

# Beyond the Checklist

## Assessing Understanding for HIV Vaccine Trial Participation in South Africa

Graham Lindegger, PhD,\* Cecilia Milford, MSocSc,\* Catherine Slack, MA,\*  
Michael Quayle, MA,\* Xolani Xaba, BSocSc (Honours),\*  
and Eftyhia Vardas, BSc (Honours), MB BCh, DTM&H, DPH, FCPATH, MMed†

**Objectives:** Informed consent and understanding are essential ethical requirements for clinical trial participation. Traditional binary measures of understanding may be limited and not be the best measures of level of understanding. This study designed and compared 4 measures of understanding for potential participants being prepared for enrollment in South African HIV vaccine trials, using detailed operational scoring criteria.

**Methods:** Assessment of understanding of 7 key trial components was compared via self-report, checklist, vignettes, and narrative measures. Fifty-nine participants, including members of vaccine preparedness groups and 1 HIV vaccine trial, took part.

**Results:** There were significant differences across the measures for understanding of 5 components and for overall understanding. Highest scores were obtained on self-report and checklist measures, and lowest scores were obtained for vignettes and narrative descriptions.

**Conclusions:** The findings suggest that levels of measured understanding are dependent on the tools used. Forced-choice measures like checklists tend to yield higher scores than open-ended measures like narratives or vignettes. Consideration should be given to complementing checklists and self-reports with open-ended measures, particularly for critical trial concepts, where the consequences of misunderstanding are potentially severe.

**Key Words:** understanding, consent, HIV vaccine trials, ethics

(*J Acquir Immune Defic Syndr* 2006;43:560–566)

Informed consent (IC) is a cornerstone of clinical trials. It has been argued that “genuine consent”<sup>1</sup> requires more than satisfying legal formalities (eg, signing consent forms). Although investigators have long been required to disclose

relevant information, there has been more recent recognition of the need to demonstrate participants’ understanding<sup>2</sup> for adequate IC. Although many international ethical guidelines make little explicit reference to the need to test understanding,<sup>3</sup> this has been upheld as a core component of consent.<sup>4,5</sup> The HIV/AIDS Vaccines Ethics Group (HAVEG), part of the South African AIDS Vaccine Initiative (SAAVI), has been concerned with ensuring sound consent procedures (including assessment of understanding) for participants in HIV vaccine trials with particular reference to cultural sensitivity.

Assessment of understanding is potentially complicated. For example, some methods may test short-term recall of disclosed technical information. Although some degree of retention is probably a prerequisite for understanding, it cannot be equated with understanding.<sup>6</sup> In many studies investigators use forced-choice true-false (eg, right/wrong, agree/disagree) checklists to measure understanding. These checklists may tend to invite rote repetition of technical, product, and methodologic information rather than assess understanding of the implications of participation for participants’ personal lives,<sup>7,8</sup> which might be as important, if not more important, to enable potential participants to make thoughtful decisions about enrollment. It is possible that despite adequate checklist scores, participants have not understood trial information in terms of these personal implications. In addition, repeated administration of the same or similar checklists or questionnaires may only measure a learned response that does not necessarily reflect understanding.<sup>9</sup> Furthermore, some assessment methods (eg, those based on binary [right/wrong] approaches) may run the risk of cultural insensitivity. In some cultures, it is possible that narrative measures may be more appropriate. Such complications may potentially undermine assessment of understanding, reducing IC to a legal formality. Although there have been calls for the development of innovative materials and processes to improve understanding and its assessment,<sup>3</sup> more research is needed in this area, especially given the complexities of HIV vaccine trials conducted in developing countries.<sup>10</sup>

Assessment of understanding has been recommended as a routine part of HIV vaccine trials, particularly in less developed countries.<sup>11</sup> A number of studies have assessed understanding of important components of these trials using various methods and yielding mixed findings. Most studies seem to have used forced-choice checklists, with many finding that participants performed well,<sup>12–14</sup> although some show only

Received for publication May 18, 2006; accepted September 19, 2006.

From the \*HIV/AIDS Vaccines Ethics Group, School of Psychology, University of KwaZulu-Natal, Pietermaritzburg, Private Bag X01, Scottsville 3209, South Africa; and †HIV/AIDS Vaccine Division, Perinatal HIV Research Unit, Faculty of Health Sciences, University of the Witwatersrand, PO Box 114, Diepkloof, Johannesburg 1864, South Africa.

Funding received from the South African AIDS Vaccine Initiative.

Reprints: Graham Lindegger, PhD, HIV/AIDS Vaccines Ethics Group, School of Psychology, University of KwaZulu-Natal, Private Bag X01, Scottsville 3209, South Africa (e-mail: lindegger@ukzn.ac.za).

Copyright © 2006 by Lippincott Williams & Wilkins

moderate levels of understanding.<sup>15</sup> Other studies have used coded responses to structured interviews and found high levels of understanding.<sup>16</sup> Some studies have used a combination of methods of assessment, finding more moderate levels of understanding among participants.<sup>17</sup> Some commentators have been concerned that assessment measures should be culturally sensitive.<sup>18,19</sup>

## OBJECTIVES

This exploratory study aimed to develop and compare 4 alternative methods to assess understanding of components central to phase 1 HIV vaccine trial participation so as to make recommendations to sites in South Africa of ways of supplementing current methods to assess understanding.

## METHODS

### Sample

There were a total of 59 participants in this research. Most ( $n = 53$ ) were potential volunteers for phase 1 HIV vaccine trials participating in prescreening protocols (Vardas et al, manuscript in preparation). There were also 6 participants who were already taking part in an HIV vaccine trial. All participants were opportunistically sampled and referred to the study by trial staff. Recruitment success was between 65% and 75%. Because the main aim of the study was to compare 4 measures of understanding using a within-subjects design, we could recruit participants from any stage of the trial process.

The final sample included 26 male and 33 female respondents. Of these, 92% ( $n = 54$ ) had completed high school (matric). The age range was 16 to 55 years. Only 20% ( $n = 12$ ) of participants were employed. With regard to exposure to vaccine knowledge, 44% ( $n = 26$ ) had attended 1 to 2 information sessions, whereas 56% ( $n = 33$ ) had attended 3 or more sessions at a trial site. Seventy-eight percent ( $n = 46$ ) had attended at least 1 individual education session with a trial counselor. In addition, 95% ( $n = 56$ ) had received some written material on the HIV vaccine trial.

### Design

The lack of prior research comparing open-ended and closed-ended techniques made it impossible to estimate effect size with integrity, which meant that a priori power and sample size calculations were ruled out. A within-subjects design was used, because the aim of the study was to compare different methods of assessing understanding on the same volunteers. Such designs are quite efficient. For example, it has been calculated that when the correlation between conditions is 0.8, a within-subjects design can detect a medium effect size with approximately one tenth of the sample size required by an equivalent between-subjects design.<sup>21</sup> Maximizing power, such designs ameliorate the problems associated with non-random sampling and serve to control potential confounding variables, such as level of education, because each participant acts as his or her own control.<sup>22</sup>

### Development of Assessment Methods

We began the study by determining those components essential for participants to understand so as to participate in

phase 1 HIV vaccine trials. After reviews of literature and guidelines, we extensively consulted with trial staff (eg, investigators, educators) and community representatives, using the guiding question, "What do you consider essential for participants to understand to enroll in HIV vaccine trials?" From a longer list, 7 components were selected in consultation with trial staff and community representatives to decrease the contact time required with our study participants. The components or concepts included trial aims; eligibility to participate; risks of testing antibody-positive if vaccinated ("false positivity"); risk of falsely believing the vaccine would protect one from infection, and thus increasing one's risk behavior ("false sense of security"); methodologic considerations such as randomization, placebo, and blinding; compensation for research-related injury; and right to withdraw.

Four different methods were developed to assess understanding of each of these components and included the following:

1. Self-report: participants were asked to estimate their level of personal understanding of each component as "little/no understanding" or "good enough understanding." For example, for "right to withdraw," participants were asked, "How would you rate your understanding of whether you have a right to no longer take part in the study if you don't want to?"
2. Forced-choice checklist: participants were required to complete a self-administered questionnaire that asked whether statements about trial participation were true or false. There were 3 statements for each component. For example, for the right to withdraw, one of the statements read, "If you leave the study, then you may lose some of the benefits related to trial participation."
3. Vignettes: participants were asked to respond to scenarios embodying specific aspects of each component that used a fictitious character. The vignettes were read to the participants by the researcher. For example, for right to withdraw, the following vignette was presented: "After 6 months of being enrolled in the HIV vaccine trial, Mrs. Dlamini (or Mrs. Jones) is tired and wants to get out of the trial. But she is afraid what will happen to her and how the trial staff will respond if she decides to leave. What do you think she should do, and what would happen if she leaves?" When necessary, standardized prompts were used to elicit information about the component at hand.
4. Narratives: participants were asked to describe participation in an HIV vaccine trial as if they were telling a friend about such participation. They responded to a standard question that was read to them by the researcher. Again, when necessary, standardized prompts were used; for example, "What kind of choice do people have to be involved in or to leave the trial?" and "What might happen if one decided to withdraw/leave?" Responses to vignettes and narrative descriptions were tape recorded with participants' consent.

### Developing and Using the Scoring System

Detailed operational criteria for scoring responses to vignettes and narratives were developed in collaboration with site staff. Development of these criteria was guided by the question, "What would satisfy us that a participant understood

enough of this component to be enrolled in the trial?" This would represent good enough understanding for trial participation. Site staff varied in their criteria for good understanding; however, extensive discussions between collaborators led to consensus about scoring criteria for each component/concept listed previously. These criteria can therefore be seen as a "gold standard" for what participants should understand for each component identified so as to be qualified to participate in trials at the sites where this research took place.

For the self-reported measure, participants estimated their own level of understanding. For the checklist, participants had to score correctly on certain items to demonstrate good enough understanding. Based on the operational criteria described previously, responses to the open-ended assessment methods were scored on a 2-point scale of understanding (no/little understanding and good enough understanding) in accordance with the requirement that HIV vaccine trial participants should have good enough understanding to enroll in a trial. For each measure, the combined scores for components were taken to represent an estimate for participants' overall level of understanding.

Despite the establishment of clear operational criteria for the scoring of open-ended responses, there was the possibility of some measure of subjectivity in scoring. As a scoring reliability check, it was decided to use 2 independent raters and to calculate the level of interrater agreement between the scorers, using the Pearson correlation coefficient (Table 1). For 12 of 14 scores across the 2 open-ended measures, the interrater reliability exceeded 0.715 but was lower for the narrative measures of trial aims ( $r = 0.675$ ) and right to withdraw ( $r = 0.567$ ). Closer examination of those measures shows that the larger (although not unacceptable) discrepancies between scorers may have been attributable to missing data and interpretation of translated transcripts. Where larger discrepancies occurred, however, scorers carefully rescored the responses until consensus was reached on all scores.

## Ethical Approval

Ethical approval was obtained from 3 local ethics committees in South Africa that had jurisdiction over the HIV vaccine trial sites where this research was conducted.

## Procedures

A pilot study was conducted with community advisory board members before the research to refine aspects of the

study, including the clarity, simplicity, and relevance of the assessment measures.

After the pilot study, potential participants were approached to participate in this research in a least disruptive manner approved by trial staff. The confidentiality of potential HIV vaccine trial participants was maintained. Once potential participants indicated their willingness to participate and had given consent, they were asked to complete all 4 assessment methods in a single visit so as to control for the possible effect of gaining or losing information over time.

One weakness of a repeated-measures design is the sequence effect, where order of presentation can have an impact on response to later methods. Therefore, the order of presentation of different methods of assessment was varied. The self-report was presented first to control for the realization, after completing the other methods, that participants' understanding was better or worse than they had thought. This was followed by presentation of the vignettes, narrative, and checklist, in varied order, to control for the effects of the order of presentation of the assessment methods. Because of the small sample size, however, the exact impact of this variation in presentation order could not be tested.

Responses to the narrative and vignette were tape recorded when consent was granted ( $n = 57$ ), or detailed notes were taken when the participant did not agree to being tape recorded ( $n = 2$ ). Responses to the vignette and narrative were transcribed, and all responses for each of the 7 components were scored, making use of the operational criteria described previously.

## Analysis

A score was calculated for each participant's understanding of each component on each assessment method (7 components across 4 methods). Overall understanding was calculated as a sum of these separate measures and expressed as a percentage of the possible total score.

Comparisons were made across the different measures of assessment for each component and for total scores of understanding, using SPSS 11.5 (SPSS Inc., Chicago, IL). The nature of the data and the small sample size necessitated the use of nonparametric techniques. The Friedman nonparametric equivalent of the repeated-measures analysis of variance (ANOVA) was used to compare overall understanding scores, and the Cochran Q was used to compare individual components across different measures. The Cohen Kappa<sup>23</sup> was used to explore further the level of agreement between the different components of each measure, yielding estimates of the degree to which the different measures result in equivalent estimates of participants' understanding of each component.

## Findings

The results reveal variability in scores for overall understanding and understanding of each component across the 4 assessment methods (Table 2). There was a significant difference between scores of overall understanding across different methods of assessment (using the Friedman rank test:  $\chi^2_F = 48.424$ ;  $df = 4$ ;  $P < 0.0005$ ). The highest proportions of good enough scores for overall understanding were on the self-reported measure of understanding (86%), followed by

**TABLE 1.** Interrater Reliability Coefficients for Scoring Open-Ended Responses

Component	Vignette	Narrative
Trial aims	0.820	0.675
Eligibility to participate	0.932	1.000
Methodologic considerations	0.878	0.918
Risk (false sense of security)	0.815	0.825
Risk (false positivity)	1.000	0.890
Compensation for research-related injury	0.765	0.715
Right to withdraw	0.730	0.567

All values were significant at a 95% confidence interval.

**TABLE 2.** Frequency of Good Enough Understanding for Each Subscale

Component	Percentage of Participants Rated as Good Enough per Section			
	Self-Report	Checklist	Vignette	Narrative
Trial aims	95%	93%	53%	60%
Eligibility to participate	96%	86%	84%	95%
Methodologic considerations	72%	83%	67%	67%
Risk (false sense of security)	88%	67%	72%	76%
Risk (false positivity)	87%	75%	55%	60%
Compensation for research-related injury	70%	85%	44%	43%
Right to withdraw	96%	95%	93%	98%
Overall/total understanding (%)	86%	82%	67%	69%

the checklist (82%), narrative (69%), and, finally, the vignette (67%).

The Cochran Q reveals that the 4 measures resulted in significantly different proportions of good enough answers for 5 of the 7 components\* (all except methodologic considerations and right to withdraw). In general, a greater proportion of participants emerged with ratings of good enough understanding on the self-report and checklist measures. The self-report assessment method yielded the highest proportion of good enough responses for 4 of the 7 components, and the checklist yielded the highest proportion of good enough responses for 2 of 7 components. The lowest proportion of good enough ratings was observed on the vignette measure for 3 components and on the narrative measure for 1 of 7 components. Methodologic considerations were tied lowest on the narrative and vignette measures; however, the narrative measure yielded the highest proportion of good enough responses for right to withdraw.

Given their longer and more extensive exposure to trial-related information, the 6 trial participants' scores, when compared with those of the discussion group participants, tended to be better on all 4 measures. Mean scores for each method of assessment were slightly higher when trial participants' scores were included than when they were excluded. Findings from the Cochran Q yielded the same results, however, with the same levels of significance when trial participants were included and excluded in analysis.

Breaking the data down by component, some components elicited consistently lower proportions of good enough responses across measures of understanding than others (eg, compensation for research-related injury, risks, methodologic considerations, trial aims). Others, for example, a participant's right to withdraw and eligibility to participate (which includes HIV-negative status) elicited consistently higher proportions of good enough responses across different measures of understanding.

To test the a priori expectation that differences between measures would most likely be between the closed-ended and open-ended measures, the Cohen Kappa was first calculated

within the closed-ended (ie, self-report, checklist) and open-ended (ie, narrative, vignette) measures and then between the closed-ended and open-ended measures. The results, as displayed in Figure 1, demonstrate a significant (although not necessarily strong) agreement within each assessment method (closed-ended or open-ended) for each component. There is no significant agreement between assessment methods for 5 of the 7 components. Nevertheless, there is significant agreement between methods for methodologic considerations (this is confirmed by earlier results from the Cochran Q) and the risk of a false-positive test result.

To explore the speculation that differences between open-ended and closed-ended measures were disproportionately driven by the self-report measure, comparisons were repeated for only the checklist, vignette, and narrative using the Cochran Q. Low- and high-scoring components were the same as those reported earlier in this report. Therefore, the differences between measures cannot be reduced simply to the impact of self-report measures.

Table 3 illustrates how some participants' (in the early stages of their education process) scores of understanding differed across methods of assessment for the same components. The narrative responses in this example all demonstrate poor understanding of the component, but they also provide richer insight into a participant's understanding, enabling the interviewer to probe appropriately.

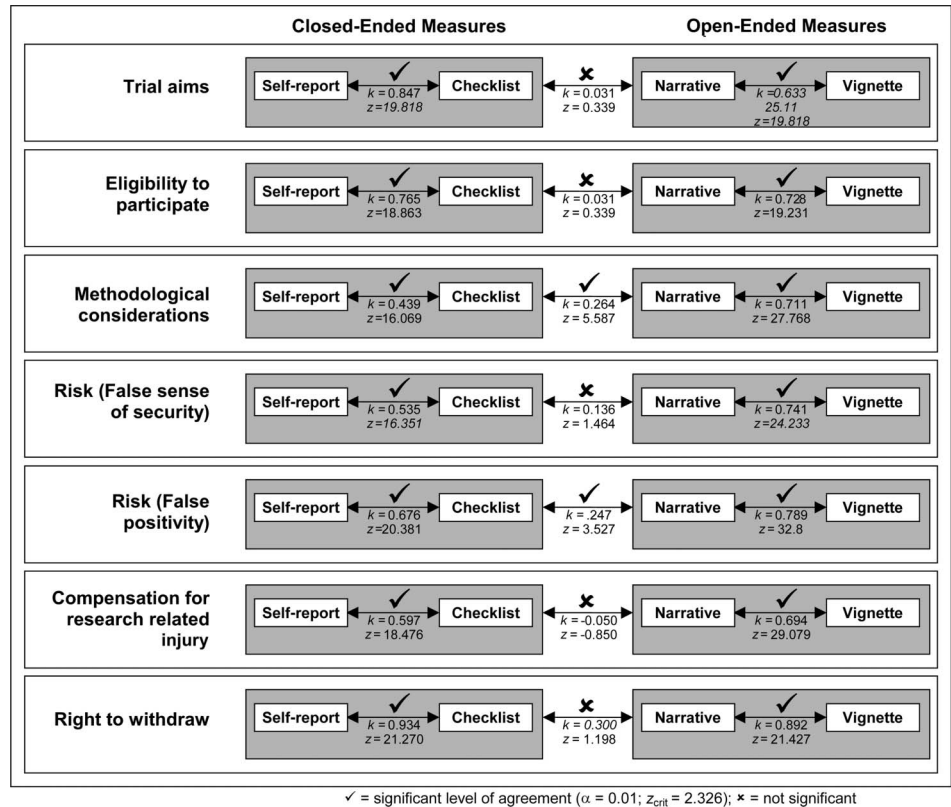
On average, the narrative measure took longest to complete (15.1 minutes, with a range of 3–41 minutes). This was followed by the vignette (average of 10.9 minutes, with a range of 4–23 minutes), the checklist (average of 5.3 minutes, with a range of 1–14 minutes), and, finally, the self-report (average of 2.5 minutes, with a range of 1–5 minutes).

## DISCUSSION

Overall, the different methods of assessment yielded different estimates of participants' levels of understanding. Scores derived from self-report measures consistently yielded the highest estimates of understanding, followed by forced-choice questionnaires and then narrative and vignette descriptions.

Considering that the criteria established for the scoring of open-ended responses (vignette and narrative) were agreed on by trial staff as the essential minimum level of understanding for entry into the trial, these can be regarded as the gold standard of understanding. Therefore, by this standard,

\*For "trial aims,"  $Q = 45.192$ ,  $df = 3$ , and  $P < 0.0005$ ; for "eligibility to participate,"  $Q = 7.929$ ,  $df = 3$ , and  $P = 0.048$ ; for "risk (false sense of security),"  $Q = 8.509$ ,  $df = 3$  and  $P = 0.037$ ; for "risk (false positivity),"  $Q = 23.76$ ,  $df = 3$ , and  $P < 0.0005$ ; and for "compensation for research-related injury,"  $Q = 32.353$ ,  $df = 3$ , and  $P < 0.0005$ .



**FIGURE 1.** Agreement between measures.

the findings suggest that self-report and forced-choice checklists may overestimate levels of understanding by participants of essential components of HIV vaccine trials.

There may be various explanations for the findings. Self-reported understanding is likely to be strongly affected by social desirability (ie, the desire to show vaccine educators the efficacy of their education) or the desire to enter trials and may not be a reliable measure of actual understanding.<sup>18</sup> In many research settings, it is likely that comprehension is assessed merely by asking participants, “Do you understand what this means?”

It is possible that responses to forced-choice questions reflect an acontextual rote memorizing of technical information about trials. The findings suggest that participants

master this task fairly well; however, when requested to respond to more contextually embedded tasks such as vignettes, their scores of understanding are substantially lower. Qualitative observations during the assessment process lent some support to the latter speculation. For example, a few participants were observed to be guessing answers (albeit correctly) on the checklist; however, when confronted with the same component on vignettes or narratives, these participants spontaneously declared that they understood a component less well than they thought.

The relatively high scores on checklists are consistent with previous studies showing that these measures can result in high scores.<sup>12–14</sup> The findings are also consistent with studies showing that levels of measured understanding based

**TABLE 3.** Example of Different Scores for Components Between Closed-Ended (Self-Report and Checklist) and Open-Ended (Narrative) Measures

Concept	Self-Report	Checklist	Narrative (Poor Understanding)
Placebo	Poor	Good enough	“I cannot differentiate between a placebo and the vaccine. But I think it (is) clear that both protect.”
Compensation for research-related injury	Good enough	Poor	“Like when they’re playing the rugby, you must know you’re gonna get hurt ... you choose to come, nobody forced you, you’re asked to play so you must (expect that there’s going to be some harm and you mustn’t expect them to help you out).”
False positivity	Good enough	Good enough	I: “When she gets her test back, it shows that she has tested HIV-positive. What could this mean?” P: “It means Mrs Dlamini was not too much into the vaccine trials, it means that she was not faithful ... It means it (was) not working.”
Eligibility to participate	Good enough	Good enough	“Everybody can participate except the people over 70 years of age.” Later: “One month up to 59.”

on mixed methods of assessment yielded scores that were lower than those using only forced-choice questionnaires.<sup>17</sup> The consistently low scores for components such as methodology are also consistent with previous literature.<sup>24</sup>

The analysis of agreement (see Fig. 1) also shows that the closed-ended and open-ended measures have fairly low levels of consensus between them. This suggests that the measures are not simply over- or underrepresenting different levels of a monolithic construct called “understanding” but may be accessing qualitatively different aspects of a multidimensional construct. For example, closed-ended measures may be accessing aspects such as recognition and recall, whereas open-ended measures may access more complex features of comprehension. This does not automatically imply a value judgment between the different methods of assessment but indicates that measures should be selected carefully according to their strengths and function.

The results reported here should caution us that closed-ended measures may overestimate levels of understanding and that open-ended measures may yield more accurate measures of desired levels of understanding. The use of open-ended measures involves some additional costs, however; namely, they are more resource-intensive, require skilled interviewing and analysis, and are more difficult to justify as “objective” measures. Furthermore, in phase 3 HIV vaccine studies, where enrollment numbers are high (a few thousand), the labor-intensive processes of open-ended methods may be difficult to implement. Additionally, the more conservative estimates of understanding yielded by open-ended measures might require more participants to receive additional education.

Why, then, should trial sites concern themselves with open-ended measures of understanding? We argue that the choice of method of assessment should be weighed against the need for adequate understanding determined, in part, by potential risks for participants and potential consequences of poor understanding. More specifically, critical issues (eg, falsely believing that the product confers protection against the condition under study) could be assessed using resource-intensive open-ended measures, whereas less imperative aspects (eg, number of trial visits) could be assessed using traditional closed-ended approaches.

### LIMITATIONS

The sample for this study was not randomly drawn. Because of the scarcity of suitable participants and difficulty in obtaining permission to sample participants at all stages of participation, opportunistic sampling was used. Levels of uptake from potential participants were reasonably good, however. Although the sample size was fairly small, the within-subject design meant that this was not a severe limitation, as argued earlier. With regard to the self-report measure, in retrospect, it was realized that phrasing of the self-report questions was a little clumsy.

### CONCLUSIONS

IC is fundamental to the ethical conduct of research.<sup>3</sup> As part of this process, it is necessary to develop valid and reliable measures of understanding that are also culturally sensitive

and appropriate and free of some of the limitations of traditional measures.

This study suggests that levels of measured understanding are dependent on methods of assessment used and that closed-ended measures like self-reports and checklists may overestimate understanding in comparison with more open-ended measures. These findings suggest that standard checklists, with strengths of objectivity and ease of administration, should be complemented with open-ended measures, especially for critical concepts that, if misunderstood, could have serious consequences for participants. Although the latter are more resource-intensive, they may provide more accurate measures of what participants actually understand. Furthermore, there is preliminary evidence that such extended discussion and interactions with trial staff are the best ways of improving understanding,<sup>9</sup> therefore, open-ended techniques to probe and assess understanding may, in fact, have the effect of enhancing comprehension, automatically adding value to trial preparations.

The aim of this research was to test and compare measures of understanding rather than to assess individual participants' understanding per se. Future research should include participants at different stages of HIV vaccine trials so as to broaden results and make them more generalizable to HIV vaccine trials. It may also be useful to conduct similar research in other non-HIV research. This study contributes initial information to a complex ethical issue, and it is hoped that the use of alternate measures of understanding may be further explored to enhance the quality of IC in clinical trials.

### ACKNOWLEDGMENTS

*The authors thank their sponsor, the SAAVI, for funding this study. They also thank Dr. Andrew Robinson and Dr. Pat Fast for their useful comments on the manuscript. They are grateful to International AIDS Vaccine Initiative and HIV vaccine trial site staff for assisting them with the research process and to the participants in this research and the trial staff and community representatives who helped to develop the essential components and criteria for understanding. Finally, they thank Dr. Nancy Kass for helpful discussions on this issue.*

### REFERENCES

1. Nuffield Council on Bioethics. *The Ethics of Research Related to Healthcare in Developing Countries*. London: Nuffield Council on Bioethics; 2002.
2. Beauchamp TL, Childress JF. *Principles of Biomedical Ethics*. New York: Oxford University Press; 2001.
3. Bhutta ZA. Beyond informed consent. *Bull World Health Organ*. 2004;82:771–778.
4. Agre P, Rapkin B. Improving informed consent: a comparison of four consent tools. *IRB*. 2003;25:1–7.
5. Levine RJ. *Ethics and Regulation of Clinical Research*. 2nd ed. New Haven, CT: Urban & Schwarzenberg; 1986.
6. Meisel A, Roth L. Toward an informed discussion of informed consent: a review and critique of the empirical studies. *Ariz Law Rev*. 1983;25:266–346.
7. Benatar SR. Reflections and recommendations on research ethics in developing countries. *Soc Sci Med*. 2002;54:1131–1141.
8. Rosnow R, Rosenthal R. *People Studying People: Artifacts and Ethics in Behavioural Research*. New York: WH Freeman; 1997.
9. Flory J, Emanuel E. Interventions to improve research participants' understanding in informed consent for research: a systematic review. *JAMA*. 2004;292:1593–1601.

10. UNAIDS. *Ethical Considerations in HIV Preventive Vaccine Research: UNAIDS Guidance Document*. Geneva: UNAIDS; 2000.
11. Fitzgerald D, Marotte C, Verdier I, et al. Comprehension during informed consent in a less-developed country. *Lancet*. 2002;360:1301–1302.
12. Fureman I, Meyers K, McLellan AT, et al. Evaluation of a video-supplement to informed consent: injection drug users and preventive HIV vaccine trials. *AIDS Educ Prev*. 1997;9:330–341.
13. Harrison K, Vlahov K, Jones K, et al. Medical eligibility, comprehension of the consent process and retention of injection drug users recruited for an HIV vaccine trial. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1995; 10:386–390.
14. MacQueen KM, Vanichseni S, Kitayaporn D, et al. Willingness of injection drug users to participate in an HIV vaccine efficacy trial in Bangkok, Thailand. *J Acquir Immune Defic Syndr*. 1999;21:243–251.
15. Querioz da Fonseca O, Lie RK. Comprehension of the informed consent form and general knowledge of vaccines among potential participants for an HIV vaccine trial in Brazil. *Int J Pharm Med*. 1999;12: 2–8.
16. McGrath JW, George K, Svilar G, et al. Knowledge about vaccine trials and willingness to participate in an HIV/AIDS vaccine study in the Ugandan military. *J Acquir Immune Defic Syndr*. 2001;27:381–388.
17. Querioz da Fonseca O, Lie RK. Informed consent to AIDS—vaccine trials in Brazil. *AIDS Public Policy J*. 1995;10:22–26.
18. Lindegger G, Richter L. HIV vaccine trials: critical issues in informed consent. *South African Journal of Science*. 2000;96:313–317.
19. Richter LM, Lindegger GC, Abdool Karim Q, et al. *Guidelines for the Development of Culturally Sensitive Approaches to Obtaining Informed Consent for Participation in HIV Vaccine Related Trials. Discussion Document*. Geneva: UNAIDS; 1999.
20. Deleted in proof.
21. Venter A, Maxwell SE. Maximising power in randomised designs when sample size is small. In: Hoyle R, ed. *Statistical Strategies for Small Sample Research*. Thousand Oaks, CA: Sage; 1999:31–58.
22. West SG, Biesanz JC, Kwok OM. Within-subject and longitudinal experiments: design and analysis issues. In: Sansone C, Morf CC, Panter AT, eds. *The Sage Handbook of Methods in Social Psychology*. Thousand Oaks, CA: Sage; 2004.
23. Siegel S, Castellan NJ. *Nonparametric Statistics for the Behavioral Sciences*. 2nd ed. New York: McGraw-Hill; 1988:285.
24. Ashcroft RE, Chadwick DW, Clark SR, et al. Implications of socio-cultural contexts for the ethics of clinical trials. *Health Technol Assess*. 1997;1:1–65.