



Published in final edited form as:

*J Empir Res Hum Res Ethics*. 2009 September ; 4(3): 21–31. doi:10.1525/jer.2009.4.3.21.

## Challenges in Measuring a New Construct: Perception of Voluntariness for Research and Treatment Decision Making

**Victoria A. Miller,**

The Children's Hospital of Philadelphia

**William W. Reynolds,**

Richard Stockton College of New Jersey

**Richard F. Ittenbach,**

Cincinnati Children's Hospital Medical Center

**Mary Frances Luce,**

Duke University

**Tom L. Beauchamp, and**

Georgetown University

**Robert M. Nelson**

U.S. Food and Drug Administration

### Abstract

RELIABLE AND VALID MEASURES OF RELEVANT constructs are critical in the developing field of the empirical study of research ethics. The early phases of scale development for such constructs can be complex. We describe the methodological challenges of construct definition and operationalization and how we addressed them in our study to develop a measure of perception of voluntariness. We also briefly present our conceptual approach to the construct of voluntariness, which we defined as the perception of control over decision making. Our multifaceted approach to scale development ensured that we would develop a construct definition of sufficient breadth and depth, that our new measure of voluntariness would be applicable across disciplines, and that there was a clear link between our construct definition and items. The strategies discussed here can be adapted by other researchers who are considering a scale development study related to the empirical study of ethics.

### Keywords

voluntariness; decision making; methodology; scale development

---

SCALE DEVELOPMENT IS CONCERNED WITH measuring phenomena that we believe to exist but that cannot be observed directly (DeVellis, 1991). In the social sciences, these phenomena are referred to as constructs (often said to be concepts in the humanities), and measures are considered proxies for these constructs. When we assess the relationships between measures, we infer relationships between the constructs they are intended to measure.

Empirical research ethics has been a growing area, and with that growth comes the need for reliable and valid measures of phenomena that are relevant to ethical problem solving. The process of scale development for subjective phenomena involved in ethical problem solving, such as quality of life and decision making, is complex, especially with imprecisely defined or inadequately measured constructs. In this situation, refining the construct definition and operationalizing the construct can be challenging. Failure to fully address these processes can result in a number of unanticipated consequences, including problems with validity and biased estimates of relationships between variables (MacKenzie, 2003). Although many papers describe the development of new questionnaires to measure subjective health-related constructs, most focus on the psychometrics of the final measure and lack important details about the early stages of scale development. To move forward and to provide useful insights for solving ethical problems in real-world research settings, attention to these early phases of scale development for research ethics-related constructs is critical. Without reliable and valid measures of such constructs, research findings are of questionable value.

The purpose of this paper is to describe how we addressed the challenges of construct definition and operationalization when developing an instrument to measure the perception of voluntariness in parents making protocol-based treatment decisions (defined below) for their seriously ill children. We defined perception of voluntariness as the perception of control the individual perceives over a specific decision. Our goal is not to discuss well-accepted principles of measurement such as scaling responses or statistical evaluation of items, but to expand on the details of a project to operationalize a construct that has never been measured before and about which there may be some debate in the literature. We will show the way in which a multidisciplinary, multifaceted approach facilitated our progress in the early phases of scale development. The strategies we describe can be considered by others when faced with similar circumstances. First, we briefly summarize the background and methods of the project. Next, we discuss the importance of construct definition and the strategies we used to define the construct of voluntariness. We then discuss construct operationalization and the strategies used to develop items for a self-report measure of voluntariness. Finally, we describe the construction of the final experimental item pool, which was then tested for its psychometric properties.

## Description of Project

A fundamental principle of research ethics is that the validity of consent reflects the degree to which consent is informed and voluntary. The empirical literature on consent has focused on disclosure, understanding, and decision making, with little attention paid either to whether the decision was made voluntarily or to the conceptual conditions of voluntariness. Absent a measure of whether consent decisions are voluntary, we are left with only a partial picture of the adequacy of the consent process. The goal of the overall project was to develop a reliable and valid instrument to measure the perception of voluntariness in parents making decisions about protocol-based treatment for their seriously ill children. Protocol-based treatment refers to protocols that offer an intervention to the child, whether proven or unproven (i.e., research), and require a signed consent document. These decisions can be quite stressful for parents, as the decisions typically involve a wealth of potentially complex information, must be made within a short period of time, involve uncertain outcomes, and often occur in the emotional context of a new diagnosis, relapse (e.g., cancer), or illness exacerbation (e.g., cystic fibrosis). There has been much debate about the extent to which individuals with a life-threatening illness can make voluntary decisions and resist enticements or manipulations of hope; this debate extends to parents of seriously ill children and applies to both research and treatment decisions. We intend for our instrument to be used in either research or treatment contexts, as we expect that the *measurement* of perception of voluntariness is similar for these two types of decisions. However, the *factors that may influence* perceptions of voluntariness are likely to be different

across contexts. This question can be tested in future research that utilizes our new instrument. This new instrument will redress the current imbalance in the assessment of both treatment and research decision making and allow us to explore the factors that influence perception of voluntariness. Research using this new tool will also facilitate the development of guidelines for ensuring the voluntariness of decisions made in treatment and research settings.

Our collaborative research team consists of three principal investigators from three institutions, as well as three additional individuals from the primary site. The three principal investigators include a physician with expertise in research ethics (RN), a philosopher who has written extensively on informed consent (TB), and a researcher in marketing with a focus on decision making and affect (MFL). Additional on-site members include a clinical psychologist (VM), a qualitative social science researcher (WR), and a quantitative psychologist (RI).<sup>1</sup>

In the first phase of the study, focus groups were conducted to examine themes related to voluntariness. Individual interviews supplemented the use of focus groups due to scheduling difficulties and concerns about socially desirable responding in a group format. In addition, we determined that we would not lose any valuable information about perception of voluntariness by shifting to the use of interviews. The focus group data up to that point allowed us to identify what questions were most important to ask. Participants included parents who had made protocol-based treatment decisions for their seriously ill children, clinician-investigators who were involved in obtaining permission from these parents, and study coordinators and other non-physician clinicians involved in the consent process (e.g., social workers, nurses). Participants were approached from several divisions at a pediatric teaching hospital, including oncology, the neonatal intensive care unit, and the cardiac intensive care unit. Focus groups and interviews were audio-taped, transcribed, and analyzed for themes related to voluntariness. These data were used to refine our construct definition and generate candidate items to be included in the experimental item pool. In the second phase of the study, we administered the experimental item pool and additional instruments to parents within 10 days of making a decision for protocol-based treatment for their seriously ill child. The items were tested for their psychometric properties to determine which items would remain in the final instrument. (Empirical data and information about informed consent procedures and IRB approval will be reported elsewhere.)

### Defining the Construct of Perception of Voluntariness

**OVERVIEW OF THE CHALLENGE**—The first step when developing a new scale is to define (i.e., specify the meaning of) the construct of interest. A clear definition of the construct and its boundaries is critical, because this definition will guide the generation and selection of items to be included in the experimental item pool. Different construct definitions will lead to different sets of items. “The boundaries of the phenomenon must be recognized so that the content of the scale does not inadvertently drift into unintended domains” (DeVellis, 1991, p. 51). In addition, an ambiguous definition creates difficulty when deciding about the inclusion and exclusion of specific items based on their psychometric properties. For example, low correlations between individual items and the total score of a questionnaire can result from poor items or items that assess a multidimensional construct. The wrong theory about the construct can also result in items that end up having no predictive or explanatory power, which will have meant wasted time and resources (Streiner & Norman, 1995). In addition, the process of construct validation, which is concerned with whether the indicators appropriately reflect the construct (Wallander, 1992), depends on a clear and unambiguous definition.

<sup>1</sup>The social science researcher (WR) and quantitative psychologist (RI) moved to other institutions but remained involved in the project.

A challenge related to construct definition is how to differentiate the construct of interest from other constructs. This differentiation will be important when testing later for convergent and discriminant validity, both of which are aspects of construct validity. Convergent validity refers to how closely the new measure is related to measures of similar constructs, or other measures of the same construct. These correlations should be moderate to high, but not so high that the two instruments appear to be measuring exactly the same thing. Conversely, discriminant validity refers to whether the measure is unrelated to measures of dissimilar constructs. If there is an unexpected correlation between measures, this may have resulted from faulty measurement or an incorrect theory of the construct of interest. When defining the construct, it is important to consider how the construct of interest is expected to relate (or not relate) to other constructs (Campbell & Fiske, 1959; Cronbach & Meehl, 1955). This pattern of relationships will provide evidence regarding how well the new instrument and the construct it purports to measure perform in relationship to other constructs.

Other constructs may: (1) be related to the construct of interest, (2) confuse interpretation of the construct of interest, or (3) moderate relationships between the construct of interest and other constructs. Distinguishing between constructs is important to determine in advance, because this will guide item development and the analytic plan. For example, items that reflect correlated constructs may generate a scale that meets the statistical criteria for unidimensionality, especially when the number of items is high (McGrath, 2005). That is, covariation of items does not necessarily imply conceptual redundancy. As such, construct clarity at the outset is crucial so that items reflecting different constructs are excluded from the item pool. Typically, the identification of related constructs is an ongoing process that emerges in conjunction with the development of a concise definition of the construct of interest. In our study, we took a multifaceted approach to this process, through the use of a literature review, focus groups and interviews, and multidisciplinary group discussions. Each facet contributed to the development of our definition of perception of voluntariness, and laid the foundation for enhanced construct representation and decreased construct misspecification in the new instrument.

### **STRATEGIES USED TO REFINE THE CONSTRUCT DEFINITION**

**Literature Review:** We started with an examination of voluntariness in the literature. One broad definition is that voluntariness is the free power of choice without undue influence or coercion (Nuremberg Code, 1949), but this definition fails to specify what it means to be free of undue influence and coercion. Building on the notion of “choice,” Wall (2001) argues that “voluntariness is the degree of control that an agent has over his own behavior” (Wall, 2001, p. 130). In the case of informed consent, the relevant behavior of interest would be decision making. Similarly, Faden and Beauchamp (1986) analyze voluntariness as the degree to which an individual controls a decision and link the notion of control to the individual’s resistance to influence. Their approach favors a subjective interpretation of influence, because there is likely to be wide variability in individuals’ resistance to particular influences (Faden & Beauchamp, 1986). We can know only if an influence was controlling the individual’s decision by assessing the individual’s subjective perception; the mere presence of an influence cannot tell us anything about the individual’s response to that influence. The distinction between a subjective and objective interpretation of influence was important from the outset of the study, as we proposed to develop a scale for measuring voluntariness from a parent’s subjective perspective. Our approach focuses on perception of voluntariness, but we also recognize that some conditions are incompatible with voluntariness (e.g., lying with the intent to lead a person to believe what is false and to act on the false belief renders the action non-voluntary) even if the individual is unaware of the influence and perceives the decision as voluntary.

In contrast to a conceptualization of perception of voluntariness as control over the decision, Roberts presents multiple definitions of voluntariness, including ‘the individual’s ability to act in accordance with one’s authentic sense of what is good, right, and best in light of one’s situation, values, and prior history’ (Roberts, 2002, p. 707). Faden and Beauchamp (1986) frame authenticity more broadly as requiring either “stability or consistency” in the values underlying choice. Although quite different from the notion of control and not widely discussed in the literature, we decided to consider authenticity as a potentially important aspect of perception of voluntariness that could be explored further in focus groups. We did not want to reject prematurely any themes that might be relevant to the construct. Thus, at this early point, we framed perception of voluntariness as consisting of themes of control *and* authenticity.

There is a paucity of literature related to authenticity; the theme of control was the primary focus of further literature review. We examined control-related constructs in the psychological literature, such as locus of control and perceived behavioral control. Locus of control, as originally conceptualized by Rotter (1966), is related to beliefs about whether outcomes are causally dependent on variations in one’s own behavior versus external factors. In contrast, perceived behavioral control refers to the perception of control over behaviors, rather than the outcomes of those behaviors (Ajzen, 2002). According to Ajzen (2002), “perceived behavioral control ... is comprised of two components: self-efficacy (dealing largely with the ease or difficulty of performing a behavior) and controllability (the extent to which performance is up to the actor)” (p. 680). The examination of these two constructs was deemed useful, by helping us to identify aspects of control that were potentially relevant to perception of voluntariness, particularly because philosophical treatments of voluntariness may not translate easily into a measurable psychological construct. Our focus on psychological constructs related to control was also consistent with our emphasis on the perceptions of the individual making the decision rather than what is actually happening in the environment.

**Focus Groups and Interviews:** The various conceptualizations of voluntariness and related constructs (i.e., freedom from coercion, control/non-control, locus of control, perceived behavioral control, self-efficacy, and authenticity) were explored further in focus groups and individual interviews. We were also interested in generating additional themes related to perception of voluntariness, because participants’ views about voluntariness may differ from psychological and philosophical treatments of the construct. Prior to the focus groups, we developed a detailed interview guide to ensure that we explored key themes identified in the literature. The groups began with open-ended questions about the consent process. We avoided use of the words “voluntariness,” “control,” and “authenticity” at the beginning of the groups, to prevent biasing participants to view the decision in a particular way and to facilitate the generation of new themes. As each group progressed, questions became more specific and prompted for reactions to different definitions of voluntariness.

A review of the data suggested that key stakeholders in the informed consent process have varying interpretations of voluntariness. In general, research personnel and clinician-investigators approached voluntariness from a regulatory perspective and discussed potential influences on voluntariness (e.g., inducements, time pressure), rather than the construct of voluntariness *per se*. Overall, data from the parent focus groups and interviews were more useful, presumably because parents are better able to reflect on their experiences regarding decision making for their seriously ill children. Parents reported that a sense of control meant they perceived that the decision was up to them, were aware of the option to reject or withdraw from the protocol, and perceived that the decision was made without feeling pressured, intimidated, or coerced. One new theme that parents identified was the experience of dissociation from the decision-making process as a result of feeling shocked by their child’s recent medical diagnosis. Similarly, physicians noted that parents are often “paralyzed” at the time of diagnosis and that voluntary decision making appears difficult in this context.

Presumably, this is one way in which a lack of control can be manifested. The theme of authenticity did not resonate with participants. This appeared to be the result of difficulties parents encountered in the attempt to grasp and apply the notion of authenticity to decision making about their children's serious illness. Overall, the focus group data allowed us to explore specific factors that may reflect a perception of control (or lack thereof) in the context of protocol-based treatment decisions.

**Multidisciplinary Group Discussions:** Multidisciplinary group discussions were utilized to discuss the literature review and focus group data and served as a vehicle through which our definition of perception of voluntariness evolved over time. The use of a multidisciplinary group of experts has several advantages, including access to the most recent thinking in the field and a range of knowledge, experience, and opinions about the construct (Streiner & Norman, 1995). Furthermore, this approach assures that the construct is tackled from multiple angles, rather than one particular viewpoint (Streiner & Norman, 1995). This is particularly important when the construct intersects with several disciplines (e.g., philosophy, medicine, psychology, bioethics). If approached from a single point of view, the work would have limited usefulness and applicability to a broader audience. In our study, a multidisciplinary approach facilitated a definition of perception of voluntariness that would be coherent and acceptable from multiple perspectives. Meetings with the on-site team members occurred weekly, while meetings with the entire team were held on a quarterly basis.

As we refined our construct definition, the group addressed the differentiation between perception of voluntariness and related constructs. This differentiation is important so that the measure does not become tainted by unrelated factors (MacKenzie, 2003). For example, the decision about whether to include authenticity in our definition of voluntariness was a source of debate in the early stages of the project and was compounded by the fact that it is not widely discussed with clarity and precision in the literature. Some team members were hesitant to discard authenticity as a component of voluntariness before we had actual data to shed light on this issue. Others argued for a more concise definition of voluntariness before the item pool was finalized and data collection began. This assertion was due to the concern that the inclusion of multiple themes in our definition of voluntariness would result in a scale that would be contaminated by other constructs and difficult to interpret. Group discussions addressed the tensions among these different perspectives until a consensus could be reached. We eventually reached consensus that the perception of authenticity, although of interest both philosophically and psychologically, is largely irrelevant to a study of perception of voluntariness. An inauthentic choice may be voluntary, and a non-voluntary choice may be consistent with 'one's situation, values, and prior history' (p. 707; Roberts, 2002).

**FINAL CONSTRUCT DEFINITION—**Our final construct definition, consistent with the views of Faden and Beauchamp (1986) and Wall (2001), was that perception of voluntariness is the degree of control the individual perceives that he or she has over the specific decision about protocol-based treatment. Both Wall (2001) and Faden and Beauchamp (1986) refer to control as a continuum; these conceptualizations are consistent with control-related constructs in the psychological literature (e.g., Ajzen, 2002).

### **Operationalizing the Construct**

**OVERVIEW OF THE CHALLENGE—**Operationalization is the process of linking a construct definition to one or more specific, concrete indicators that can be measured, such as items on a self-report questionnaire. This can be challenging when there are no existing instruments to measure the construct. In addition, the development of actual items must attend to issues of clarity and content validity, to assure that the items assess all relevant content (i.e., content validity) and avoid tapping into other constructs (i.e., discriminant validity). Potential

sources of items include focus groups, key informant interviews, clinical observation, theory, expert opinion, and existing measures (DeVellis, 1991; Streiner & Norman, 1995). As mentioned earlier, the most important guide to choosing items is the definition of the construct. In classical scale development, the assumption is that responses to items are caused by the construct, such that each item should give an indication of the strength of the latent variable (DeVellis, 1991). That is, higher scores on any single item are indicative of higher levels of the construct of interest. In addition, the instrument should include multiple items to assess similar content. Redundancy in assessing the construct is important for developing a scale with adequate psychometric properties (DeVellis, 1991). Because items with poor psychometric properties are omitted after testing, the experimental scale is usually longer than the final scale. For this reason, the experimental item pool can tolerate more redundancy than the final instrument. However, the need for redundancy should also be balanced with the overall length of the item pool. When the response burden to participants is too high, recruitment may become more difficult and/or the quality of responses may decline (Schmitt & Stuits, 1985). Strategies we used to operationalize perception of voluntariness included adaptation of existing instruments and generation of new items based on focus group data and literature review.

### STRATEGIES USED TO OPERATIONALIZE THE CONSTRUCT

**Adaptation of Existing Instruments:** The process of operationalization was made easier by linking perception of voluntariness to control-related constructs in the psychological literature. Particular attention during item development was paid to measures of locus of control and self-efficacy. Additional instruments related to decision making were also examined for their relevance to perception of voluntariness. Relevant instruments that we examined included the Self-Determination Scale (Sheldon, Ryan, & Reis, 1996), the Rotter Locus of Control Scale (Rotter, 1966), the Multidimensional Health Locus of Control Scale (Wallston, Wallston, & DeVellis, 1978), the General Self-Efficacy Scale (Bosscher & Smit, 1998; Sherer et al., 1982), the Decision Self-Efficacy Scale (Bunn & O'Connor, 1996; Cranney et al., 2002), the Admission Experience Survey-Short Form (Gardner et al., 1993), and the Decisional Conflict Scale (Bunn & O'Connor, 1996; O'Connor, 1995). Items from these instruments were used to facilitate the generation of items that assessed perception of voluntariness.

We created a spreadsheet that included verbatim items from the instruments listed above; items were linked with their original source so that we could track the derivation and original meaning of items as they were edited. In group discussions, each item was examined for its potential usefulness as an indicator of perception of voluntariness (e.g., whether responses to items were indicative of voluntariness according to our construct definition). Items were discarded when they bore no relevance to the construct, and items were retained and adapted when the team agreed that they might be potential indicators of perception of voluntariness. Table 1 shows examples of original source items from these instruments and the adapted items used in our final experimental item pool. For example, items from the locus of control scales were relevant because they assessed beliefs about control, but they were not, in and of themselves, appropriate indicators of perception of voluntariness. These items assess a trait-like characteristic of individuals (e.g., an enduring characteristic of the person that applies across situations), refer to general domains of functioning, and assess perceptions of control over important outcomes. As such, items related to locus of control were adapted so that they (1) assessed control during a specific period of time, (2) referred to the specific decision about protocol-based treatment or research, and (3) assessed perceptions of the decision-making process and the decision itself, not the outcomes associated with the decision. Items from the other measures were adapted in a similar manner, so that they eventually bore little resemblance to their source items but instead reflected the construct of voluntariness as we had defined it.

The goal of this stage of scale development was to ensure that we were being inclusive in our approach to item generation, so items were discarded only after extensive group discussion. For example, we eventually concluded that items reflecting self-efficacy were not relevant indicators of perception of voluntariness, because self-efficacy refers to confidence in one's ability to perform a task or behavior in the future, not a task that has already been completed (e.g., a decision that was made).

**Generation of New Items:** In addition to discarding items that were not relevant to perception of voluntariness, we generated items to assess content that was not addressed by existing measures. For example, some focus group participants noted that their lack of control during decision making was expressed through feelings of detachment and passivity. As such, we added items to assess these experiences to the experimental item pool (e.g., "I was passive in the face of this decision"). New items were added to the spreadsheet as they were generated, to ensure that we did not discard potentially relevant items prematurely. Similar to the items adapted from existing instruments, these items were revisited over time and discarded or edited as appropriate.

The use of a multidisciplinary team ensured that we targeted all relevant content and that items made sense conceptually from multiple perspectives. Throughout this process, team members struggled with differentiating perception of voluntariness from related constructs, which was important for determining which items would go into the experimental item pool. For example, some team members felt that items related to the perception of options (e.g., "There were multiple options worth considering") should be included in the item pool. Other team members, coming from a measurement perspective, asked if this was consistent with our definition of perception of voluntariness. That is, would we be willing to say that agreement with this statement was indicative of higher levels of voluntariness? In addition, some team members noted that the focus group data suggested that the availability of options is independent of the individual's control over the decision. By discussing these questions, eventually we decided that these items should not be included in the item pool, but that it would be desirable to include them somewhere in the questionnaire packet. We agreed that responses to these items could inform our understanding of voluntariness and generate hypotheses for future research, by providing preliminary data about the conditions under which perception of voluntariness is more or less likely.

To clarify our thinking about how to use additional items to inform our understanding of perception of voluntariness, we utilized a matrix, with one construct (e.g., perception of options) on the horizontal axis and the other (perception of voluntariness) on the vertical axis (Table 2). The four individual cells reflect the various intersections of the two constructs; for example, parents who perceived that there were no options but still felt in control of the decision are reflected in the bottom left cell of the matrix. The matrix illustrated how, after data collection was complete, we could assign parents to one of the four categories, based on their perception of voluntariness scores and their responses to the item about perception of options. This approach has enabled us to identify separate constructs and the boundaries of voluntariness and to examine the factors that differentiate between the parents in the different cells. For example, in cases in which it was not possible to fill in off-diagonal cells, we were not able to conclude that the relevant construct overlapped with voluntariness.

### Construction of the Experimental Form

In tandem with the adaptation of existing instruments and generation of new items, we edited and deleted items for clarity in order to arrive at our final experimental item pool. Attention was paid to reading level, jargon, ambiguity, length of items, and double-barreled items, which contain more than one question, each of which can be answered in a different way. Each of



these factors can threaten the validity of the instrument if it impacts how individuals respond to the items. Items were revised by individual team members and discussed and edited further in group meetings. Any single item may have gone through several revisions, as we refined our construct definition, adapted and generated items, and examined the items for redundancy, structure, and clarity. While devising the experimental item pool, we also finalized the additional instruments and demographic variables that would be administered with the questionnaire packet. These instruments were chosen with an eye toward establishing the construct validity of the new scale and informing our understanding of perception of voluntariness. We included measures of mood (Profile of Mood States–Bipolar; Lorr & McNair, 1988), coping (Coping Inventory for Stressful Situations; Endler & Parker, 1990), decision-making preferences (Krantz Health Opinion Survey; Krantz, Baum, & Wideman, 1980) (Autonomy Preference Index; Ende et al., 1989), trust (Trust in Physician Scale; Pearson & Raeke, 2000) (Trust in Medical Researchers; Hall et al., 2006), self-efficacy for decision making (Decision Self-Efficacy Scale; Bunn & O'Connor, 1996, Cranney et al., 2002), and social desirability (Social Desirability Scale-17; Stober, 2001). These instruments were adapted using minor wording changes so that they were applicable to parents making medical decisions for their children.

After piloting the experimental item pool and other questionnaires with 17 parents from the target population, items were discarded or revised as appropriate. Reasons for discarding or revising items included repetitiveness, double-barreled format, and ambiguous meaning. After this process, the final experimental item pool contained 28 items, all of which assessed the parent's perception of control over decision making. The next phase of the study involved administration of the final experimental item pool and additional instruments to parents, within 10 days of making a decision for protocol-based treatment for their seriously ill children. Classical test theory guided our approach to scale construction (Streiner & Norman, 1995). This approach includes exploratory and confirmatory factor analysis, which is considered to be one source of information regarding the construct validity of an instrument (Thompson & Daniel, 1996).<sup>2</sup>

## Summary

We have highlighted two challenges that can arise when developing an instrument to measure a construct that is not clearly defined in the literature and has never been measured before: refining the construct definition and operationalizing the construct. Addressing these two challenges constituted a major portion of the project, spanning a period of ten months from the first focus group to the final item pool. This time period does not include the work that went into exploring potential meanings of the construct and sources of items prior to the first focus group. It is surprising that construct definition and operationalization are rarely discussed in much detail, given their importance for accurate measurement and validity. The development of reliable and valid measures of key constructs in human research ethics is essential to helping IRB members and researchers assess decision making (e.g., voluntariness, understanding) in ethically problematic settings. In addition, our new measure of perception of voluntariness can be used in future research to test important ethical questions related to undue influences on both treatment and research decision making in medical settings.

## Best Practices

A number of recommendations can be made regarding scale development for constructs relevant to human research ethics. The use of a multidisciplinary team was particularly fruitful

---

<sup>2</sup>The final instrument will be available on the University of California Press website once it is published, pending appropriate copyright approval.

for us. For example, this approach enabled us to do collective interpretive work that could not have been achieved individually (Siltanen, Willis, & Scobie, 2008). Given the complex nature of voluntariness, the benefits of a team approach were enhanced by the fact that team members viewed the issue through different lenses. This allowed us to give careful thought to existing literature and avoid the biases that may result when viewing a construct from one particular view-point. We were also careful not to reject prematurely potential themes or items that might be relevant to the construct. Although this process created tension at certain points and involved revisiting the same debates over time, a multidisciplinary approach allowed us to develop a more comprehensive understanding of our construct and to ensure that the end product (i.e., an instrument to measure perception of voluntariness) would be acceptable across different audiences. Both our construct definition and item pool were subjected to multiple revisions throughout this process. The dissemination of empirical findings to multiple audiences to which human subjects research is relevant is consistent with our multidisciplinary approach.

Additional strategies included literature reviews, qualitative methods, and adaptation of existing measures. For example, the use of focus groups and interviews ensured that our definition of perception of voluntariness was informed by the views of key stake-holders in the informed consent process. The integration of all of these components was critical in the early phases of scale development, by facilitating careful exploration of the construct definition, assuring that our item pool assessed all relevant domains of the construct, placing our construct definition and instrument in the context of existing literature, and providing a strong foundation for enhancing the construct validity of the final instrument.

Though this process took time, it must not be short-changed in the instrument development process. Additional research, using the new measure of perception of voluntariness, is needed before recommendations can be made for enhancing the voluntariness of both research and treatment decisions made in medical settings.

## Research Agenda

Once the Decision Making Control Instrument is complete, additional research using the instrument is needed to explore different questions related to informed consent. For example, how does the presence of a serious medical condition influence voluntary choice? Under what conditions do financial inducements for research participation pose an undue influence to informed consent? How do features of the informed consent process affect the voluntariness of a choice? Are there patient or clinician characteristics that make voluntary choice less likely? Answers to these questions are needed to develop interventions to enhance the voluntariness of both research and treatment decisions made in medical settings.

Interventions aimed at improving informed consent would be foolhardy unless soundly based on empirical research using reliable and valid measures of the key constructs that are relevant to human subject research and clinical ethics. It is likely that rigorous scale development studies need to be undertaken for constructs such as autonomy, vulnerability, and cultural competence. In addition, it would be beneficial to conduct systematic reviews of existing measures of decision making competence and understanding and appreciation of the elements of consent. Attention to psychometrics, validity, appropriate settings for use, and feasibility would be important, so that researchers in this area can choose which measures will best fit their needs.

## Educational Implications

It is important for investigators working in the area of human research ethics to understand key concepts related to instrument development and use instruments with acceptable levels of reliability and validity. Journal clubs to discuss these issues are one good forum in which to become familiar with currently used instruments in empirical ethics research, critique those

instruments, and generate ideas for the participants' own work. Workshops or conferences can also be used in this manner. As mentioned earlier, literature reviews should be undertaken to review existing instruments and identify gaps that should be addressed in future instrument development studies. These reviews are important for investigators and would also be useful for familiarizing other members of the research ethics community, such as IRB members, with the benefits and shortcomings of available measures.

## Acknowledgments

This research was supported in part by grants from the National Science Foundation (SES-0527618; Drs. Nelson, Luce, and Beauchamp) and the National Institutes of Health/National Cancer Institute (R21CA118377-01A1; Dr. Nelson) and an Institutional Development Grant to The Center for Research Integrity, Department of Anesthesiology and Critical Care Medicine, The Children's Hospital of Philadelphia.

## Authors' Biographical Sketches

**Victoria A. Miller** is a pediatric psychologist at The Children's Hospital of Philadelphia and Assistant Professor of Anesthesiology and Critical Care Medicine at the University of Pennsylvania School of Medicine. Her current research focuses on informed consent, child assent, and parent-child collaborative decision making for the management of childhood chronic illness.

**William Reynolds** is Assistant Professor of Social Work at The Richard Stockton College of New Jersey. His research has focused on the informed consent and assent decisions of parents, adolescents, and children in the context of pediatric medical research. He is a co-investigator on the project described in the present article.

**Richard F. Ittenbach** is Associate Professor of Pediatrics in the Division of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center. His current research focuses on the intersection of traditional scale development methods and generalized linear models and applications. Increasingly, his work emphasizes operational definitions of emerging biobehavioral constructs for children with life-threatening illnesses and disabilities.

**Mary Frances Luce** is Thomas A. Finch, Jr. Professor of Business Administration at The Fuqua School of Business at Duke University. She is interested in consumer behavior, medical decision making, and the effects of negative emotion on decision behavior. She co-authored *Emotional Decisions: Tradeoff Difficulty and Coping in Consumer Choice*.

**Tom L. Beauchamp** is a member of the Department of Philosophy and also Senior Research Scholar at the Kennedy Institute of Ethics, Georgetown University. In 1976, he wrote the bulk of *The Belmont Report* (1978) for the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. He co-authored (with James Childress) *Principles of Biomedical Ethics* (Oxford University Press, 1st ed., 1979; 6th ed., 2009), co-authored (with Ruth Faden) *A History and Theory of Informed Consent* (Oxford University Press, 1986), and authored *Philosophical Ethics* (McGraw Hill, 1st ed., 1982; 3rd ed., 2001). In 2004, Beauchamp was given the Lifetime Achievement Award of the American Society of Bioethics and Humanities (ASBH) in recognition of his contributions and publications in bioethics and the humanities. Earlier, in 1994, Indiana University awarded Beauchamp its "Memorial Award for Furthering Greater Understanding and Exchange of Opinions between the Professions of Law and Medicine."

**Robert M. Nelson** is Pediatric Ethicist in the Office of Pediatric Therapeutics, Office of the Commissioner at the U.S. Food and Drug Administration. The work reported in this article was conducted prior to Dr. Nelson joining the U.S. Food and Drug Administration, and do not

represent the views and/or policies of the FDA or the U.S. Department of Health and Human Services. Dr. Nelson's research explores aspects of child assent and parental permission such as adolescent risk perception, the development of a child's capacity to assent, and the degree to which parental choice is perceived as voluntary.

## References

- Ajzen I. Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology* 2002;32(4):665–683.
- Bosscher RJ, Smit JH. Confirmatory factor analysis of the general self-efficacy scale. *Behavior Research and Therapy* 1998;36(3):339–343.
- Bunn H, O'Connor A. Validation of client decision-making instruments in the context of psychiatry. *Canadian Journal of Nursing Research* 1996;28(3):13–27. [PubMed: 8997937]
- Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin* 1959;56(2):81–105. [PubMed: 13634291]
- Cranney A, O'Connor AM, Jacobsen MJ, Tugwell P, Adachi JD, Ooi DS, Waldegger L, Goldstein R, Wells GA. Development and pilot testing of a decision aid for postmenopausal women with osteoporosis. *Patient Education and Counseling* 2002;47(3):245–255. [PubMed: 12088603]
- Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychological Bulletin* 1955;52(4):281–302. [PubMed: 13245896]
- Devellis, RF. *Scale Development: Theory and Applications*. Vol. Vol. 26. Sage Publications; Newbury Park, CA: 1991. Applied Social Research Methods Series
- Ende J, Kazis L, Ash A, Moskowitz MA. Measuring patients' desire for autonomy: Decision making and information-seeking preferences among medical patients. *Journal of General Internal Medicine* 1989;4(1):23–30. [PubMed: 2644407]
- Endler NS, Parker JD. Multidimensional assessment of coping: A critical evaluation. *Journal of Personality and Social Psychology* 1990;58(5):844–854. [PubMed: 2348372]
- Faden, RR.; Beauchamp, TL. *A history and theory of informed consent*. Oxford University Press; New York: 1986.
- Gardner W, Hoge SK, Bennett N, Roth LH, Lidz CW, Monahan J, et al. Two scales for measuring patients' performance and perceptions of coercion during hospital admission. *Behavioral Sciences and the Law* 1993;20:307–321. [PubMed: 10150233]
- Hall MA, Camacho F, Lawlor JS, Depuy V, Sugarman J, Weinfurt K. Measuring trust in medical researchers. *Medical Care* 2006;44(11):1048–1053. [PubMed: 17063137]
- Krantz DS, Baum A, Wideman MV. Assessment of preferences for self-treatment and information in health care. *Journal of Personality and Social Psychology* 1980;39(5):977–990. [PubMed: 7441487]
- Lorr, M.; McNair, DM. *Profile of Mood States: Bipolar Form*. Educational and Industrial Testing Service; San Diego, CA: 1988. (manual)
- Mackenzie SB. The dangers of poor construct conceptualization. *Journal of the Academy of Marketing Science* 2003;31(3):323–326.
- Mcgrath RE. Conceptual complexity and construct validity. *Journal of Personality Assessment* 2005;85(2):112–124. [PubMed: 16171410]
- Nuremberg Code. *Trials of War Criminals before the Nuremberg Military Tribunals under Control Council Law No 10, Vol. 2, Nuremberg, October 1946–April 1949*. Washington, DC: U.S. Government Printing Office; 1949.
- O'Connor AM. Validation of a decisional conflict scale. *Medical Decision Making* 1995;15(1):25–30. [PubMed: 7898294]
- Pearson SD, Raeke LH. Patients' trust in physicians: Many theories, few measures, and little data. *Journal of General Internal Medicine* 2000;15(7):509–513. [PubMed: 10940139]
- Roberts LW. Informed consent and the capacity for voluntarism. *American Journal of Psychiatry* 2002;159(5):705–712. [PubMed: 11986120]
- Rotter JB. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs* 1966;80:1–28. [PubMed: 5340840]

- Schmitt N, Stuits DM. Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement* 1985;9(4):367–373.
- Sheldon KM, Ryan R, Reis HT. What makes for a good day? Competence and autonomy in the day and in the person. *Personality and Social Psychology Bulletin* 1996;22(12):1270–1279.
- Sherer M, Maddux JE, Mercadante B, Prentice-dunn S, Jacobs B, Rogers RW. The self-efficacy scale: Construction and validation. *Psychological Reports* 1982;51:663–671.
- Siltanen J, Willis A, Scobie W. Separately together: Working reflexively as a team. *International Journal of Social Research Methodology* 2008;11(1):45–61.
- Stöber J. The Social Desirability Scale-17 (SDS-17): Convergent validity, discriminant validity, and relationship with age. *European Journal of Psychological Assessment* 2001;17(3):222–232.
- Streiner, DL.; Norman, GR. *Health Measurement Scales: A Practical Guide to their Development and Use*. Oxford University Press; New York: 1995.
- Thompson B, Daniel LG. Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement* 1996;56(2):197–208.
- Wall E. Voluntary action. *Philosophia* 2001;28(1–4):127–136.
- Wallander JL. Theory-driven research in pediatric psychology: A little bit on why and how. *Journal of Pediatric Psychology* 1992;17(5):521–535. [PubMed: 1432479]
- Wallston KA, Wallston BS, Devellis R. Development of the multidimensional health locus of control (MHLC) scales. *Health Education Monographs* 1978;6:160–170. [PubMed: 689890]

## Adapted Instruments and Items.

TABLE 1

Instrument	What It Measures	Example of Original Item	1st Revision for New Instrument	Final Revision for New Instrument
Self-determination Scale	Beliefs about whether one perceives oneself as the origin of one's actions ("trait autonomy.")	"What I do is often not what I'd choose to do."	"I was not really free to decide what I wanted."	<ul style="list-style-type: none"> <li>"I was not really free to decide what I wanted. (reverse-scored)"</li> <li>"I was free to decide what I wanted."</li> </ul>
Rotter Locus of Control Scale	Beliefs about whether outcomes are causally dependent on variations in one's own behavior	"By taking an active part in political and social affairs, the people can control world events."	"By taking an active part in decision part in decision making, I was able to control the decision that was made."	"I was actively involved in this decision."
Multi-dimensional Health Locus of Control Scale	Beliefs about whether one's health outcomes are causally dependent on variations in one's own behavior.	"If I get sick, it is my own behavior which determines how soon I get well again."	"It was my own behavior that determined the decision that was made."	"I had an influence on the decision about the protocol."
General Self-efficacy Scale	Beliefs about one's ability to successfully execute behaviors required to produce particular outcomes.	"When I make plans, I am certain I can make them work."	"When confronted with this decision, I was certain that I could make the decision for myself."	Omitted
Decision Self-efficacy Scale	Beliefs about one's ability to make effective decisions about treatment options.	"I feel confident that I can express my concerns about each choice."	"When making this decision, I felt confident that I could express my medical team."	"I was able to express my point of view about the decision."
Decisional Conflict Scale	Beliefs about one's uncertainty and effective decision making (e.g., feeling the choice is informed, values-based, and likely to be implemented) for a specific decision.	"I have enough support from others to make a choice."	"When making this decision, I felt supported in making my own choice."	"I was the one to choose."
Admission Experiences Survey-Short Form	Perceptions of coercion when being admitted for psychiatric hospitalization.	"People tried to force me to come to the hospital."	"Someone took this decision away from me."	<ul style="list-style-type: none"> <li>"Someone took this decision away from me." (reverse-scored)</li> <li>"Others made this decision against my wishes." (reverse-scored)</li> </ul>

**TABLE 2**

Example of Matrix Approach to Differentiating Voluntariness from Related Constructs.

<u>Perception of Voluntariness</u>			
		<b>Yes</b>	<b>No</b>
Perception of Options	Yes	There were several options, and I was in control of choosing one.	There were several options, but I was not in control of choosing one.
	No	There was only one option, and I was in control of choosing it (e.g., It was my choice but I had no choice)	There was only one option, and I was not in control of choosing it.