ijbd

International Journal of Behavioral Development
2015, Vol. 39(1) 77–86
© The Author(s) 2014
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0165025414551761
ijbd.sagepub.com
SAGE

Nominating under constraints: A systematic comparison of unlimited and limited peer nomination methodologies in elementary school

Rob Gommans[1,2] and Antonius H. N. Cillessen[1]

Abstract

Children's peer relationships are frequently assessed with peer nominations. An important methodological issue is whether to collect unlimited or limited nominations. Some researchers have argued that the psychometric differences between both methods are negligible, while others have claimed that one is superior over the other. The current study compared both methods directly in a counterbalanced design among 112 8–12-year-old elementary school children. Overall, both methods revealed comparable results, although some significant and noteworthy differences were found. The use of unlimited nominations was recommended for questions related to social status (preference, popularity). Some method differences varied by gender. Implications for future peer relations research were discussed.

Keywords

peer assessment, peer relations, research methods, sociometric measurement

Sociometric methods (Moreno, 1934) are commonly used in the study of peer relations to assess the positive and negative links between children or adolescents (Bukowski & Cillessen, 1998; Cillessen, 2009). Sociometric methods use multiple informants (group members) to gain insight in a group's social structure. Ideally, a round-robin design is used, allowing all group members to evaluate each other on one or more sociometric questions or criteria. The rationale behind this procedure is that each group member has knowledge of the status, behaviours and reputations of the other group members.

Various methods are available to collect sociometric data, in particular peer nominations, peer ratings, rank ordering, and paired comparisons. Of these methods, peer nominations are most commonly used (Maassen, van Boxtel, & Goossens, 2005; Poulin & Dishion, 2008). In this method, group members are asked to nominate other group members for various constructs of peer affiliation and social behaviour. The peer nomination method is relatively easy to implement. Peer nominations are easy to understand for participants and can be implemented in such a way that they are not exceptionally time-consuming.

Despite the frequent use of nominations in peer research, an important but unsettled methodological issue is whether to collect unlimited or limited nominations (Cillessen, 2009; Hallinan, 1981; Hymel, Vaillancourt, McDougall, & Renshaw, 2002; Santo, Mayman, Lopez, & Bukowski, 2012; Terry, 2000). In the limited method, children make a fixed and relatively small number of nominations (usually three). In the unlimited method, they are allowed to nominate as many or as few peers as they see fit for each question. A disadvantage of a paper-and-pencil application of this method with children is that it may take them very long to complete the measure. Therefore, the number of nominations is often capped at a number that is large enough to be considered unlimited but provides some restriction to help children move forward (frequently 10). We might call this method quasi-unlimited. The quasi-unlimited method is referred to as "unlimited" in the remainder of this article for the sake of readability.

There are potential advantages to unlimited nominations. First, they may yield more ecologically valid data than limited nominations (see Cillessen, 2009). This was originally the most important reason to shift from limited to unlimited nominations. During the practice of sociometric data collection with limited questions, children often indicated either that they knew more peers for certain questions or that they did not know as many for other questions than the fixed number requested (Terry, 2000). Especially for positive sociometric questions, children often know more peers than just three. In this case, restricting participants to three would lead to measurement error (Holland & Leinhardt, 1973). This suggests that researchers should be concerned about artificial limits placed on positive sociometric questions such as liking and friendship, where children are expected to nominate more than the fixed number allowed (Holland & Leinhardt, 1973; Moreno, 1951).

Second, evidence suggests that sociometric scores derived from unlimited nominations are more stable and reliable than those derived from limited nominations. A meta-analysis of longitudinal sociometric data showed that the stabilities of sociometric scores

[1] Radboud University Nijmegen, The Netherlands
[2] Utrecht University, The Netherlands

Corresponding author:
Rob Gommans, Behavioural Science Institute, Radboud University, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands.
Email: r.gommans@bsi.ru.nl

Downloaded from jbd.sagepub.com at University Library Utrecht on March 18, 2015

derived from unlimited nominations were at the high end of the distribution of stability coefficients (Jiang & Cillessen, 2005). For measures of social behaviour, unlimited nominations also yielded more stable scores and higher inter-construct correlations than limited nominations (Terry, 2000).

In addition to these advantages, researchers also have expressed concerns about unlimited methods. Data collection and analysis have been thought of as more straightforward and less time-consuming with limited nominations than with unlimited nominations (Newcomb & Bukowski, 1983). If children name too many peers with unlimited nominations and become less selective, this would reduce the validity of the data (Terry, 2000). Further, unlimited nominations might cause fatigue or frustration with the task (Poulin & Dishion, 2008). Writing down names or code numbers of many peers is more time consuming and laborious than writing down only a few.

Despite these arguments for and against both methods, the available empirical evidence is scarce. The existing evidence comes primarily from indirect comparisons of data collected in different samples of different studies (Cillessen, 2009). So far, only two studies directly compared both methods. Terry (2000) found that the unlimited method yielded superior distributional properties (less skew, more variability) and a superior model fit to the data than the limited method for sociometric questions of peer status (acceptance, rejection). Santo and colleagues (2012) reported better psychometric properties, goodness-of-fit, effect sizes, and power to detect interactions with unlimited than limited nominations of peer victimization. However, the limited method is still used frequently (e.g., Bellmore, Jiang, & Juvonen, 2010; Caravita, Di Blasio, & Salmivalli, 2008; Rose, Swenson, & Waller, 2004). Further, the fact that unlimited nominations have some advantages over limited nominations does not imply that limited nominations are invalid or unreliable in all situations.

The overall aim of this research was to extend the two previous studies by directly comparing quasi-unlimited nominations (capped at nine) with limited nominations (set at three) using 18 sociometric questions collected within the same sample in a counterbalanced design. The current study focused on complete networks by allowing nominations both within and across classrooms, including classrooms from different grades.

There were three specific goals. The first goal was to compare the means and intercorrelations of six sociometric questions assessed with both methods measuring social preference (liked most, liked least), popularity (most popular, least popular), and bullying involvement (bully, victim). We expected significantly higher numbers of nominations received in the unlimited method than in the limited method for the social preference and popularity questions. Based on previous findings, and given the average size and complexity of children's peer networks, we expected children on average to have more than three peers fitting each of these criteria. Related to this, we expected relatively smaller intercorrelations for these questions (positive and/or status-related). For bullying involvement, we expected no differences between methods, as children in general nominate only a few peers for them. As such, we also expected relatively larger intercorrelations for these (negative) questions.

The second goal was to assess the validity of both methods by examining the correlations between the six sociometric questions (preference, popularity, bullying involvement) assessed with each method and 12 peer reputation questions (e.g., aggression, prosocial behaviour, social exclusion) assessed with unlimited nominations. We hypothesized that the unlimited method would yield higher correlations between the six sociometric questions and the 12 peer reputation questions than the limited method for two reasons. First, with unlimited nominations, the reliability of the resulting construct scores will improve because more nominations (more data points) are available. Second, if the unlimited method allows children to more accurately express their "true" choices, validity of the construct scores will also improve. Lower skewness and kurtosis (i.e., more desirable psychometric properties) should yield a more pronounced pattern of correlations (less noise). By not constraining children's choices, they are expected to be better able to differentiate their peer relationships, resulting in a higher validity than with limited nominations (Holland & Leinhardt, 1973; Hymel & Asher, 1977; Terry, 2000).

The third goal was to examine more precisely whether the associations of peer acceptance (liked most) and rejection (liked least) with the 12 reputation questions differed between both methods. Exploratory structural equation modelling was used for this purpose. Similar to the second goal, we expected that the unlimited method would result in more significant and stronger associations of the relatively stable measures of social status with the 12 behavioural descriptors. In addition, we expected this method effect to be more pronounced for the positive status-related question (liked most) than for the negative status-related question (liked least).

Method

Participants

Participants were 112 elementary school children aged 8 to 12 years ($M_{age} = 9.93$; $SD = 1.25$; 52.7% boys) from four classrooms across four grades (Grade 3: $M_{age} = 8.46$, $SD = .59$, 41.7% boys; Grade 4: $M_{age} = 9.28$, $SD = .52$, 40.6% boys; Grade 5: $M_{age} = 10.24$, $SD = .52$, 56% boys; Grade 6: $M_{age} = 11.48$, $SD = .51$, 71% boys) of one elementary school in The Netherlands. These were the four classes for Grades 3 to 6 at this elementary school. Passive consent procedures requested by the school administrators for regular testing in the school context were followed. All students from the selected grades participated in the study. Participants were given a small toy after participation. The majority of the participants were native Dutch-speaking residents. The school is located in a neighbourhood with a SES factor score within 1 *SD* of the general mean, indicating the neighbourhood did not differ in socio-economically conditions from the Dutch average (SCP, 2006).

Procedure

Two classroom assessments were held in the fall of the school year, separated by a 2-week interval. Two paper-and-pencil sociometric questionnaires, one quasi-unlimited capped at nine, and one limited set at three, were administered in counterbalanced order. Two classrooms (Grades 3 and 6) completed the unlimited questionnaire in the first session and the limited questionnaire in the second session. The other two classrooms (Grades 4 and 5) completed the limited questionnaire first and the unlimited questionnaire second. The study was designed for the sole purpose of comparing unlimited and limited nominations. During the first session, children received no information about the second session to limit testing effects. Teachers also were instructed not to inform the students about the second session.

Downloaded from jbd.sagepub.com at University Library Utrecht on March 18, 2015

Table 1. Means, *SD*s, correlation coefficients and *Z* scores for raw sociometric scores by sex, unlimited and limited nominations received.

	Sociometric method								Mean differences		Correlation differences	
	Unlimited				Limited							
	M	*SD*	*Skew.* [a]	*Kurt.* [a]	*M*	*SD*	*Skew.* [a]	*Kurt.* [a]	ΔM (95% CI) [b]	*d*	*r* (95% CI) [c]	*q* [d]
Boys ($n = 59$)												
Liked most	7.02	3.44	.15	.14	2.97	1.83	.35	−.53	4.05 (3.40, 4.70)	1.47	.71 (.55, .82)	.33
Liked least	**4.81**	6.33	2.39	6.18	**3.59**	5.74	2.57	6.93	1.22 (.77, 1.67)	.20	.96 (.94, .98)	1.05
Most popular	3.39	5.24	2.71	8.07	2.14	4.23	3.52	13.58	1.25 (.76, 1.75)	.26	.94 (.90, .96)	.36
Least popular	3.15	5.35	2.45	5.54	2.00	4.46	3.12	9.49	1.15 (.57, 1.73)	.24	.91 (.86, .95)	.24
Bully	**3.73**	6.75	3.21	11.78	**2.80**	5.59	2.74	7.37	.93 (.25, 1.62)	.15	.93 (.88, .95)	.44
Victim	**2.27**	5.65	6.27	43.88	**1.76**	5.31	6.23	43.34	.51 (.19, .82)	.10	.98 (.96, .99)	1.35
Girls ($n = 53$)												
Liked most	7.55	3.04	−.26	−.22	2.91	2.02	.69	.38	4.64 (3.91, 5.38)	1.80	.50 (.27, .68)	
Liked least	**1.79**	2.20	2.19	5.30	**1.19**	1.84	3.18	12.39	.60 (.22, .98)	.30	.78 (.65, .87)	
Most popular	2.26	3.00	1.73	2.70	1.47	2.36	2.45	6.83	.79 (.39, 1.19)	.29	.88 (.80, .93)	
Least popular	2.45	2.95	2.40	6.21	1.62	3.00	3.53	13.34	.83 (.39, 1.27)	.28	.85 (.76, .91)	
Bully	**.28**	.93	4.95	28.50	**.28**	.87	4.07	18.58	.00 (−.14, .14)	.00	.83 (.73, .90)	
Victim	**.77**	1.40	2.10	3.89	**.72**	1.65	3.52	13.55	.05 (−.26, .37)	.03	.73 (.58, .84)	

Note. Significant gender differences at $p < .05$ are shown in bold.
[a] For boys, $SE_S = .311$, $SE_K = .613$; for girls, $SE_S = .327$, $SE_K = .644$.
[b] All means are significantly different between methods at $p < .01$, except for bully and victim for girls (ns).
[c] All correlation coefficients are significant at $p < .001$.
[d] Effect size *q* for gender differences in inter-method correlations.

Prior to the assessment in the first session, students were informed about the study during a 30-minute instruction period. They were told that the goal of the study was to understand children's peer relationships at school. They were also instructed how to complete the questionnaire. Confidentiality and privacy of answers were emphasized. Students were not allowed to talk during the assessment as they sat at their own desks in a test arrangement with adequate space between desks. An investigator was present in the classroom to answer questions and make sure instructions were followed. The teacher was also present for the majority of the time to make sure order was maintained.

Each participant received a roster with the names of all 112 participants (Grades 3 to 6), sorted alphabetically by first name, grouped by classroom, and preceded by a unique code number. Each questionnaire started with a short repetition of the instructions. Although some sociometric questions measured behaviours bound to the classroom (e.g., disrupts class), participants were allowed to nominate any student from any of the four classrooms for all questions. Previous research has shown that nominations received from other classrooms add unique information about the structure of the peer network (Poulin & Dishion, 2008). Further, students, particularly young adolescents, often have cross-grade friendships (Bowker & Spencer, 2010). In Dutch elementary schools, children usually remain in the same classroom throughout Grades 1 to 6 and thus know each other well. The school of this study was relatively small (all Grade 3 to 6 classrooms participated) and interactions across classrooms and grades were common in the schoolyard before and after school and during breaks. It follows that students knew each other well across classrooms and grades.

The number of blank lines provided for each question in both the limited (3) and unlimited (9) questionnaires indicated the maximum number of nominations allowed. Participants were instructed to use one blank line for each nomination. In the unlimited method, participants did not receive any information regarding the minimum number of nominations requested. If a participant expressed concern about who to nominate (if anyone), she or he was privately motivated once to nominate at least one peer for each question. In the limited method, participants were instructed to nominate three peers for each question. However, this instruction was not strictly enforced. That is, fewer nominations were accepted. Very few participants (< 5%) nominated more than the maximum (e.g., by adding their own blank lines). Such additional nominations were not included during data processing. Both same- and other-sex choices were allowed, as well as self-nominations to allow participants to indicate if they were victimized. Very few participants (< 5%) nominated themselves, and self-nominations were also not included during data processing.

Sociometric measures

For all questions in both instruments, the number of nominations received from all peers (across grades) were counted. These counts were used in all analyses. Both instruments started with the six questions for preference, popularity, and bullying. The unlimited instrument continued with the 12 reputational questions for peer affiliation, physical, relational and verbal aggression, social isolation and social exclusion, prosocial behaviour, disruptive behaviour, and leadership. Appendix A lists all questions.

Results

Descriptive statistics

First, we examined the distributions of the scores derived from both methods. The unlimited method yielded more desirable distributional properties as indicated by more variability and less skewness and kurtosis than the limited method, especially for the status measures (most and least liked; most and least popular, see Table 1). For bullying, these distributional properties were more desirable in the

Downloaded from jbd.sagepub.com at University Library Utrecht on March 18, 2015

limited method for both boys and girls. For victimization, skewness and kurtosis was smaller in the unlimited method for girls.

Second, we examined the number of nominations received for both methods (Goal 1; see Table 1). On average, participants received two to seven nominations for each question. As expected, the largest number of nominations occurred for liked most (acceptance). The negative bully and victim questions yielded the fewest nominations. Significant gender differences were found for these questions and for liked least (rejection). For these questions, girls received significantly fewer nominations than boys did. Nominations received were given by children in any of the four classrooms/grades. Cross-grade nominations were given particularly for negative questions (liked least, not hang around with, bully, calls names). On average, 21% of all nominations given, and 35% of all nominations given in response to negative questions, were cross-grade with the highest percentage of cross-grade nominations for the bully question (62%). For behaviours typically bound to the classroom, the percentage of cross-grade nominations given was low (disrupts class, 3%; cooperates, 4%).

Third, we examined differences in the number of nominations received between both methods using paired *t* tests (Goal 1). For most questions, the difference between methods was significant but small (see Table 1). In general, children received slightly more nominations in the unlimited method than in the limited method (ΔM = .00–1.25; Cohen's ds = .00–.30), except for liked most (acceptance). For liked most, children received about three nominations in the limited method and about seven in the unlimited method (ΔM_{boys} = 4.05, 95% CI_{boys}: 3.40–4.71, Cohen's d_{boys} = 1.47; ΔM_{girls} = 4.64, 95% CI_{girls}: 3.91–5.38, Cohen's d_{girls} = 1.80). For girls, there was no difference between methods in the number of nominations received for bully and victim.

Fourth, correlations between both methods were examined (see Table 1). All inter-method correlations were significant and moderate to large in size. However, one correlation was considerably smaller than the other correlations. As expected, the lowest inter-method correlation and largest departure from correspondence was observed for liked most (acceptance). Further, correlations between methods were lower for girls than for boys (effect size qs = .24–1.35) with the largest differences for liked least (q = 1.05) and victim (q = 1.35). For girls, the correlations between both methods were generally smaller than for boys.

Correlations between sociometric variables

To examine differences in validity (Goal 2), Pearson correlations were computed between the six sociometric questions assessed with both methods and the 12 additional reputation questions (Table 2). A Bonferroni correction was applied to control for family-wise error rate (FWER) in multiple comparisons ($p < \alpha/144$), meaning that associations that were marked as significant at $p < .001$ were actually significant at $p < .001/144 = .000007$. Of the 144 correlation contrasts conducted, 80 reached statistical significance before correction ($p < .001$). Moderate to strong correlations were found for both methods; the directions of the correlations were as expected. Specifically, hits or kicks, gossips or excludes, calls names, starts fights, disrupts class, talks badly, and leader correlated positively with liked least, most popular, and bully. Not play or work with and plays alone correlated positively with liked least, least popular, and victim. Furthermore, hang around with correlated positively with liked most and most popular, and negatively with least popular. Not hang around with correlated positively with liked least, least popular, bully, and victim. After correction, 15 out of 80 significant correlation contrasts became non-significant ($p >$ .000007; see Table 2): six for the unlimited method and nine for the limited method. Changes after correction were most pronounced for correlations with most popular.

Fisher's *r*-to-*Z* transformations were conducted to test whether corresponding correlations differed between methods. In general, correlations did not differ between methods (most Fisher's *z* scores were not significant) with one exception. The correlation between liked most and hang around with was significantly larger in the unlimited method than in the limited method, $z = 4.19$, $p < .001$.

The same correlations were examined for boys and girls separately. For boys, no further differences were found between methods, beyond the method effect for liked most and hang around with in the total sample. For girls, the correlation between liked most and not hang around with was higher in the unlimited method than in the limited method, $z = -1.99$, $p < .05$. This was also the case for the correlations between liked least and not hang around with ($z = 2.45$, $p < .05$), between liked most and plays alone ($z = -2.02$, $p < .05$), and between liked least and leader ($z = 2.01$, $p < .05$). Contrary to the higher correlations in the unlimited method, the negative correlation between victim and hang around with was larger in the limited method than in the unlimited method, $z = -2.86$, $p < .01$ (girls only).

Assessment of explanatory relationships and model fit

Path analyses were conducted in AMOS 18.0 to examine whether the associations of liked most and liked least with the reputation questions differed between methods (Goal 3). The exogenous variables were the standardized scores for liked most and liked least according to both methods; the endogenous variables were the standardized scores for the 12 peer reputation questions.

First, a fully saturated model ($df = 0$) was run for both methods separately. Standard error estimates were between 4% and 55% smaller in the unlimited method than in the limited method. A test for the equality of regression estimates (Paternoster, Brame, Mazerolle, & Piquero, 1998) showed that three estimates differed significantly between methods, with stronger estimates in the unlimited method, namely: hang around with regressed on liked most and liked least, and leader regressed on liked most.

Second, we tested a fully constrained model ($df = 24$) in which regression paths (e.g., from liked least to gossips or excludes) were set equal between both methods. The fit of this initial model was $\chi^2(24) = 58.92$, $p < .001$, RMSEA = .114 (CI: .078–.152, p = .003), CFI = .985, SRMR = .026. Table 3 gives the standardized estimates of this model. Analyses of the adequacy of the parameter estimates (Byrne, 2001) revealed no concerns.

Third, parameters with significant modification indexes (> 3.84) were freed one by one between methods in order of the size of the modification indexes. After each modification, goodness-of-fit was examined. If the χ^2 difference between the new model and the more constrained model was not significant ($\Delta\chi^2 < 3.84$ for $\Delta df = 1$), the process ended and the immediately preceding model was retained. If all parameters with significant modification indexes were freed, the process ended and the current model was retained.

In total, three parameters were freely estimated in the final model. Freeing them yielded a final model with good fit, $\chi^2(21) = 16.95$, $p = .714$, RMSEA = .000 (CI: .000–.062, p = .904),

Downloaded from jbd.sagepub.com at University Library Utrecht on March 18, 2015

Table 2. Correlation coefficients and 95% confidence intervals for sociometric scores by method.

	Liked most		Liked least		Most popular		Least popular		Bully		Victim	
	Unlimited	Limited	Unlimited	Limited	Unlimited	Limited	Unlimited	Limited	Unlimited	Limited	Unlimited	Limited
Hang around with	**.82***	**.54***	−.20	−.21	.41†	.36†	−.54*	−.48*	.01	−.09	−.26	−.32
	(.75, .88)	(.39, .66)	(−.37, −.02)	(−.38, −.02)	(.24, .55)	(.18, .51)	(−.66, −.40)	(−.61, −.33)	(−.18, .19)	(−.27, .10)	(−.42, −.07)	(−.48, −.15)
Not hang around with	−.45*	−.35†	.92*	.92*	.24	.23	.42*	.42†	.74*	.82*	.69*	.67*
	(−.59, −.29)	(−.50, −.17)	(.89, .94)	(.88, .94)	(.05, .40)	(.05, .40)	(.26, .57)	(.26, .56)	(.64, .81)	(.74, .87)	(.58, .78)	(.55, .76)
Cooperates	.46*	.38†	−.20	−.19	.19	.15	−.19	−.14	−.16	−.13	−.09	−.10
	(.29, .59)	(.21, .53)	(−.37, −.01)	(−.36, .00)	(.00, .36)	(−.03, .33)	(−.36, −.01)	(−.31, .05)	(−.33, .03)	(−.31, .05)	(−.27, .10)	(−.28, .09)
Not play or work with	−.43*	−.30	.58*	.61*	−.09	−.04	.69*	.74*	.26	.41†	.88*	.92*
	(−.57, −.27)	(−.46, −.12)	(.44, .69)	(.48, .72)	(−.27, .10)	(−.23, .14)	(.58, .78)	(.64, .81)	(.08, .42)	(.24, .55)	(.83, .91)	(.88, .94)
Plays alone	−.42†	−.23	.33†	.37†	−.17	−.11	.86*	.90*	.07	.19	.64*	.73*
	(−.56, −.25)	(−.40, −.05)	(.15, .49)	(.20, .52)	(−.34, .02)	(−.29, .08)	(.81, .90)	(.86, .93)	(−.12, .25)	(.00, .36)	(.52, .74)	(.64, .81)
Hits or kicks	−.27	−.21	.88*	.87*	.44*	.39†	.08	.07	.91*	.93*	.42*	.36†
	(−.43, −.09)	(−.38, −.02)	(.83, .91)	(.81, .91)	(.27, .57)	(.22, .54)	(−.11, .26)	(−.12, .25)	(.88, .94)	(.91, .95)	(.26, .57)	(.19, .51)
Gossips or excludes	−.21	−.16	.79*	.77*	.55*	.49*	.01	.03	.81*	.80*	.32	.27
	(−.38, −.03)	(−.34, .02)	(.70, .85)	(.68, .83)	(.41, .67)	(.33, .62)	(−.18, .19)	(−.16, .21)	(.73, .86)	(.72, .86)	(.14, .48)	(.10, .44)
Calls names	−.26	−.22	.81*	.80*	.51*	.44*	.02	.02	.93*	.90*	.29	.24
	(−.42, −.07)	(−.39, −.03)	(.74, .87)	(.73, .86)	(.35, .63)	(.28, .58)	(−.17, .20)	(−.17, .20)	(.90, .95)	(.86, .93)	(.11, .45)	(.06, .41)
Starts fights	−.31	−.24	.88*	.88*	.39†	.37†	.15	.13	.86*	.91*	.49*	.43*
	(−.47, −.13)	(−.41, −.06)	(.83, .92)	(.83, .92)	(.22, .54)	(.19, .52)	(−.04, .33)	(−.06, .31)	(.80, .90)	(.87, .94)	(.33, .62)	(.26, .57)
Disrupts class	−.10	−.11	.56*	.56*	.41†	.44*	.01	−.03	.59*	.59*	.08	.05
	(−.28, .09)	(−.29, .08)	(.41, .67)	(.41, .67)	(.25, .56)	(.27, .58)	(−.18, .19)	(−.21, .16)	(.46, .70)	(.45, .70)	(−.11, .26)	(−.13, .24)
Talks badly	−.24	−.22	.82*	.81*	.53*	.49*	.05	.04	.85*	.87*	.35†	.31
	(−.41, −.06)	(−.39, −.03)	(.75, .87)	(.73, .86)	(.38, .65)	(.34, .62)	(−.14, .23)	(−.14, .23)	(.79, .89)	(.81, .91)	(.18, .50)	(.13, .47)
Leader	.11	−.03	.58*	.53*	.78*	.79*	−.15	−.11	.67*	.60*	.15	.09
	(−.08, .29)	(−.21, .16)	(.44, .69)	(.38, .65)	(.70, .85)	(.71, .85)	(−.33, .04)	(−.29, .07)	(.57, .77)	(.46, .71)	(−.03, .33)	(−.10, .27)

Note. * $p < .001$. Significant differences in correlation coefficients between methods at $p < .001$ are shown in bold. A Bonferroni correction was applied to control for FWER in multiple comparisons ($p < \alpha/144$). Those correlation coefficients marked with a † were significant at $p < .001$ before correction, but became non-significant after correction.

Downloaded from jbd.sagepub.com at University Library Utrecht on March 18, 2015

Table 3. Standardized regression estimates and 95% confidence intervals predicting liked most and liked least from peer reputation questions by method.

	Fully saturated model ($df = 0$)				Constrained model ($df = 24$) [b]		Final model ($df = 21$) [c]			
	Liked most[a]		Liked least[a]		Liked most	Liked least	Liked most		Liked least	
	Unlimited	Limited	Unlimited	Limited	(Un)limited	(Un)limited	Unlimited	Limited	Unlimited	Limited
Hang around with	.87	.52***	.11	−.06*	.44	.04	.79	.11	.06	
	(.75, .98)	(.36, .69)	(.00, .23)	(−.22, .11)	(.37, .51)	(−.02, .11)	(.66, .91)	(−.01, .23)	(.00, .11)	
Not hang around with	−.13	−.10	.87	.89	−.07	.45	−.07		.45	
	(−.21, −.06)	(−.17, −.02)	(.80, .95)	(.81, .97)	(−.11, −.03)	(.41, .48)	(−.11, −.03)		(.41, .48)	
Cooperates	.44	.35	−.03	−.09	.25	−.01	.25		−.01	
	(.26, .62)	(.17, .53)	(−.21, .15)	(−.27, .09)	(.15, .35)	(−.11, .08)	(.15, .35)		(−.11, .08)	
Not play or work with	−.26	−.14	.48	.57	−.12	.26	−.12		.26	
	(−.41, −.10)	(−.29, .02)	(.32, .64)	(.42, .73)	(−.21, −.03)	(.18, .34)	(−.21, −.03)		(.18, .34)	
Plays alone	−.34	−.14†	.20	.33	−.15	.13	−.15		.13	
	(−.52, −.16)	(−.32, .04)	(.02, .39)	(.15, .52)	(−.25, −.05)	(.04, .22)	(−.25, −.05)		(.04, .22)	
Hits or kicks	.06	.04	.90	.88	.04	.46	.04		.46	
	(−.04, .16)	(−.06, .13)	(.80, 1.00)	(.78, .98)	(−.02, .09)	(.41, .51)	(−.02, .09)		(.41, .51)	
Gossips or excludes	.08	.05	.82	.78	.05	.41	.05		.41	
	(−.04, .21)	(−.07, .18)	(.69, .94)	(.66, .91)	(−.02, .12)	(.35, .48)	(−.02, .12)		(.35, .48)	
Calls names	.05	.01	.83	.81	.02	.42	.02		.42	
	(−.06, .17)	(−.11, .13)	(.72, .95)	(.69, .92)	(−.04, .09)	(.37, .48)	(−.04, .09)		(.37, .48)	
Starts fights	.01	.01	.88	.88	.01	.45	.01		.45	
	(−.08, .11)	(−.09, .10)	(.79, .98)	(.79, .98)	(−.04, .06)	(.41, .50)	(−.04, .06)		(.41, .50)	
Disrupts class	.12	.05	.60	.57	.06	.31	.06		.31	
	(−.04, .29)	(−.11, .21)	(.44, .77)	(.41, .73)	(−.03, .15)	(.22, .39)	(−.03, .15)		(.22, .39)	
Talks badly	.07	.01	.84	.81	.03	.43	.03		.43	
	(−.05, .18)	(−.10, .12)	(.73, .96)	(.70, .93)	(−.03, .09)	(.37, .48)	(−.03, .09)		(.37, .48)	
Leader	.37	.13*	.71	.57	.16	.34	.37	−.03	.75	−.06
	(.22, .52)	(−.03, .30)	(.56, .86)	(.40, .73)	(.07, .25)	(.26, .42)	(.22, .51)	(−.17, .11)	(.41, 1.10)	(−.40, .28)

[a] Significance of differences between both regression estimates obtained in two fully saturated models (unlimited versus limited).
[b] Regression estimates are equal between both methods as all regression paths were constrained in the constrained model.
[c] Differences in regression estimates for three parameters freed in final model.
† $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.

Downloaded from jbd.sagepub.com at University Library Utrecht on March 18, 2015

CFI = 1.000, SRMR = .017. Table 3 presents the standardized estimates of this final model. The parameter estimates are presented in the unlimited method column only, because they were equal between methods.

The method differences in the freely estimated regression weights show that the unlimited method resulted in significant estimates for hang around with regressed on liked most, and leader regressed on liked most and liked least. In the limited method, these estimates were smaller and not significant.

Discussion

The main goal of this study was to directly compare the use of unlimited versus limited nominations in sociometric research. It was hypothesized that the relative advantage of unlimited nominations is larger for positive sociometric questions and for sociometric questions related to social status and general reputation (e.g., liking, popularity). Multiple analyses were conducted to assess differences and similarities between both methods (Goal 1). A major emphasis was placed on the validity of both methods by analysing correlations between the six sociometric questions assessed with both methods and the 12 peer reputation questions assessed with unlimited nominations (Goal 2). The unlimited method was expected to outperform the limited method in this respect. Additional path analyses were conducted to analyse the predictive relationships between social preference measures (liked most and liked least) assessed with both methods and peer reputations assessed with unlimited nominations (Goal 3).

Inter-method differences in distributional properties

In line with previous research (Terry, 2000) and our expectations, the unlimited method provided more variability and less skewness and kurtosis than the limited method, especially for liked most, liked least, most popular, and least popular. Since these questions are critical for sociometric classification and the assessment of popularity, more favourable distributional properties (i.e., a closer approximation of normality) will result in more accurate status determination. As such, for these constructs, ecological validity seemed higher for scores derived from the unlimited method than the limited method (Goal 1).

Typically, exerting more external control in a study, for instance constraining participants in the number of nominations allowed, results in less ecological validity and lower generalizability from the controlled setting to the "real world." More skew in the distribution of sociometric scores in the limited method is more an artefact of being constrained than a "true" representation of the distribution in the population in their natural environment. Once the most noticeable peers are chosen, the remaining peers will no longer be considered for nomination (Parkhurst & Asher, 1992), leading to a considerable loss of data and inability to make inferences (Babad, 2001). Since the assumption of normality underlies many of the statistics used to study the validity of a construct, a closer approximation of normality is better. When we collect this type of sociometric data, there is a tendency for variables to be not normally distributed and it is thus important to examine the distributional properties of variables from all sociometric methods.

Overall, distributional properties for unlimited data were generally more favourable than for limited data. However, this was only the case for questions related to social status and general reputation (e.g., liking, popularity). In contrast, distributional properties for the negative bully questions were more favourable in the limited method. For the negative victim question, distributional properties were more favourable in the limited method only for boys.

Inter-method differences in number of nominations received

Mean differences between methods were significant though small for most questions. Nominations received for bully and victim did not differ between methods for girls. However, nominations received for liked most were substantially higher in the unlimited method for both boys and girls. This substantiates the hypothesis that for positive, status-related questions, the unlimited method was more valid than the limited method (Goal 1). Hence, one can assume that for these types of questions, participants know (many) more peers than three (the maximum in the limited method). These results were as expected and confirm what has been shown previously in an indirect comparison of both methods (Jiang & Cillessen, 2005).

Restricting participants in the number of nominations for positive sociometric questions will lead to measurement error (Holland & Leinhardt, 1973; Hymel & Asher, 1977) and lower construct validity, for instance when participants want to nominate four equally liked friends and only three nominations are allowed. Participants then are forced to leave one friend out at random, which would result in a distortion of their "true" evaluation of their peer relationships (lower construct validity). Further, due to measurement error, effect sizes are attenuated and parameter estimates are less accurate and consistent.

Inter-method correlations

All inter-method correlations were significant and moderate to large in size. However, the considerably lower inter-method correlation for liked most suggests that, at least for this type of question (i.e., positive, status-related), both methods are not interchangeable (Goal 1). That is, a low inter-method correlation suggests that both methods do not measure exactly the same construct. This is comparable to the results Terry (2000) found concerning this sociometric measure. Because both methods were counterbalanced and assessed across a two-week interval, cross-over effects and changes in group dynamics cannot account for this relative lack of correspondence.

Lack of correspondence between methods was most pronounced for girls. For girls, the lower inter-method correlations were most likely due to a less restricted evaluation of peers in the unlimited method, which enabled them to give a more nuanced opinion of their peers. It appears that girls benefit most from being able to give a more detailed peer evaluation, which is in line with previous findings on gender differences in peer relationship processes (e.g., girls are more network- and connection-oriented than boys) and girls' greater concerns about receiving negative evaluations from peers than boys (see Rose & Rudolph, 2006).

In addition, girls were possibly more affected by the way the task was formulated. In contrast to boys, girls did not nominate a lot more peers for the majority of the questions in the unlimited method compared to the limited method (except for liked most). However, at the same time, lack of correspondence between methods was most pronounced for girls. The fact that unlimited nominations were allowed and explicitly expressed in the task instructions

Downloaded from jbd.sagepub.com at University Library Utrecht on March 18, 2015

may have resulted in girls thinking longer and more thoroughly about who to nominate, which may have led to differences in their choices between methods.

Correlations between sociometric variables

Overall, correlations between the six sociometric questions assessed with both methods and the 12 peer reputation questions assessed with unlimited nominations did not differ significantly between methods except for the correlation between liked most and hang around with (i.e., significantly larger in the unlimited method; Goal 2). However, it must be noted that this comparison between the unlimited and limited methods is to some degree unwarranted, since the 12 reputation questions were only assessed with unlimited nominations. A closer examination of the correlations by gender revealed some differences; absolute differences in correlation coefficients between methods appeared larger for girls, while for boys the absolute differences in correlations between both methods were small and not significant. This may well be explained by gender differences in peer evaluation tendencies (see Rose & Rudolph, 2006).

Differences in regression estimates

Given the significant mean and correlational differences between both methods, path analyses examined method differences in the predictive relationships between liked most and liked least and the reputation questions more precisely (Goal 3). Three significant method effects were found, confirming our hypothesis that the unlimited method would yield stronger predictive power than the limited method for relatively stable measures of social status. Liked most predicted hang around with, and liked most and liked least both predicted leader in the unlimited method but not in the limited method. These effects were expected because hang around with is a positive sociometric criterion and leader is a general reputation (that compounds specific behaviour). In other words, ''leader'' might be more a (relatively stable) sociometric role than a behaviour.

Limitations and directions for future research

The current study raised four additional questions for future research. First, we used a limited sample size and a limited age range including only elementary school children. Future research should extend the sample size and age range by including secondary school adolescents, as method effects may be moderated by age. The current study offered some indication of moderation by age (results not shown); future research should address this more thoroughly.

A second important and related direction for future research is the use of reference groups varying in size and the use of true unlimited questionnaires or quasi-unlimited questionnaires capped at a different number than nine. In the current study, the complete roster contained 112 names and cross-grade nominations were allowed. Results may differ if smaller or larger reference groups are used and if fewer or more nominations are allowed.

A third direction for future research is to extend the list of sociometric questions assessed with both methods. We expected the unlimited method to outperform the limited method on positive questions and questions of status and general reputation. This hypothesis was partially confirmed, specifically for positive, status-related questions and more pronounced for girls than for boys. The reverse method-effect, however, was found for the negative bully and victim questions. It is interesting to know whether the relative advantage of the limited method also applies to other negative questions that measure behaviour. Hence, it may be that the relative advantage of limited nominations only applies to bullying or only to questions depicting a certain (negative) social ''role'' expressed with nouns (e.g., loner, fighter) rather than verbs (e.g., ''who plays alone or keeps to themselves,'' ''who starts fights''), perhaps due to perceived temporal differences and/or differences in frequency or intensity; labelling someone as a ''loner'' is more rigorous and definite than saying someone ''plays alone or keeps to him/herself.''

Fourth, future research should investigate whether results are similar when the 12 peer reputation questions are assessed with both methods and not just with the unlimited questionnaire. The current study also could be extended by using, for example, teacher ratings of behaviours associated with peer reputation. Ideally, such additional peer reputation measures should be assessed separately from or in counterbalanced order with measures of social preference and popularity.

Conclusion and methodological implications

In general, analyses revealed comparable results for both methods. However, some significant and noteworthy differences were found between unlimited and limited nominations. First, unlimited nominations offered more favourable distributional properties (i.e., a closer approximation of normality) than limited nominations for social preference (liked most and liked least) and popularity (most popular and least popular). In contrast, limited nominations offered more favourable distributional properties for the bully and victim questions. Based on the results and in reference to previous research on the distinct behaviours of well-liked and rejected children (e.g., Cillessen & Mayeux, 2004; Coie, Dodge, & Coppotelli, 1982; Sandstrom & Cillessen, 2006), it is likely that for questions related to social status, ecological, construct, and criterion validity will be higher when unlimited nominations are used. The argument that unlimited nominations would decrease selectivity and differentiation between children for these types of questions, because participants would nominate an excessive number of peers, was not substantiated. Both boys and girls benefited from unlimited nominations, as they were less constrained and more able to give nuanced opinions about their peers regarding their social status.

To conclude, based on this study and previous research, it is advisable to use unlimited peer nominations for sociometric questions related to social status (i.e., preference, popularity). Unlimited nominations yield more favourable distributional properties for these types of questions, a more pronounced pattern of correlations (less noise), and as such more accurate status determination. Future research should examine whether unlimited nominations are also more advisable for other positive sociometric questions (e.g., friendship, acquaintanceship, cooperativeness) and sociometric questions related to a general reputation that compounds specific behaviour (e.g., leader, clown, coolness).

Although there are obvious advantages to the unlimited nomination method, the current study should not be interpreted as evidence that limited nominations are an invalid or unreliable assessment practice for these and other types of questions. That is, the current study did not provide substantial evidence to conclude that the

Downloaded from jbd.sagepub.com at University Library Utrecht on March 18, 2015

limited method is in such cases ineffective. However, the unlimited method may be advised for other sociometric questions as well, as it is likely that participants know more than three peers to nominate for a variety of questions. In those cases, restricting them to a fixed number may induce measurement error and reduce ecological validity.

Optionally, to improve the feasibility of unlimited nominations and to preserve some degree of voter selectiveness, the maximum number of nominations allowed could be capped at a fixed number or a percentage of reference group size, depending on data collection procedures (e.g., computerized vs. paper-and-pencil assessment, size of reference group, writing down names of peers vs. selecting peers from rosters).

The results from this study suggest the use of limited nominations for negative nominations related to bullying, although when nominations for these questions were unlimited, increases in the number of nominations and decreases in their selectivity were small, indicating that unlimited nominations would yield comparable results when used for bullying-related questions. Limited nominations may also be recommended for other negative nomination questions. Future research should determine whether the relative advantage of limited nominations also applies to other negative sociometric questions.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

Babad, E. (2001). On the conception and measurement of popularity: More facts and some straight conclusions. *Social Psychology of Education*, *5*, 3–29.

Bellmore, A. D., Jiang, X. L., & Juvonen, J. (2010). Utilizing peer nominations in middle school: A longitudinal comparison between complete classroom-based and random list methods. *Journal of Research on Adolescence*, *20*, 538–550.

Bukowski, W. M., & Cillessen, A. H. N. (Eds.) (1998). Sociometry then and now: Building on six decades of measuring children's experiences with the peer group. San Francisco, CA: Jossey-Bass.

Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. New York, NY: Routledge.

Caravita, S. C. S., Di Blasio, P., & Salmivalli, C. (2008). Unique and interactive effects of empathy and social status on involvement in bullying. *Social Development*, *18*, 140–163.

Cillessen, A. H. N. (2009). Sociometric methods. In K. H. Rubin, W. M. Bukowski & B. Laursen (Eds.), *Handbook of peer interactions, relationships, and groups* (pp. 63–61). New York, NY: Guilford.

Cillessen, A. H. N., & Mayeux, L. (2004). From censure to reinforcement: Developmental changes in the association between aggression and social status. *Child Development*, *75*, 147–163.

Coie, J. D., Dodge, K. A., & Coppotelli, H. (1982). Dimensions and types of social status: A cross-age perspective. *Developmental Psychology*, *18*, 557–570.

Hallinan, M. T. (1981). Recent advances in sociometry. In S. R. Asher & J. M. Gottman (Eds.), *The development of children's friendships* (pp. 91–115). New York, NY: Cambridge University Press.

Holland, P. W., & Leinhardt, S. (1973). The structural implications of measurement error in sociometry. *Journal of Mathematical Sociology*, *3*, 85–111.

Hymel, S., & Asher, S. R. (1977, March). *Assessment and training of isolated children's social skills*. Paper presented at the biennial meeting of the Society for Research in Child Development, New Orleans, Louisiana. Retrieved from ERIC database (ED136930).

Hymel, S., Vaillancourt, T., McDougall, P., & Renshaw, P. D. (2002). Peer acceptance and rejection in childhood. In P. K. Smith & C. H. Hart (Eds.), *Blackwell handbook of child social development* (pp. 265–284). Hoboken, NJ: Wiley-Blackwell.

Jiang, X. L., & Cillessen, A. H. N. (2005). Stability of continuous measures of sociometric status: A meta-analysis. *Developmental Review*, *25*, 1–25.

Maassen, G. H., van Boxtel, H. W., & Goossens, F. A. (2005). Reliability of nominations and two-dimensional rating scale methods for sociometric status determination. *Journal of Applied Developmental Psychology*, *26*, 51–68.

Moreno, J. L. (1934). *Who shall survive? A new approach to the problem of human interrelations*. Washington, DC: Nervous and Mental Disease Publishing Company.

Moreno, J. L. (1951). Sociometry, experimental method and the science of society. Beacon, NY: Beacon House.

Newcomb, A. F., & Bukowski, W. M. (1983). Social impact and social preference as determinants of children's peer group status. *Developmental Psychology*, *19*, 856–867.

Parkhurst, J. T., & Asher, S. R. (1992). Peer rejection in middle school: Subgroup differences in behavior, loneliness, and interpersonal concerns. *Developmental Psychology*, *28*, 231–241.

Paternoster, R., Brame, R., Mazerolle, P., & Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology*, *36*, 859–866.

Poulin, F., & Dishion, T. J. (2008). Methodological issues in the use of peer sociometric nominations with middle school youth. *Social Development*, *17*, 908–921.

Rose, A. J., & Rudolph, K. D. (2006). A review of sex differences in peer relationship processes: Potential trade-offs for the emotional and behavioral development of girls and boys. *Psychological Bulletin*, *132*, 98–131.

Rose, A. J., Swenson, L. P., & Waller, E. M. (2004). Overt and relational aggression and perceived popularity: Developmental differences in concurrent and prospective relations. *Developmental Psychology*, *40*, 378–387.

Sandstrom, M. J., & Cillessen, A. H. N. (2006). Likeable versus popular: Distinct implications for adolescent adjustment. *International Journal of Behavioral Development*, *30*, 305–314.

Santo, J. B., Mayman, S. B., Lopez, L. S., & Bukowski, W. M. (2012, March). *An analysis of the differences between using limited versus unlimited peer nominations of victimization among Colombian early adolescents*. Paper presented at the biennial meeting of the Society for Research on Adolescence, Vancouver, British Columbia, Canada.

SCP. (2006). *Socio-economic status in The Netherlands 2006* [Data file]. Retrieved from http://www.scp.nl/download/statusscores_zip

Terry, R. (2000). Recent advances in measurement theory and the use of sociometric techniques. *New Directions for Child and Adolescent Development*, *88*, 27–53.

Appendix A: Sociometric questions

Sociometric questions assessed with both methods

1. Who do you like most? (liked most)
2. Who do you like least? (liked least)
3. Who is most popular? (most popular)
4. Who is least popular? (least popular)

Downloaded from jbd.sagepub.com at University Library Utrecht on March 18, 2015

5. Who bullies other children? (bully)
6. Who is bullied by other children? (victim)

Additional peer reputation questions assessed with the unlimited method

1. Who do you hang around with? (hang around with)
2. Who do you rather not hang around with? (not hang around with)
3. Who is willing to help/cooperate with other children on a regular basis? (cooperates)
4. With whom does almost nobody want to play or work with? (not play or work with)
5. Who works or plays alone often? (plays alone)
6. Who hits, kicks or shoves other children? (hits or kicks)
7. Who gossips about other children or ignores or excludes other children? (gossips or excludes)
8. Who calls other children names? (calls names)
9. Who starts fights with other children? (starts fights)
10. Who disrupts class on a regular basis? (disrupts class)
11. Who talks badly about other children? (talks badly)
12. Who is a leader or often takes the lead? (leader)

Downloaded from jbd.sagepub.com at University Library Utrecht on March 18, 2015