Proxy Reports and Their Value

by Gordon Scott Bonham, Ph.D.

Bonham Research

July 10, 2006

Abstract

The quality of life of people with intellectual disabilities is difficult to measure, and questions have been raised about both self-responses and proxy responses. The FY2004 Ask Me! Survey collected data from 1,091 self-respondents, and 764 proxies for 407 people who could not respond for themselves. Peer interviewers decided who could respond for themselves. Self-respondents answered more questions, answered the same questions as consistently, and had more reliable scales than did proxies. Two proxies for the same person agreed most on emotional well-being and least on self-determination. Two day staff proxies agreed most; family and staff proxies agreed least. Self and proxy response data can be combined for most analysis with the inclusion of a statistical control for who responded.

Background

Research since the late 1980's has asked people with intellectual disabilities to express their own views about their quality of life (Schalock and Verdugo, 2002). Their increased participation, however, has raised four basic issues: 1) who determines the capability of people to respond for themselves, 2) can self-respondents provide valid information, 3) can proxies provide valid information, and 4) how should those who cannot respond for themselves be represented?

Sometimes staff or families who support people with intellectual disabilities decide if they can understand and answer questions (Rapley, Ridgway and Beyer, 1997; Stancliffe, 1999). Frequently non-disabled interviewers make the decision of who can respond for themselves, whether these are professional interviewers of the general population (Research and Training Center on Community Living, 2004) or interviewers specifically trained to interview people with intellectual disabilities (Sigelman, et al., 1980). Perry and Felce (2004) found that an interviewer with intellectual disabilities could collect information as effectively as a person without intellectual disabilities. Sometimes the decision-maker has a set of pre-interview questions to help them determine who can adequately respond for themselves (Sigelman et al., 1980; Perry and Felce, 2002; Stancliffe, 1995). Often the basis for deciding who can respond is not specified (Research and Training Center, 2004). The explicit or implicit decision criterion about who is able to respond for themselves affects the information that will be obtained for those individuals. Siegelman, et al. (1980) found that only three of twenty people with profound retardation could respond at all to interviews, and dropped those with profound retardation from the study. Perry and Felce (2002) determined that none of those with the least intellectual ability could respond for themselves, while all those with average ability could. The National Health Interview Survey

had responses for 60% of people with mental retardation before interviewers were allowed to use proxies (Research and Training Center on Community Living, 2004). After interviewers could use proxies, representation increased to 90% but only 41% responded for themselves. Rapley (2000) suggests that quality of life is a social construct and its measurement is an interaction between the interviewer and interviewee. Participatory action research calls for involvement of people in research affecting them (Whitney-Thomas, 1996). The Ask Me! Survey began with two premises: 1) people with developmental disabilities should be asked directly about their own life, and 2) interviewers with developmental disabilities were in the best position to elicit meaningful responses from their peers (Bonham, Basehart, Schalock, Marchand, Kirchner and Rumenap, 2004; Schalock, Bonham and Marchand, 2000). The survey used peer interviewers to decide if people had the ability to respond for themselves, avoiding potential bias of staff who might select those that might make their agency look good, and potential bias of families who might underestimate abilities. This paper presents some results to address Research Question One, *can peer interviewers make appropriate judgements of the abilities of people to respond for themselves.*

A number of researchers have studied the reliability of answers given by people with intellectual disabilities (Chung, Yu, Martin, Havapiak and Garinger, 2000; Heal and Sigelman, 1996; Perry and Felce, 2002; Sigelman et al., 1980; Stancliffe, 1995). They have looked at acquiescence, first and last response bias, and stability of answers over time. Suggestions to reduce these problems involve simple wording in questions, having more than two fixed responses, balance between positive and negative wording, and use of pictures and aids (Finlay and Lyons, 2002; Heal and Sigelman, 1996; Sigelman, Budd, Winer, Schoenrock and Martin, 1982). Antaki and Rapley (1996) suggest the problems lie with the interview context, rather than

just the person with intellectual disabilities, as both interviewers and respondents work and rework questions and answers. Sigelman et al. (1982), Heal and Sigelman (1996), and Perry and Felce (2002) applied four tests in determining validity of self-responses: 1) the proportion of questions answered, 2) the correspondence between answers to the same question on two occasions, 3) the consistency of answers with the same meaning but with different words or format, and 4) the agreement of information from various sources. The Ask Me! Survey was designed to maximize reliability and validity of self-response, rather than to rigorously apply the four tests. However, the information it provides can address Research Question Two, *can people with intellectual disabilities provide reliable and valid data on their quality of life?*

Studies on the reliability of responses by people with intellectual disabilities often compare them to responses of proxies, without first testing proxy responses for reliability (Chung et al., 2000; Heal and Sigelman, 1996; Stancliffe, 1995; Perry and Felce, 2002). Lunsky and Benson (1997) found that staff members did not respond any more consistently across a number of instruments than did people with mild mental retardation. The National Health Interview Survey (Research and Training Center, 2004) found self-respondents skipped questions less frequently than proxies, particularly on subjective questions. Lunsky and Benson (1997) and Stancliffe (1999) found that staff members tended to agree more with each other than with consumers. Gaudet, Pulos, Crethar and Burger (2002) found low correlations between family members and provider staff, similar to the low correlations between self-responses and proxy-responses. Umb-Carlsson (2005) found relative and staff proxies contributed dissimilar information. Cummins (2002), Dudley (2001), and Perry and Felce (2002) found that the more subjective the measure, the lower the correlation between self and proxy respondents. Cummins (2002) reviewed the literature about proxies responding for subjective well-being and concluded

that proxy responses for people with severe disabilities "cannot be regarded as valid under any circumstances." Proxies have limited information on how individuals who cannot respond to questions and express themselves think. In these situations, proxies cannot escape projecting their own thoughts and prejudices. Cummins argues that any reported reliability between self-respondents and proxies, or among proxies, is a result of shared cognition and life experiences rather than specific response validity. As a result, staff and caregivers may be the least reliable proxies in those areas that reflect upon their responsibilities. Two proxies are frequently used, and Stancliffe (1999) found that the average of two staff proxies agreed better with self-response than a single staff response. Cummins (2002), however, argues that two proxies double the response effort without increasing validity. The Ask Me! Survey attempted to collected information from two proxies for people who could not respond for themselves. It cannot address questions of self and proxy concordance, nor if two proxies are better than one. However, it can apply the same tests of validity to proxy as self-reporting to address Research Question Three, *are staff and family any more reliable and valid sources of information on quality of life than the person with intellectual disabilities?*

Cummins (2002) argues that proxy responses are least valid for people who cannot respond for themselves and should not be used to make decisions concerning a person's life. Stancliffe (2000) also cautions against projecting findings on the validity of proxy reporting from people who can respond for themselves to those who cannot respond for themselves. Parsons, Baum and Johnson (2000), however, argue that a proxy respondent is preferable to complete non-response, especially when the person with intellectual disabilities is willing to participate. Cummins (2002) concedes that proxy responses may be acceptable when the primary objective is to reflect generally shared views. Rapley et al. (1997) suggest that simple regression adjustments

can be made in some domains for proxy responses, while Stancliffe (2000) cautions that not enough is known to assume a simple linear adjustment is adequate. The Ask Me! Survey was designed to influence system change and provide agencies with information on consumers' quality of life to help them enhance programs. It was also intended to help people with intellectual disabilities and their families select the most appropriate service providers. Not collecting proxy information would provide a limited view of the quality of life in Maryland, would not help agencies that support large numbers of people with the most severe disabilities, and would not help families, of those with the most severe disabilities, choose service providers. This paper addresses Research Question Four, *can data from self-responders and proxies provide sufficiently similar information on quality of life that they can be combined with reasonable statistical controls?*

The Maryland Developmental Disabilities Administration (DDA) contracts with The Arc of Maryland to conduct the Ask Me! Survey. The survey collects information on the quality of life each year for a sample of about 1,500 adults with developmental disabilities who receive support in the community. It was designed to maximize the probability of self-response and reduce the probability of unreliable responses in several ways: 1) using peer interviewers, 2) asking questions important to self-advocates, 3) using direct and simple language (average of 9.9 words per questions, 1.3 syllables per word, and rated by WordPerfect at a 3.6 grade level with the lowest verbal complexity score of 1.0); 4) providing three fixed responses for each question (favorable, neutral and unfavorable); and 5) associating happy, neutral and sad facial representation with the three responses (Bonham, et al., 2004). Annual reports have consistently shown significant differences in the quality of life as reported by self-respondents and by proxies (Bonham, Basehart & Marchand, 2002, 2003, 2004). Data were collected in FY2004 about the

type of proxies to better understand the observed differences.

Methods

The FY2004 Ask Me! Survey collected data for 1,498 adults out of a population of 12,971 people supported by DDA. Of these, 1,031 responded for themselves without any help, 60 responded with someone providing help, 357 had two proxy respondents, and 50 had one proxy respondent. Unless specifically noted, this paper excludes people responding for themselves with help from a staff or family member, as analysis suggested their responses displayed a blend of self-response and proxy response. The proxy respondents included 326 day habilitation or employment support direct-care staff, 195 residential direct-care staff, 123 family members or friends, 107 other support staff (trainers, case mangers, supervisors, nurses, etc.), and 13 resource coordinators from an agency that provides no other services. This paper excludes the small number of service coordinator proxies, and combines other support staff with day habilitation and employment direct-support staff into a single category of day staff. The 357 proxy-pairs include 133 sets of day staff and residential staff, 92 sets of day staff and family members, 20 sets of residential staff and family members, 82 sets of two day staff, 17 sets of two residential staff, 9 resource coordinator with some type of staff, 3 sets of two family proxies, and 1 family and resource coordinator. The last three groups are excluded from the pair-analysis because of their small numbers. A category of two similar staff combines the sets of two day staff and of two residential staff.

Half (55%) of all the sampled individuals were male; their mean age was 39.8 years, s.d. = 13.3; 7% had autism; and 22% had epilepsy and seizure disorders. Agency staff reported the level of retardation for 80% of the people: 12% had profound retardation, 13% severe, 23% moderate, 26% mild, and 7% had borderline or no retardation.

Forty adults were randomly selected from DDA files for each of 48 community agencies at the beginning of FY2004. The agencies included all ten that supported 300 or more adults, half of the 22 agencies that supported 130-299 adults, one-fourth of the 40 agencies that supported 55-129 adults, and 17 of the 42 agencies that supported 10-54 adults. The sample frame excluded adults who received all their services from state institutions or from small agencies that supported fewer than ten adults. Ask Me! staff sent each agency their sample list, so that the agencies could contact the individuals or their guardians, obtain general consent, and schedule the interviews. Guardians for 2% of the selected individuals refused to let them participate, 10% of the selected individuals refused for themselves, and 10% were not interviewed due to illness, not receive services, moving from the state, or whereabouts unknown. The response rate, including proxies, varied among agencies from 43% to 97%. Most (76%) of the interviews occurred at day habilitation or employment locations, while 12% occurred at residences and 12% occurred at other locations or over the telephone. An Arc of Maryland supervisor explained the survey background and purpose to the interviewees, and then assigned them to two-person teams of peer interviewers. The peer interviewers took each person to a private room to ensure confidentiality and reduce the likelihood of affirmative answers to please someone on whom they depended. Staff or family could accompany the individual only at the individual's request, and only to act as a translator. The 35 peer interviewers during FY2004 had an average of 4.0 years of interviewing experience with Ask Me!

The institutional review board for the DDA required that each person interviewed had the ability to consent to the interview. Peer interviewers determined ability to consent by asking six pre-interview questions:

1. Do you understand you will be answering questions? (Yes, no)

2. Do you understand you can skip questions if you do not want to answer them? (Yes, no)
3. Do you understand you can stop the interview at any time? (Yes, no)
4. Let me ask you a question from the interview. Would you say that you are a happy person? (Yes, sometimes, no)
5. Would you like to answer more questions? (Yes, no)
6. Do you understand that you will be answering questions about your life? (Yes, no)

These questions effectively determined who could respond to the interview for themselves. If the interviewers believed the interviewees understood and answered these screening questions, the interviewers signed the form and proceeded with the interview. If they believed the interviewee did not understand and could not communicate, they notified their supervisor to identify (with the agency's help) two proxies to interview. The self-response rate at agencies varied from 24% to 100%, reflecting differences, among agencies, in the intellectual abilities of the people they supported.

The Ask Me! Survey has 56 questions, six for each of eight quality of life domains identified by Schalock and Verdugo (2002), five questions on the perceived availability of transportation, and three repeated questions to check internal consistency of responses. Each question could be answered with a positive response (☺ 1), a neutral response (😐 2) or a negative response (☹ 3). Respondents could answer with words, by general gesture, or by pointing to a face on a card associated with their answer. An identically worded question on earnings, in the material well-being domain, occurred 13 questions after its first location ("Do you have the chance to earn good money?"). A second questions on general happiness, in the emotional well-being domain, had only slight differences in wording and was placed 43 questions after the original ("In general, how happy are you with you life?" and "How happy are you with your life

overall?"). A second question about house-mate choice, in the self-determination domain, addressed the same concept using different words and was placed 23 questions after the original ("How much choice did you have in whom you live with?" and "Did you pick who you live with?"). A *Disagreement Score* was calculated as the absolute difference between two answers a responder gave to the same or similar questions, or two proxies gave to the same question: 0 when the answers were the same, 1 for a favorable/neutral or unfavorable/neutral combination, and 2 for a favorable/unfavorable combination. Disagreement scores were not calculated if responses to one or both questions were missing.

Most self and proxy interviews occurred at the site of the day or employment program, and took about 30 minutes. The first proxy interviews generally took place immediately after peer interviewers determined that people could not respond for themselves. Therefore, 77% of the first proxy interviews were with day staff, e.g., direct support staff, supervisors, trainers, case managers and nurses. The second proxies were more diverse: 36% were residential staff, 34% were day staff, 28% were family or friends, and 2% were service coordinators. Interviewers conducted 83% of the first proxy interviews face-to-face and conducted 56% of the second proxy interviews over the telephone. The second proxy interview occurred an average of 30 days after the first proxy interview. Some differences between proxies could be due to different modes of interviewing and different time frames.

The analysis used SPSS (Statistical Package for the Social Sciences) Scale Reliability to calculate Cronbach alphas for each of the eight quality of life domains. It used independent sample t-tests to compare means and percents at a 0.05 level of statistical significance. Forward stepwise multiple regressions used 0.05 as the criteria for entrance and 0.10 for variable removal. Missing data were excluded pairwise. Data shown in this analysis were not weighted, although

people were selected with different probabilities. The findings address the four research questions.

Ability to Respond for Self

Peer interviewers determined that 69% of the people had the ability to respond for themselves and 4% could respond with help. Non-disabled interviewers for the National Health Interview Survey found only 60% of people with mental retardation could respond to their interview (Research and Training Center, 2004). Almost everyone classified with mild retardation, borderline retardation or normal intellectual ability responded for themselves to the Ask Me! Survey. (See Figure 1.) Complete self-response became less frequent as the level of intellectual ability decreased: 39% of people classified with severe retardation and 8% of those classified with profound retardation responded for themselves. Using a more precise measurement of intellectual ability, Perry and Felce (2002) found that 100% with Adaptive Behavior Scale scores above 90 could respond to the pretest questions, while no one with scores below 21 could respond. The Ask Me! findings suggest peer interviewers may make it possible for a high percent of people to respond for themselves, but self-response will still remain affected by the individual's level of intellectual ability.

INSERT FIGURE 1 ABOUT HERE

Stepwise multiple regression showed that almost half of the variability, *adjusted* R^2 = .482, $F(4, 1151) = 269.94$, $p < .001$, in self-response could be explained by four characteristics. Intellectual ability, measured by five levels of retardation, was the most important, $\beta = .25$, $\Delta R^2 = .467$, $p < .001$. Dichotomous measures of the presence of epilepsy and seizure disorders ($\beta = -.11$, $\Delta R^2 = .009$, $p < .001$), and autism ($\beta = -.14$, $\Delta R^2 = .003$, $p = .001$) decreased self-response. Six categories of age, primarily in 10-year groups, showed that self-response decreased with

older age (β = -.02, ΔR^2 = .003, p = .005). Dichotomous variables for other types of disabilities (language impairments, visual impairments, hearing impairments, orthopedic impairments, mental disorders, behavioral problems, cerebral palsy, and other neurological impairments) had no relation to self-response.

Validity of Response

Self-respondents answered an average of 49.8 of the 53 non-duplicated questions, with 67% answering all questions. Self-respondents answered significantly more questions than did day staff (48.4, $t(1462) = 2.85$, p = .004), significantly fewer questions than did residential staff (51.2, $t(637) = 3.21$, p = .001), and about the same number as did family members (49.3). Only 35% of day staff, 54% of residential staff and 46% of family answered all the questions. On the test of the proportion of questions answered, people judged by peer interviewers as being able to respond for themselves provided as valid, or more valid, data as provided by the majority of proxies.

Self-respondents gave the favorable, and first, response to 33.8 of the 49.8 questions they answered. Day staff gave the favorable response to 30.8 questions, fewer than self-respondents, $t(1191) = 5.60$, p = .001. Residential staff gave the favorable response to 35.4 questions, more than self-respondents, $t(483) = 2.64$, p = .008. Family members gave the favorable response to 34.2 questions, about the same as self-respondents. These averages do not suggest high levels of acquiescence, or primacy bias, among self-respondents in comparison with proxies. However, the percent of self-respondents who answered 49 or more of the questions with the favorable response does suggest some acquiescence, particularly among those with the least ability: 21% of self-respondents with profound retardation, 9% with severe, 8% with moderate, 8% with mild, and 4% of self-respondents with borderline or no retardation. Only one day staff answered 49 or

more questions with the positive response. Acquiescence may be a problem for some people with intellectual disabilities, but favorable responses do not seem to be higher for self-respondents than for proxies.

Self-respondents gave the unfavorable answer to 7.3 questions on average. They gave more unfavorable responses than did residential staff (6.4 questions, $t(644) = 2.63$, $p = .009$) or day staff (6.6 questions, $t(1260) = 2.18$, $p = .030$), but no different from family (7.3 questions). No one, self-respondent or proxy, gave the unfavorable response to 49 or more questions. Nay-saying or recent bias does not appear to be a problem for anyone, but self-respondents and family members did see things more negatively than did staff.

Self-respondents had disagreement scores of .25 on the pair of questions with identical wording and .26 on the pair of questions with similar wording. (See Table 1.) They had disagreement scores of 0.51 about house mate choice where the wording differed substantially. Disagreement scores of self-respondents did not consistently vary with intellectual ability, although the few self-respondents with profound retardation tended to give different answers to pairs of questions more frequently than did other self-respondents. Proxies were statistically more consistent than self-respondents in answering the pairs of questions about earning good money, $M = .16$, $t(1545) = 3.83$, $p < .001$, and happiness with life, $M = .18$, $t(1643) = 3.75$, $p < .001$, but were no more consistent in answering the two questions on house mate choice ($M = .48$). In addition, while more proxies (94%) than self-respondents (91%) answered both questions on happiness ($t(1741) = 2.17$, $p = .030$), fewer proxies than self-respondents answered both questions on earning good money (85% and 89%, $t(1486) = 2.43$, $p = .015$) and house mate choice (84% and 92%, $t(1322) = 5.01$, $p < .001$). Only 79% of day staff could answer the questions on picking house mates.

INSERT TABLE 1 ABOUT HERE

Six questions comprised each quality of life scale. Scale reliability, measured by Cronbach's alpha, ranged from 0.62 to 0.73 for self-respondents on the eight quality of life domain scales, averaging 0.67. (See Table 2.) Self-respondents with severe or profound retardation had an average scale reliability of 0.73, higher than those with moderate (0.67), mild (0.64) and borderline retardation (0.64). Proxies had an average scale reliability of 0.55, ranging from 0.44 to 0.63, and always lower than self-respondents. Day staff proxies had the lowest average scale reliability (0.44), but had the same or higher scaling reliability than self-respondents on rights and self-determination. Family members had the highest scale reliability among proxies (0.53) and had the same or higher reliability as self-respondents on personal development, material well-being and emotional well-being.

INSERT TABLE 2 ABOUT HERE

Proxies for people who could not respond for themselves reported a higher quality of life in four of the eight domains than did self-respondents, and lower quality of life in three domains. On a scale of 0 to 10, proxies reported almost perfect (M = 9.33) physical well-being of the person for whom they were responding, significantly higher than self-respondents reported, 8.46, $t(1511) = 12.41$, $p < .001$. This is consistent with the higher reporting of proxies than self-respondents on feelings of safety that Dudley (2001) found, and safety in the neighborhood was one question in the Ask Me! physical well-being scale. Proxies also reported higher quality of life than did self-respondents in the domains of emotional well-being (9.02 and 8.37, $t(1745) = 8.67$, $p < .001$), material well-being (7.76 and 7.50, $t(1541) = 2.72$, $p = .009$) and interpersonal relations (8.14 and 7.79, $t(1693) = 3.90$, $p < .001$). (See Table 3.) Among self-respondents, the quality of life in these four domains was not correlated with intellectual ability. Proxies reported

significantly lower quality of life than did self-respondents in three domains: personal development (7.23 and 7.75, $t(1694) = 5.40$, $p < .001$), self-determination (6.19 and 7.62, $t(1569) = 13.21$, $p < .001$), and rights (4.99 and 6.93, $t(1456) = 16.16$, $p < .001$). Among self-respondents, those with lower intellectual ability reported lower levels of self-determination ($r = .10$, $p = .004$), but reported higher levels of social inclusion ($r = -.08$, $p = .021$). Thus only in the domain of self-determination do the data suggest a relationship between quality of life and intellectual ability that is reflected in both self and proxy responses, assuming people who cannot respond for themselves have the lowest level of intellectual ability. Stancliffe (1995) also found that self-respondents reported higher levels of self-determination than proxies reported. Acquiescence could be a problem among the 14 people with profound retardation that peer interviewers judged capable of responding for themselves, as they reported statistically higher quality of life than did self-respondents with severe retardation in three of the eight domains: physical well-being, interpersonal relations, and social inclusion. The pattern of increasing favorable reporting with decreasing intellectual ability, that acquiescence would produce, was generally observed among self-respondents only in the domain of social inclusion. In this domain, however, proxies for those with the least intellectual abilities reported no differently than did self-respondents.

INSERT TABLE 3 ABOUT HERE

Validity of Proxy Information

Two proxies for the same person should agree on answers to questions if they are reporting accurate information specific to the person. The Ask Me! Survey had 357 pairs of proxies, with 344 pairs that can be classified into three groups large enough for separate analysis. The two proxies for the same person had an average disagreement score of .37 across all survey

questions. This is approximately equal to both proxies answering two-thirds of the questions the same way, and one proxy answering the remaining questions with the neutral response while the other answered with a different response, most frequently the favorable one. Two proxies agreed most on individual questions in the domains of physical (.16) and emotional (.23) well-being. Proxies disagreed with each other most in the domains of rights (.50) and self-determination (.60). Disagreement scores between proxies on questions in the other four domains ranged from .32 to .38. For the eight domain scale scores, the Pearson correlations between the two proxy responses ranged from r = .21 for interpersonal relations to r = .37 for physical well-being, all statistically different from zero at $p < .01$, but all substantially smaller than Stancliffe (1999) found, and smaller than four of the five that Rapley et al. (1997) found.

The various combinations of proxy pairs had about the same low levels of disagreement in the domains of physical and emotional well-being. (See Figure 2.) However, family and staff proxies disagreed more than did two similar staff in the other six domains: material well-being (difference M = .09, $t(209) = 2.01$, $p = .046$), interpersonal relations (difference M = .10, $t(209) = 2.09$, $p = .002$) social inclusion (difference M = .09, $t(209) = 2.47$, $p = .014$), personal development (difference M = .12, $t(209) = 3.06$, $p = .002$), self-determination (difference M = .24, $t(209) = 5.10$, $p < .001$), and rights (Difference M = .17, $t(209) = 3.42$, $p = .001$). Family and staff proxies had statistically higher disagreement scores than did residential and day staff proxies in four domains: material well-being (difference M = .09, $t(243) = 2.22$, $p = .027$), social inclusions (difference M = .13, $t(243) = 3.47$, $p = .001$), self-determination (difference M = .12, $t(243) = 2.61$, $p = .010$), and rights (difference M = .10, $t(243) = 2.27$, $p = .024$). Residential and day staff disagreed more than two similar staff in the domain of self-determination (difference M = .11, $t(230) = 2.34$, $p = .020$).

INSERT FIGURE 2 ABOUT HERE

Proxy responses to three individual questions with the greatest proxy disagreement (.65-.71) illustrate different perspectives by different types of proxies. Two-fifths (46%) of day staff reported that the individuals could be alone when they wanted, compared to 66% of residential staff and 74% of family. Two-thirds (68%) of day staff reported it was easy for the person to say something when they had a problem with staff, compared 57% of residential staff and 20% of family members. On job training, 31% of day staff, 43% of residential staff and 71% of family proxies said the person received what was needed. These findings suggest that proxies report from different perspectives, and the more dissimilar their perspectives, the more dissimilar their reports.

Combining Self and Proxy Data

If two proxies for the same person report differently, and both report differently from self-respondents, does it make sense to include proxy with self-reported data? Excluding proxy information means that no information is available about the quality of life of people with the least intellectual abilities and unable to respond for themselves. Table 4 shows the results of three linear regressions for each quality of life domain on intellectual ability: self-respondents only, proxy respondents only, and self and proxy responses combined with a dichotomous variable to indicate proxy responses. Self-reported data estimated the physical well-being of persons with profound retardation as 8.18 on a scale from 0 to 10. Higher levels of intellectual ability had little relation to physical well-being. Proxy data estimated the physical well-being of persons with profound retardation as 9.37, but showed a similar insignificant relation between intellectual ability and physical well-being. The combined linear regression reflected the same insignificant relation to intellectual ability. The combined data estimated the physical well-being

of people with profound retardation as 8.35 for self-respondents (Proxy = 0) and 9.29 with proxy data (Proxy = 1), a statistically significant difference. Thus, the combined regression that identified who responded yielded almost the same information as did separate analysis of self and proxy data. Estimates from the combined data may be more valid as they represented the full range of intellectual abilities, whereas self-response data most reflected adults with moderate and less retardation, while proxy data most reflected adults with profound and severe retardation.

INSERT TABLE 4 ABOUT HERE

The quality of life in three other domains, emotional well-being, material well-being and interpersonal relations also had no relation to intellectual ability, whether data from self-respondents and proxies were analyzed separately or combined. Self-determination increased significantly with intellectual ability, regardless of whether self and proxy responses were analyzed separately or combined.

The validity of combined data was less clear for three quality of life domains. Perceptions of social inclusion declined with intellectual ability among self-respondents and with the combined data, but had no relation to intellectual ability when proxies reported. Personal development had no relation to intellectual abilities among self-respondents and in the combined data, but proxy data showed a significant relation. In the domain of rights, the combined data reflected the significant relation with intellectual ability found in proxy data that was not found in self-responses.

Discussion

The Ask Me! Survey began with the assumption that well-trained peer interviewers had the best ability to decide if a person with intellectual disabilities was capable of understanding questions and expressing views on his or her own quality of life. An Ability to Consent Form,

required by the institutional review board for the protection of human subjects, provided a standard procedure to help the interviewer team make the decision. It also reminded them of the importance of making a correct decision. Peer interviewers had a stake in producing accurate information that might result in system changes that affect them. The percent of the Ask Me! sample who respond for themselves appears to be as high as or higher than in other studies, suggesting that peer interviewers may facilitate self-response better than other interviewers. However, further investigations like that of Perry and Felce (2004) are needed to empirically test the ability and value of peer interviewers. Having peer interviewers decide who can and cannot respond for themselves had the additional value of taking this decision out of the hands of agency staff. It reduced the possibility of agency staff deliberately or unconsciously affecting the survey outcomes, and insulated agencies from suspicions of bias when the quality of life among the people they support is compared to the quality of life of people supported by other agencies. The Ask Me! procedures, however, still relied on agency staff to inform the individuals, or their guardians, about the survey and to secure cooperation. The freely expressed choice of an individual not to participate must be honored as an expression of their rights. However questions can be raised whether staff-reported refusals truly reflected the informed choices of people or their guardians, or reflect a lack of interest by caregivers who might be affected by the responses? Additional study is needed on reasons and effects of non-participation, what might increase participation, and how to encourage self-response once their willingness to participate has been established.

This study presented three types of evidence for the reliability of self-respondents' answers. The first evidence was interview completion. Two-thirds of the self-respondents answered every single question and provided more complete data than day staff proxies. A

second measure of reliability compared answers to repeated questions. About three-fourth of the self-respondents gave the same answer to two questions with the same or similar wordings, and about one-half gave the same answer to two questions with the same meaning but different words. This type of test has been used to judge self-respondents as incapable of providing useful data. However, while proxies were a little more consistent in answering questions with the same wording, proxies were not more consistent than self-respondents in their answers to the same concept with different wording. Few would suggest that the families and staff are incapable of providing useful information. People with intellectual disabilities should not be held to a higher standard of reliability than family and staff. Scale reliability provided a third measure on reliability of responses–how similarly the person responded to the set of six questions designed by the investigators as indicators of a quality of life domain. The scales had greater statistical reliability for self-respondents than for proxies. This may reflect that the indicator questions were derived from a pool of questions identified as important by self-advocates (People on the Go, 1996) rather than from questions family or staff might consider important.

Reliability of data does not guarantee its validity, and the Ask Me! Survey collected no external data to test the validity of either self-response or proxy response. It incorporated many aspects suggested by previous survey research to increase the probability of validity, yet found some indications of acquiescence among self-respondents with the least intellectual ability. Rapley (2000) and Papley and Antaki (1998) argue that "acquiescence bias" is not a simple phenomenon, and is more a social construction in the interview context than a characteristic of an individual. Peer interviewers would therefore be expected to introduce less social construction bias than other types of interviewers when interviewing people with intellectual disabilities. Whether family and staff tend toward an acquiescence bias in the context of being interviewed by

a person with disability is not known. However, much higher physical and emotional well-being reported for people not able to respond for themselves compared to self-respondents suggests that favorable response bias is a bigger issue for proxy responses than for self responses, particularly in domains of their responsibility.

This study interviewed proxies only when people could not respond for themselves. Therefore, it had no direct way to separate differences in response related to intellectual abilities and differences related to third-party reporting. This is a well-recognized limitation, and Stancliffe (2000) cautions against projecting self and proxy agreement findings to people unable to respond for themselves. The statistical controls used in this study, however, suggest that quality of life reporting is more related to who responded than to the intellectual ability documented in agency records. All levels of intellectual ability available to the study were represented only in self-respondent data, and had a significant relation to self-respondent quality of life only in the domains of social inclusion and self-determination. The primary variability in intellectual ability among those represented by proxies was between profound and severe retardation, and proxy data indicated this difference related only to personal development, self-determination and rights. Early years of the Ask Me! found that agencies could not consistently report greater detail of intellectual ability, and the scores they did report were based on various tests administered at various times in the past. A consistently administered measure of intellectual disability that provided a greater range of differentiation might show different relationships between intellectual ability and quality of life. It may be that the ability to respond for oneself, as determined by peer interviewers, is a better measure of people's intellectual ability than what is recorded in agency files.

Cummins (2002) hypothesized that proxy data reflected their shared life experiences and

general response reliability rather than information specific to the individuals for whom they are reporting. Both family and staff have traditionally focused on physical and emotional well-being, and a great amount of research has been conducted in these domains (Schalock & Verdugo, 2002). Proxies of all types are therefore likely to have a shared perspective on the meaning of physical and emotional well-being, and a vested interest in meeting their expectations. Self-respondents appear to have a different perspective that is less frequently met. Two proxies who share the same relation to a person would be expected to share the most similar perspectives, and this study found greater agreement between two similar staff than between family and staff in six of the eight quality of life domains. Findings from this study support another hypothesis of Cummins that people cannot escape projecting their own prejudices. Most family and staff want what is best for people with disabilities, and believe that they are providing the best possible support. Proxies living with the person confidently reported that people were not hit or hurt in their homes. Proxies outside the home were less confident. In additional, staff proxies were more likely than non-staff proxies to believe that people were getting all the services they needed, and that it was easy to talk to staff about gripes.

The Ask Me! Survey was designed to collect uniform data across a large representative sample of adults living and supported in the community and to analyze differences at one point in time, and changes over time. This purpose limited the amount of detail that could be collected from or about any single individual, and required that any changes in the survey be well justified and introduced incrementally. It also limited the ability of the survey to rigorously test specific hypotheses that would require the use of different instruments and procedures with different sub-samples. Its findings, however, can be generalized to all people with intellectual and other developmental disabilities who receive publicly-funded support in their communities.

This study cannot identify which proxies, or proxy combinations, provide the best information for people who cannot respond for themselves. Ask Me! included only two proxies for a person, and did not select them randomly from a pool of all potential proxies. The study relied on agency personnel to identify two proxies, guided only by the survey's preference to have one family member when possible, and a residential staff or resource coordinator when a family member was not available. Some of the differences between types of proxies could be associated with agency differences in identifying proxies rather than to differences among types of proxies. Other differences could be associated with the relationships between proxies and the individuals, although McVilly, Burton-Smith and Davidson (2000) found little relation between proxy characteristics and subject-proxy agreement. Cummins (2002) hypothesized that "peers provide the most reliable proxy responses due to shared cognition." It might be fruitful to test how friends with intellectual disabilities respond as proxies for people who cannot respond for themselves, in comparison to family and staff.

This study compared disagreement among different types of respondents by calculating a disagreement score that was similar to the percent agreement used by Stancliffe (1995). Standcliffe, however, also used Cohen's kappa and discussed the pros and cons of the two measures of agreement when they disagreed. It is possible that slightly different conclusions might have emerged if kappas had been used in this study, although the calculations of a small set of kappas did not suggest such. The kappa for agreement with self on house mate choice was 0.31 for family proxies, 0.36 for self-respondents, and 0.58 for staff not providing direct care. Kappas of these magnitudes suggest fair self agreement at best. It is also possible that different conclusions might have emerged if different assumptions were made for missing data. This study did not calculate disagreement when one or both of the scores were missing, and this

implicitly assumes that respondents who don't give answers are similar to those who do. However, it might be argued that total lack of knowledge by either or both proxies reflects the most extreme form of disagreement.

This study found that the relation of intellectual ability with five quality of life domains weas the same whether self and proxy responses were analyzed separately or together. Therefore, combining self and proxy data does not appear to affect general conclusions as long as the type of respondent is taken into account. This study showed that adding a self/proxy dichotomous variable to a multiple regression equation provided a reasonable statistical control. Information collected from proxies permits the Ask Me! Survey to fulfill three major reasons for its establishment: 1) to guide systemwide planning, 2) to help agencies enhance their programs, and 3) to provide information to consumers and families to help them make informed choices among agency services. Survey information was never intended to be used in making decisions about individuals, reducing the concern about how much proxies report from their general perspective rather than specific knowledge of the person. Thus people could be interviewed with a promise of confidentiality, reducing potential bias of favorable response to please care givers.

There is no reason to believe that proxies consciously lie, but there is reason to believe that they are influenced by their own perspectives. Multiple proxies may provide no more valid understanding of a person's feelings than a single proxy, but multiple proxies do highlight each other's limitations and different perspectives.

Quality of life reflects external circumstances and internal value judgements. The most valid reporting, therefore, must come from people themselves. The Ask Me! Survey attempts to maximize self-response, in order to better guide the state and community agencies in enhancing support for people with intellectual and other developmental disabilities. Even with peer

interviewers, not everyone can respond for themselves. Proxies must be used if people unable to respond for themselves are to be represented. This study has shown the value of including proxies, even if just to understand proxy biases and encourage providers to involve people as much as possible in decisions that affect their lives.

References

Antaki, C., & Rapley, M. (1996). Questions and answers to psychological assessment schedules: Hidden troubles in 'quality of life' interviews. *Journal of Intellectual Disability Research, 40*, 421-437.

Bonham, G. S., Basehart, S., & Marchand, C. B. (2002). *Ask Me!^{sm} FY2002, The Quality of Life of Marylanders with Developmental Disabilities Receiving DDA Funded Support.* Annapolis, MD: The Arc of Maryland.

Bonham, G. S., Basehart, S., & Marchand, C. B. (2003). *Ask Me!^{sm} FY2003, The Quality of Life of Marylanders with Developmental Disabilities Receiving DDA Funded Support.* Annapolis, MD: The Arc of Maryland.

Bonham, G. S., Basehart, S., & Marchand, C. B. (2004). *Ask Me!^{sm} FY2004, The Quality of Life of Marylanders with Developmental Disabilities Receiving DDA Funded Support.* Annapolis, MD: The Arc of Maryland.

Bonham, G.S., Basehart, S., Schalock, R. L., Marchand, C. B., Kirchner, N., & Rumenap, S. M. (2004). Consumer-Based Quality of Life Assessments: The Maryland Ask Me! Project. *Mental Retardation, 42*, 338-355.

Chong, I., Yu, D., Martin, G., Harapiak, S. & Garinger, J. (2000). Response switching to repeated questions by individuals with developmental disabilities during interviews. *Developmental Disabilities bulletin, 28*, 56-67.

Cummins, R. A. (2002). Proxy responding for subjective well-being: A review. *International Review of Research in Mental Retardation, 25,* 183-207.

Dudley, J. R. (2001). When staff and consumers disagree about consumer satisfaction. *The NADD Bulletin, 4,(6)*, 103-106.

Finlay, W. M. L., & Lyons, E. (2002). Acquiescence in interviews with people who have mental retardation. *Mental Retardation, 40*, 14-29.

Gaudet, L., Pulos, S., Crethar, H., & Burger, S. (2002). Psychosocial concerns of adults with developmental disabilities: perspectives of the self, family member, and provider [Abstract]. *Education and Training in Mental Retardation and Developmental Disabilities, 37*. Retrieved March 15, 2005 from the World Wide Web: http:/www.addcec.org/etmrddv/TOC/etmrddv37n1.html.

Heal, L. W., & Sigelman, C. K. (1996). Methodological issues in quality of life measurement. In R. L. Schalock & G. N. Siperstein (Eds.), *Quality of life: Volume I: Conceptualization and measurement* (pp. 91-104). Washington, DC: American Association on Mental Retardation.

Lunsky, Y., & Benson, B. A. (1997). Reliability of ratings of consumers with mental retardation and their staff on multiple measures of social support. *American Journal on Mental Retardation, 102*, 280-284.

McVilly, K. R., Burton-Smith, R. M., & Davidson, J. A. (2000). Concurrence between subject and proxy ratings of quality of life for people with and without intellectual disabilities. *Journal of Intellectual & Developmental Disability, 25*, 19-39.

Parsons, J. A., Baum, S., & Johnson, T. P. (2000). Inclusion of disabled populations in social surveys: review and recommendations. Chicago, IL: Survey Research Laboratory, University of Illinois at Chicago for The National Center for Health Statistics.

People on the Go. (1996). *Signs of Quality*. Annapolis, MD: The Arc of Maryland.

Perry, J., & Felce, D. (2002). Subjective and objective quality of life assessment: Responsiveness, response bias and resident:proxy concordance. *Mental Retardation, 40*,

445-456.

Perry, J., & Felce, D. (2004). Initial findings of the involvement of people with an intellectual disability in interviewing their peers about quality of life. *Journal of Intellectual & Developmental Disability, 29,* 164-171.

Rapley, M. (2000). The social construction of quality of life: The interpersonal production of well-being revisited. In K. D. Keith & R. L. Schalock (Eds.), *Cross-cultural perspectives on quality of life* (pp. 155-172). Washington, D.C.: The American Association on Mental Retardation.

Rapley, M., Ridgway, J., & Beyer, S. (1997). Staff:staff and staff:client reliability of the Schalock & Keith (1993) Quality of Life Questionnaire. *Journal of Intellectual Disability Research, 42,* 37-42.

Research and Training Center on Community Living. (2004). Response patterns among adult respondents with mental retardation in the National Health Interview Survey, 1997-2002. *DD Data Brief, 6*(2). Institute on Community Integration (UCEDD). Retrieved March 18, 2005 from the World Wide Web: www.rtc.umn.edu/nhis/databrief10/dddb62.pdf.

Schalock, R. L., Bonham, G. S., & Marchand, C. B. (2000). Consumer based quality of life assessment: A path model of perceived satisfaction. *Evaluation and Program Planning, 23*, 77-87.

Schalock, R. L., & Verdugo, M. A. (2002). *Handbook on quality of life for human service practitioners*. Washington, DC: American Association on Mental Retardation.

Sigelman, C. K., Budd, E. C., Winer, J. W., Schoenrock, C. J., & Martin, P. W. (1982). Evaluating alternative techniques of questioning mentally retarded persons. *American Journal of Mental Deficiency, 86*, 511-518.

Sigelman, C. K., Schoenrock, C.J., Spanhel, C.L., Hromas, S.G., Winer, J.L., Budd, E.C., & Martin, P.W. (1980). Surveying mentally retarded persons: Responsiveness and response validity in three samples. *American Journal of Mental Deficiency, 84*, 479-486.

Stancliffe, R. J. (1995). Assessing opportunities for choice-making: A comparison of self- and staff reports. *Mental Retardation, 99*, 418-429.

Stancliffe, R. J. (1999). Proxy respondents and the reliability of the Quality of Life Questionnaire Empowerment factor. *Journal of Intellectual Disability Research, 43,* 185-193.

Stancliffe, R. J. (2000). Proxy respondents and quality of life. *Evaluation and Program Planning, 23*, 89-93.

Umb-Carlsson, O. (2005). Living Conditions of People with Intellectual Disabilities: A Study of Health, Housing, Work, Leisure and Social Relations in a Swedish County Population. Uppsala, Sweden: Uppsala Universitet, *Digital comprehensive summaries of Upsala Dissertations from the Faculty of Medicine 89.*

Whitney-Thomas, J. (1996). Participatory action research as an approach to enhancing quality of life for individuals with disabilities. In R. L. Schalock & G. N. Siperstein (Eds.), *Quality of life: Volume II: Application to persons with disabilities* (pp. 181-197). Washington, DC: American Association on Mental Retardation.

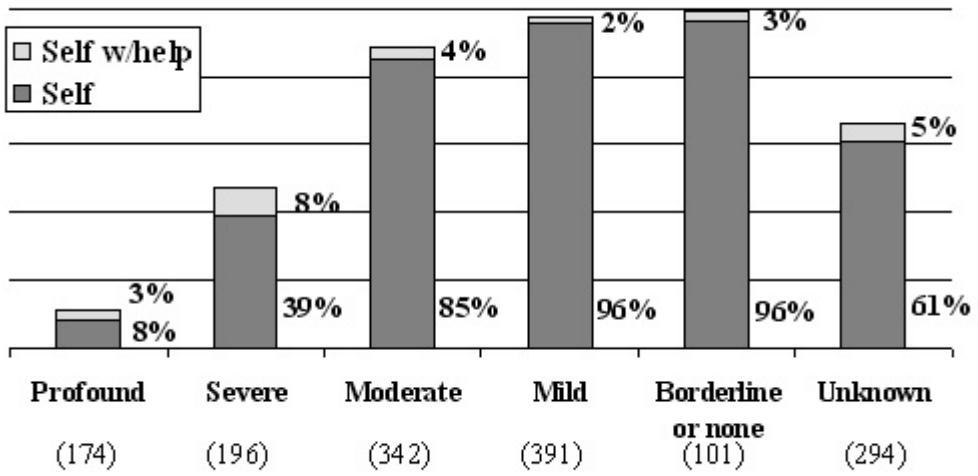

Figure 1. Percent Self-Response by Level of Retardation

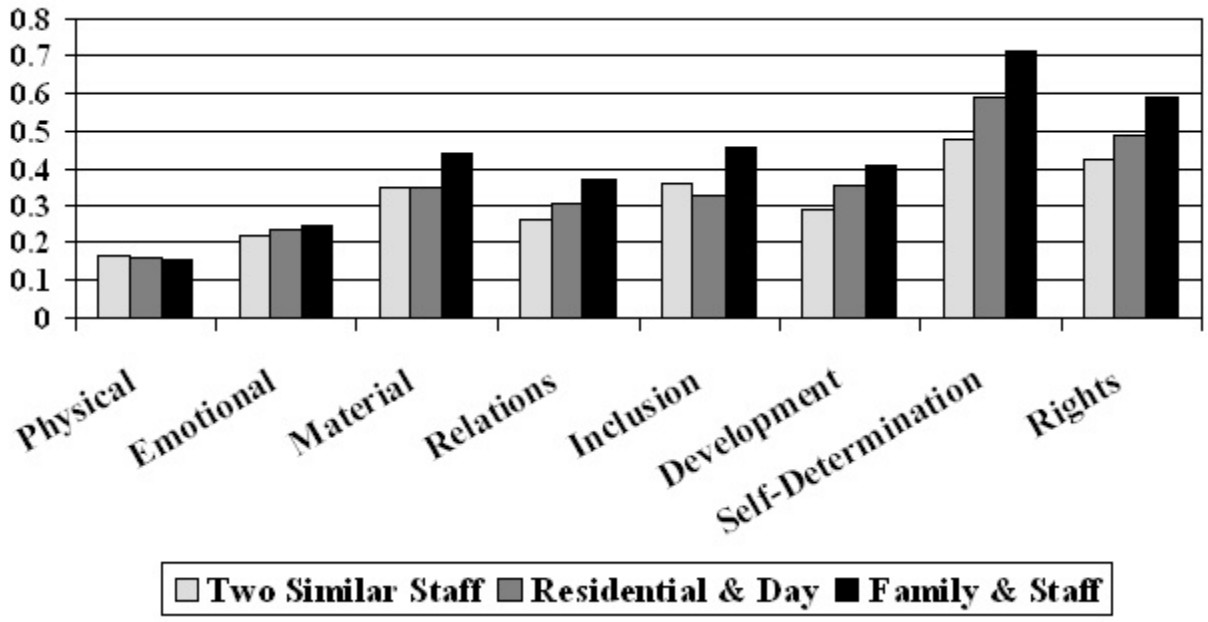

Figure 2. Disagreement Scores on Quality of Life by Proxy Combination

	Disagreement Score			***Answered Both Questions***		
	Earn good money	*Happy with life*	*Pick house mate*	*Earn good money*	*Happy with life*	*Pick house mate*
All Self-Respondents	**.25**	**.26**	**.52**	**89%**	**91%**	**92%**
Profound retardation	.40	.27	.67	71%	79%	86%
Severe retardation	.25	.33	.59	77%	79%	86%
Moderate retardation	.27	.28	.53	87%	89%	93%
Mild retardation	.28	.23	.47	93%	95%	93%
Borderline or no retardation	.20	.22	.54	88%	95%	92%
All Proxy Respondents	**.16**	**.18**	**.48**	**85%**	**94%**	**84%**
Family, friend	.17	.10	.46	84%	90%	90%
Day staff	.18	.19	.43	84%	94%	79%
Residential staff	.10	.18	.59	89%	96%	91%

Table 1. Disagreement with Self on Responses to Duplicated Questions, by Type of Respondent

		All		*Day*	*Residential*
Domain Scale	*Self*	*Proxies*	*Family*	*Staff*	*Staff*
(Number of responses)	*(1031)*	*(751)*	*(123)*	*(433)*	*(195)*
Average	0.67	0.55	0.53	0.55	0.48
Rights	0.62	0.61	0.43	0.66	0.59
Self-determination	0.63	0.56	0.38	0.63	0.55
Personal development	0.68	0.57	0.76	0.56	0.53
Social inclusion	0.70	0.53	0.35	0.62	0.42
Interpersonal relations	0.73	0.45	0.63	0.34	0.43
Material well-being	0.66	0.63	0.75	0.53	0.60
Emotional well-being	0.68	0.62	0.70	0.61	0.46
Physical well-being	0.64	0.44	0.25	0.48	0.28

Table 2. Cronbach's Alpha for Scale Reliability, by Type of Respondent

Respondent	Physical Well-being	Emotional Well-being	Material Well-being	Interpersonal Relations	Social Inclusion	Personal Development	Self-Determination	Rights
All Self-Respondents	**8.46**	**8.37**	**7.50**	**7.79**	**7.64**	**7.75**	**7.62**	**6.93**
Profound retardation (14)	9.17	8.99	7.92	9.37	9.01	8.41	7.35	7.39
Severe retardation (77)	8.14	8.02	7.34	7.05	7.42	7.29	6.90	6.48
Moderate retardation (290)	8.17	8.32	7.37	7.72	7.83	7.74	7.48	6.87
Mild retardation (374)	8.59	8.43	7.55	7.84	7.61	7.87	7.82	7.02
Borderline or no retardation (97)	8.37	8.09	7.22	7.62	7.10	7.18	7.78	7.18
Not reported (179)	8.79	8.57	7.80	8.10	7.67	7.97	7.66	6.81
All Proxy Respondents	**9.33*** **	**9.02*** **	**7.78** **	**8.14*** **	**7.52**	**7.23*** **	**6.19*** **	**4.99*** **
Family/Friend (123)	9.53***	9.52***	8.16**	8.71***	7.36	7.67	6.19***	3.70***
Day Staff (433)	9.18***	8.74***	7.38	7.97	7.50	7.04***	6.06***	5.30***
Residential Staff (195)	9.54***	9.33***	8.32***	8.15**	7.66	7.40**	6.48***	5.13***

Different from all self-respondents: * $p < .05$, ** $p < .01$, *** $p < .001$

Table 3. Average Quality of Life Score by Type of Respondent and Intellectual Ability

	Intercept (Profound)	Intellectual ability	Proxy	R^2	F(d.f.)
Physical well-being					
Self-respondents	8.18	0.11		.00	2.06(1,798)
Proxy respondents	9.37	-0.11		.01	3.41(1,299)
Combined	8.35	0.05	0.94***	.05	30.24(2,1098)***
Emotional well-being					
Self-respondents	8.38	-0.00		.00	0.00(1,849)
Proxy respondents	9.03	-0.02		.00	0.10(1,300)
Combined	8.35	0.01	0.65***	.03	16.12(2,1150)***
Material well-being					
Self-respondents	7.53	-0.01		.00	0.01(1,779)
Proxy respondents	7.81	0.01		.00	0.12(1,286)
Combined	7.54	-0.01	0.28	.01	2.41(2,1066)
Interpersonal relations					
Self-respondents	7.69	0.04		.00	0.22(1,821)
Proxy respondents	8.05	0.05		.00	1.05(1,299)
Combined	7.76	0.01	0.18	.00	2.45(2,1121)
Social inclusion					
Self-respondents	8.15	-0.20*		.01	5.31(1,831)*
Proxy respondents	7.52	-0.06		.00	0.41(1,298)
Combined	8.06	-0.17*	-0.46*	.01	3.66(2,1130)*

Personal development					
Self-respondents	7.87	-0.04		.00	0.23(1,806)
Proxy respondents	6.97	0.26**		.03	7.72(1,299)**
Combined	7.69	0.02	-0.57**	.02	10.28(2,1106)***
Self-determination					
Self-respondents	6.96	0.26**		.01	8.48(1,800)**
Proxy respondents	5.77	0.44***		.04	12.68(1,297)***
Combined	6.81	0.31***	-0.99***	.13	70.03(2,1098)***
Rights					
Self-respondents	6.55	0.15		.00	2.37(1,775)
Proxy respondents	4.51	0.55***		.06	16.99(1,294)***
Combined	6.11	0.31***	-1.50***	.16	100.55(2,1070)***

* $p < .05$ ** $p < .01$ *** $p < .001$

Table 4. Quality of Life Scores Regressed on Intellectual Ability, by Respondent