, M. (1992). The dimensions of peer evaluation in Israel: A cross-cultural perspective. *International Journal of Behavioral Development*, 15, 299-314.

Kromhout, M. & Vedder, P. (1996). Cultural inversion in children from the Antilles and Aruba in the Netherlands. *Anthropology and Education Quarterly*, 27, 568-586.

Luthar, S. & McMahon, Th. (1996). Peer reputation among inner-city adolescents: Structure and correlates. *Journal of Research on Adolescence*, 6, 581-603.

Masten, R., Morison, P. & Pellegrini, D. (1985). A revised class play method of peer assessment. *Developmental Psychology*, 21, 523-533.

Morison, P. & Masten, A. (1991). Peer reputation in middle childhood as a predictor of adaptation in adolescence: A seven-year follow-up. *Child Development*, **62**, 991-1007.

Parker, J. & Asher, S. (1987). Peer relations and later personal adjustment. *Psychological Bulletin*, **102**, 357-389.

Patrick, H. (1997). Social self-regulation: Exploring the relations between children's social relationships, academic self-regulation, and school performance. *Educational Psychologist*, **32**(4), 209-220.

Realmuto, G., August, G., Sieler, J. & Pessoa-Brandao, L. (1997). Peer assessment of social reputation in community samples of disruptive and nondisruptive children: Utility of the revised class play method. *Journal of Clinical Child Psychology*, **26** (1), 67-76.

Van Lieshout, C. & Ferguson, T. (1991) Relaties met leeftijdgenoten in het onderwijs. [Peer relationships in school]. *Pedagogische Studiën*, **68**, 45-55.

Vedder, P. (1995). *Antilliaanse kinderen. Taal, opvoeding, en onderwijs op de Antillen en in Nederland.* [Antillean children. Their language, education, and schooling] Utrecht: Jan van Arkel.

Vedder, P. (1999). Primary schoolchildren's social competence in three countries. In J.C. Lasry, J.G. Adair & K.L. Dion (Eds.) *Latest contributions to cross-cultural psychology* (318-330). Amsterdam: Swets & Zeitlinger.

Vedder, P. & Kook, H. (1993). Early childhood care and education in Curaçao. *International Journal of Early Years Education*, **1**(3), 29 - 34

Vedder, P. & O'Dowd, M. (1996). *Social Competence in Swedish Primary School Pupils. The Validity of the RCP-scale.* Stockholm: The Institute of International Education Stockholm University.

Vedder, P. & O'Dowd, M. (1999). Swedish primary school pupils' interethnic relationships. *Scandinavian Journal of Psychology*, **40**(3), 221-228.

Wentzel, K. R. & Asher, S. R. (1995). The academic lives of neglected, rejected, popular, and controversial children. *Child Development*, **66**, 754-763.