

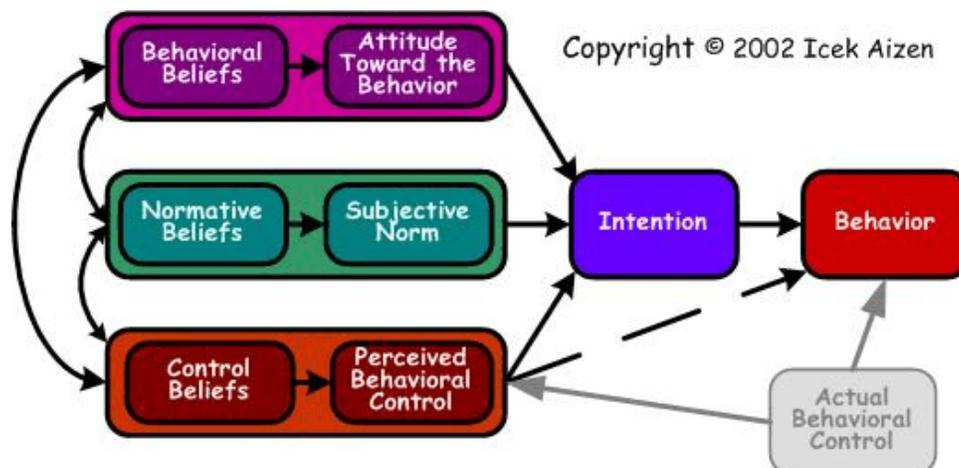
Constructing a TpB Questionnaire: Conceptual and Methodological Considerations

September, 2002

Icek Ajzen

Brief Description of the Theory of Planned Behavior

According to the theory of planned behavior, human action is guided by three kinds of considerations: beliefs about the likely outcomes of the behavior and the evaluations of these outcomes (behavioral beliefs), beliefs about the normative expectations of others and motivation to comply with these expectations (normative beliefs), and beliefs about the presence of factors that may facilitate or impede performance of the behavior and the perceived power of these factors (control beliefs). In their respective aggregates, behavioral beliefs produce a favorable or unfavorable *attitude toward the behavior*; normative beliefs result in perceived social pressure or *subjective norm*; and control beliefs give rise to *perceived behavioral control*. In combination, attitude toward the behavior, subjective norm, and perception of behavioral control lead to the formation of a behavioral *intention*. As a general rule, the more favorable the attitude and subjective norm, and the greater the perceived control, the stronger should be the person's intention to perform the behavior in question. Finally, given a sufficient degree of *actual control* over the behavior, people are expected to carry out their intentions when the opportunity arises. Intention is thus assumed to be the immediate antecedent of behavior. However, because many behaviors pose difficulties of execution that may limit volitional control, it is useful to consider perceived behavioral control in addition to intention. To the extent that perceived behavioral control is veridical, it can serve as a proxy for actual control and contribute to the prediction of the behavior in question. The following figure is a schematic representation of the theory.



Latent Variables and Manifest Indicators

The theoretical constructs shown in the above diagram are hypothetical or latent variables. They cannot be directly observed but must instead be inferred from observable responses. This is as true of actual behavior as it is of the other constructs.

Behavior

The behavior of interest is defined in terms of its **T**arget, **A**ction, **C**ontext, and **T**ime (TACT) elements. Consider the case of *walking on a treadmill in a physical fitness center for at least 30 minutes each day in the forthcoming month*. Defining the TACT elements is somewhat arbitrary. Walking is clearly part of the action element, but we should probably also include 30 minutes a day in this element. The treadmill could be considered the target and the physical fitness center the context, or we may prefer to view the fitness center as the target and the treadmill as the context. The time element refers to when the behavior is performed, and in this example it is defined as the forthcoming month.

Compatibility. No matter how the TACT elements of the behavior are defined, it is important to observe the *principle of compatibility* which requires that all other constructs (attitude, subjective norm, perceived behavioral control, and intention) be defined in terms of exactly the same elements. Thus, the attitude compatible with this behavior is the attitude toward *walking on a treadmill in a physical fitness center for at least 30 minutes each day in the forthcoming month*, the subjective norm is the perceived social pressure to do so, perceived behavior control refers to control over performing the defined behavior, and we must assess the intention to perform this very behavior.

Specificity and generality. The TACT elements in the above example are quite specific, but it is possible to increase the generality of one or more element by means of *aggregation*. In fact, the time element “in the forthcoming month” is already defined at a more general level than, say, “next Tuesday at 5:00 pm.” To obtain a measure of the behavior in our example, we have to aggregate observations over the course of a whole month. Looking at behavior on only a single occasion is usually too restrictive to be of much practical value. Similarly, in many cases we may not be particularly interested in a specific context. Thus, we may want to predict and understand the behavior of walking on a treadmill, irrespective of the context (at home, in a gym, at a friend’s place) in which it occurs. We can generalize the context element by recording how often the behavior is performed in all relevant contexts.

A comparable argument can be made with respect to the action element. We may focus on exercising in general, instead of walking on a treadmill, in which case we would have to generalize across such different forms of exercise as running, walking, swimming, and aerobics. When we do this, however, we have to explicitly describe the behavior for our respondents. Simply asking them about “exercising” is ambiguous and attitudes toward exercising can be affected by recent experiences that temporarily raise the accessibility of one or another type of exercise.

and rating scale) could be included, and an estimate of internal consistency computed.

Predictor Variables

All predictors in the theory of planned behavior can be assessed directly, by asking respondents to judge each on a set of scales. In addition, attitude toward the behavior, subjective norm, and perceived behavioral control can also be measured indirectly, on the basis of the corresponding beliefs. In either case, the measures must be directly compatible with the behavior in terms of action, target, context, and time elements.

Direct Measures

Investigators often mistakenly assume that direct measures of the theory's constructs are obtained by asking a few arbitrarily selected questions, or by adapting items used in previous studies. Although this approach often yields findings of interest, it can produce measures with relatively low reliabilities and lead to an underestimate of the relations among the theory's constructs and of its predictive validity. To secure reliable, internally consistent measures, it is necessary to select appropriate items in the formative stages of the investigation. Different items may have to be used for different behaviors and for different research populations. In the final questionnaire, the different items assessing a given construct are usually separated and presented in nonsystematic order, interspersed with items for the other constructs.

Intention

Several items are used to assess behavioral intentions, as shown in the following examples.

- ▶ I intend to walk on a treadmill for at least 30 minutes each day in the forthcoming month
extremely unlikely :____:____:____:____:____:____:____: extremely likely
- ▶ I will try to walk on a treadmill for at least 30 minutes each day in the forthcoming month
definitely true :____:____:____:____:____:____:____: definitely false
- ▶ I plan to walk on a treadmill for at least 30 minutes each day in the forthcoming month
strongly disagree :____:____:____:____:____:____:____: strongly agree

Care should be taken to ensure that the intention items selected in the pilot study have acceptable psychometric qualities. At the very least, the set of items to be used must be shown to correlate highly with each other (i.e., that the measure has high internal consistency).

Attitude Toward the Behavior

Any standard attitude scaling procedure (Likert scaling, Thurstone scaling) can be used to obtain a respondent's evaluation of the behavior, but due largely to its ease of construction, the semantic differential is most commonly employed. To make sure that the bipolar adjectives selected for inclusion are in fact evaluative in nature (for the behavior and population of interest), the investigator should start with a relatively large set, perhaps 20 to 30 scales. The initial set can be taken from the list of published adjective scales that, across concepts and populations, tend to load highly on the evaluative factor of the semantic differential (Osgood, Suci, & Tannenbaum, 1957). A small subset of scales that exhibit high internal consistency is selected for the final attitude measure. This selection can rely on item-total correlations (Likert's criterion of internal consistency), or on an analysis of reliability (e.g., Cronbach's alpha). If factor analysis is to be used, then the original set of scales should also include adjective pairs that tend to load highly on the other two major factors of the semantic differential: potency and activity.

A second criterion for item selection has to do with the qualitative aspects of evaluation represented by the adjective scales. Attitude toward a behavior is defined as a person's overall evaluation of performing the behavior in question. However, empirical research has shown that overall evaluation often contains two separable components. One component is instrumental in nature, represented by such adjective pairs as *valuable* — *worthless*, and *harmful* — *beneficial*. The second component has a more experiential quality and is reflected in such scales as *pleasant* — *unpleasant* and *enjoyable* — *unenjoyable*. It is recommended that the final set of scales include adjective pairs of both types, as well as the *good* — *bad* scale which tends to capture overall evaluation very well. Care should be taken to counterbalance positive and negative endpoints to counteract possible response sets. To illustrate, a measure of attitude toward the behavior could take the following form.

For me to walk on a treadmill for at least 30 minutes each day
in the forthcoming month is

harmful : _____ : _____ : _____ : _____ : _____ : _____ : _____ : beneficial

pleasant : _____ : _____ : _____ : _____ : _____ : _____ : _____ : unpleasant

good : _____ : _____ : _____ : _____ : _____ : _____ : _____ : bad

worthless : _____ : _____ : _____ : _____ : _____ : _____ : _____ : valuable

enjoyable : _____ : _____ : _____ : _____ : _____ : _____ : _____ : unenjoyable

Subjective Norm

Several different questions should be formulated to obtain a direct measure of subjective norm. The following items illustrate the format these questions can take.

- ▶ Most people who are important to me think that

I should : _____ : _____ : _____ : _____ : _____ : _____ : _____ : I should not
walk on a treadmill for at least 30 minutes each day
in the forthcoming month

- ▶ It is expected of me that I walk on a treadmill for at least 30 minutes each day in the forthcoming month

extremely likely : _____ : _____ : _____ : _____ : _____ : _____ : _____ : extremely unlikely

- ▶ The people in my life whose opinions I value would

approve : _____ : _____ : _____ : _____ : _____ : _____ : _____ : disapprove
of my walking on a treadmill for at least 30 minutes
each day in the forthcoming month

Items of this kind have an *injunctive* quality, consistent with the concept of subjective norm. However, responses are often found to have low variability because important others are generally perceived to approve of desirable behaviors and disapprove of undesirable behaviors. To alleviate this problem, it is recommended that the measure of subjective norm usually also include items designed to capture *descriptive* norms, i.e., whether important others themselves perform the behavior in question. Some injunctive questions can be reformulated to take on a descriptive quality.

Examples:

- ▶ Most people who are important to me walk on a treadmill for at least 30 minutes each day

completely true : _____ : _____ : _____ : _____ : _____ : _____ : _____ : completely false

- ▶ The people in my life whose opinions I value

walk : _____ : _____ : _____ : _____ : _____ : _____ : _____ : do not walk
on a treadmill for at least 30 minutes each day

As was true for the measures of behavior, intention, and attitude toward the behavior, the investigator must make sure that the final set of injunctive and descriptive items used to measure subjective norm has a high degree of internal consistency.

Perceived Behavioral Control

A direct measure of perceived behavioral control should capture people's confidence that they are capable of performing the behavior under investigation. A number of different items

have been used for this purpose. Some items have to do with the difficulty of performing the behavior, or with the likelihood that the participant could do it. Items of this kind are often said to capture the respondent's sense of *self-efficacy* with respect to performing the behavior:

- ▶ For me to walk on a treadmill for at least 30 minutes each day in the forthcoming month would be

impossible : _____ : _____ : _____ : _____ : _____ : _____ : _____ : possible

- ▶ If I wanted to I could walk on a treadmill for at least 30 minutes each day in the forthcoming month

definitely true : _____ : _____ : _____ : _____ : _____ : _____ : _____ : definitely false

Other items used to assess perceived behavioral control refer to the behavior's *controllability*. These items address people's beliefs that they have control over the behavior, that its performance is or is not up to them. The following are examples of this type of item.

- ▶ How much control do you believe you have over walking on a treadmill for at least 30 minutes each day in the forthcoming month?

no control : _____ : _____ : _____ : _____ : _____ : _____ : _____ : complete control

- ▶ It is mostly up to me whether or not I walk on a treadmill for at least 30 minutes each day in the forthcoming month

strongly agree : _____ : _____ : _____ : _____ : _____ : _____ : _____ : strongly disagree

The perceived behavioral control scale should contain self-efficacy as well as controllability items, and care should again be taken to make sure that the set of items to be used has a high degree of internal consistency.

Belief-Based Measures

Beliefs play a central role in the theory of planned behavior. They are assumed to provide the cognitive and affective foundations for attitudes, subjective norms, and perceptions of behavioral control. If this assumption is correct, then we can rely on beliefs to obtain indirect, belief-based measures of these constructs. Like the direct measures described earlier, such belief-based measures can serve as manifest indicators of the latent constructs and we would expect the two types of indicators (direct and belief-based) to be correlated. This point is sometimes misunderstood. Behavioral beliefs are assumed to determine attitude toward the behavior; they are *not* assumed to determine the direct measure of attitude. Similarly, normative beliefs determine subject norm, not the direct measure of subjective norm, and control beliefs determine perceived behavioral control, not the direct measure of perceived behavioral control.

Belief-based and direct measures are fallible indicators of the same underlying, latent constructs. When evaluating the theory of planned behavior, therefore, it would be a mistake to model the belief-based measures as antecedents or causes of the direct measures.

Accessible beliefs. There are also some important differences between direct and indirect measures. By measuring beliefs, we can, theoretically, gain insight into the underlying cognitive foundation, i.e., we can explore why people hold certain attitudes, subjective norms, and perceptions of behavioral control. This information can prove invaluable for designing effective programs of behavioral intervention. It is important to realize, however, that this explanatory function is assumed only for *salient* beliefs or, to use the currently favored term, beliefs that are readily *accessible* in memory.

Internal consistency. Another difference has to do with the way in which the reliabilities of direct and belief-based measures can be estimated. The earlier description of direct measures emphasized the need to ensure high internal consistency in our measures of behavior and in the measures of intention, attitude, subjective norm, and perceived behavioral control. This is a minimal requirement to confirm the assumption that the items selected do in fact assess the same underlying construct. Each item is, by itself, designed to be a direct measure of the theoretical construct, and the different items used to assess the same construct should correlate with each other and exhibit high internal consistency.

For theoretical reasons, this requirement is not imposed on the belief-based measures of attitude, subjective norm, and perceived behavioral control. Accessible behavioral beliefs are assumed to account for attitude toward the behavior, accessible normative beliefs for subjective norm, and accessible control beliefs for perceived behavioral control. However, no assumption is made that accessible beliefs are internally consistent. People's attitudes toward a behavior can be ambivalent if they believe that the behavior is likely to produce positive as well as negative outcomes. And the same is true of the set of accessible normative beliefs and the set of accessible control beliefs. Consequently, internal consistency is not a necessary feature of belief-based measures of attitude, subjective norm, and perceived behavioral control. It is in their aggregate that they provide a single manifest indicator of the latent construct.

Temporal stability. Internal consistency is one way of assessing a measure's reliability. This criterion can reasonably be applied to direct measures of the theory's constructs, but not to belief-based measures. A second way to estimate reliability is to examine a measure's temporal stability (test-retest reliability). This criterion can be applied to the direct as well as to the belief-based measures. Temporal stability is in fact an important characteristic in prospective studies that attempt to predict behavior at a later point in time. If measures of the theory's constructs lack temporal stability, they cannot be expected to predict later behavior.

Eliciting Accessible Behavioral Outcomes

Pilot work is required to identify accessible behavioral, normative, and control beliefs. Respondents are given a description of the behavior and are asked a series of questions illustrated

following equation:

$$A_B \propto \sum b_i e_i$$

Belief strength is multiplied by outcome evaluation, and the resulting products are summed over all accessible behavioral outcomes.

Optimal scaling. In this example, belief strength and outcome evaluation are both scored in a *unipolar* fashion, from 1 to 7, with higher numbers representing greater subjective probabilities and more favorable evaluations, respectively. Alternatively, it would be possible to use *bipolar* scoring, from –3 to +3, such that low probabilities and unfavorable evaluations would be represented by negative numbers and high probabilities and favorable evaluations by positive numbers. Even though the shift from unipolar to bipolar scoring involves a simple linear transformation (subtraction by 4), it results in a *nonlinear* transformation of the product term (be). This can be seen in the following computation where the original values of b are transformed by the addition of a constant B , and the values of e by a constant E . For simplicity, only one behavioral belief is entered into the expectancy – value equation.

$$\begin{aligned} A_B &= (b + B)(e + E) \\ &= be + Eb + Be + BE \end{aligned}$$

Because two of the new terms, Eb and Be , are not constants, the resulting expression is clearly not a linear transformation of be . The problem this creates is that moving from unipolar to bipolar scoring, or vice versa, can have a substantial impact on the correlation of the belief-based attitude estimate with any other variable. Thus, it may raise or lower by a considerable degree the correlation between belief-based and direct attitude measures, a correlation that is often used to validate the expectancy – value model itself.

Unfortunately, there is no a priori way to determine the proper scaling of belief strength and outcome evaluation. It could reasonably be argued that outcome evaluations should receive bipolar scoring because the low end of the scale represents a negative evaluation of the outcome and the high end a positive evaluation. A similar argument, however, cannot be made with respect to the measure of belief strength. To be sure, belief strength is identified with the probability that performing a behavior will produce a given outcome, and it may thus appear to require unipolar scoring, perhaps from 0 to +1, to correspond to the metric of objective probabilities. (A linear transformation to attain this scoring scheme would subtract a constant of 1 and divide by 6.) However, subjective probabilities do not necessarily have the same properties as objective probabilities. Consider, for example, the belief that walking a treadmill for 30 minutes each day can cause serious physical injury. Assuming that a respondent rates this outcome extremely negatively (–3) but believes it to be highly unlikely, should the low probability judgment be scored as 0 or become a score of –3 on the original 7-point scale? The correct scoring depends on the way in which the respondent interprets and uses the likelihood scale. If it is used in the manner of an objective probability, then this particular belief would make no contribution to the attitude (due to the multiplication by 0). In other words, because

suffering a serious physical injury is considered highly improbable, this outcome has no bearing on the attitude toward walking the treadmill. Alternatively, respondents who assign an extremely low probability to suffering a serious physical injury may regard this to be a highly positive characteristic of walking the treadmill. Perhaps in their view, a major advantage of this form of exercise, compared to other types of exercise, is the low likelihood of suffering serious physical harm. The low belief strength, combined with a negative evaluation of the outcome, should therefore make a strongly *positive* contribution to the attitude toward the behavior. This would be accomplished by using bipolar scoring for belief strength and outcome evaluation.

Because there is no obvious theory-based method to determine proper scoring, we must rely on an empirical solution. In light of the success and general acceptance of the expectancy – value model, we can make its (presumed) validity the criterion of optimal scaling. That is, we can examine the correlation between belief-based and direct measures of attitude and adopt the scoring scheme (unipolar or bipolar) that produces the better result.

A mathematical solution is described in Ajzen (1991). Let B represent the constant to be added or subtracted in the rescaling of belief strength, and E the constant to be added or subtracted in the rescaling of outcome evaluations. The expectancy-value model shown earlier can then be rewritten as follows:

$$A_B \propto \sum (b_i + B)(e_i + E)$$

Expanding, this expression and disregarding the constant BE , we obtain:

$$A_B \propto \sum b_i e_i + B \sum e_i + E \sum b_i$$

To estimate the rescaling parameters B and E , we regress the standard attitude measure, which serves as the criterion, on $\sum b_i e_i$, $\sum b_i$, and $\sum e_i$, and then divide the unstandardized regression coefficients of $\sum b_i$ and $\sum e_i$ by the coefficient obtained for $\sum b_i e_i$. The resulting value for the coefficient of $\sum e_i$ provides a least-squares estimate of B , the rescaling constant for belief strength, and the value for the coefficient of $\sum b_i$ serves as a least-squares estimate of E , the rescaling constant for outcome evaluation.

In an easier, though less precise method of estimation, the investigator experiments by comparing the results of unipolar and bipolar scoring and then retains the scores that produce the stronger correlation between belief-based and direct attitude measures.¹

¹In this approach, the same scoring scheme is applied to all behavioral beliefs, under the implicit assumption that respondents use all belief scales in the same manner. It would also be possible to apply different scoring schemes to different beliefs.

Eliciting Accessible Normative Referents

The following questions can be asked to elicit the identity of relevant referent individuals and groups that are readily accessible in memory.

- ▶ Are there any individuals or groups who would *approve* of your walking on a treadmill for at least 30 minutes each day in the forthcoming month?
- ▶ Are there any individuals or groups who would *disapprove* of your walking on a treadmill for at least 30 minutes each day in the forthcoming month?
- ▶ Are there any other individuals or groups who come to mind when you think about walking on a treadmill for at least 30 minutes each day in the forthcoming month?

Measuring Normative Beliefs

The assessment of normative beliefs follows a logic similar to that involved in the measurement of behavioral beliefs. Two questions are asked with respect to each referent.

Example: Assume that “my family” is one of the accessible referents.

- ▶ *Normative belief strength (n)*

My family thinks that

I should :_____:_____:_____:_____:_____:_____:_____: I should not
walk on a treadmill for at least 30 minutes each day in the forthcoming month

- ▶ *Motivation to comply (m)*

Generally speaking, how much do you want to do what your family thinks you should do?

not at all :_____:_____:_____:_____:_____:_____:_____: very much

Measures of normative belief strength and motivation to comply with respect to each accessible referent offer a “snap shot” of perceived normative pressures in a given population. An overall, belief-based measure of subjective norm (SN) is obtained by applying the expectancy – value formula to these measures, as shown in the following equation:

$$SN = \sum n_i m_i$$

As in the case of behavioral beliefs, optimal scaling of normative belief strength and motivation to comply must be determined empirically.

Elicitation of Accessible Control Factors

To generate a list of accessible factors that may facilitate or impede performance of the behavior, the following questions can be asked.

- ▶ What factors or circumstances would enable you to walk on a treadmill for at least 30 minutes each day in the forthcoming month?
- ▶ What factors or circumstances would make it difficult or impossible for you to walk on a treadmill for at least 30 minutes each day in the forthcoming month?
- ▶ Are there any other issues that come to mind when you think about the difficulty of walking on a treadmill for at least 30 minutes each day in the forthcoming month?

Measuring Control Beliefs

Again, two questions are asked with respect to each accessible control factor.

Example: Assume that one of the accessible control factors has to do with work-related demands on time.

- ▶ *Control belief strength (c)*

I expect that my work will place high demands on my time in the forthcoming month

strongly disagree : _____ : _____ : _____ : _____ : _____ : _____ : _____ : strongly agree

- ▶ *Control belief power (p)*

My work placing high demands on my time in the forthcoming month would make it

much more difficult : _____ : _____ : _____ : _____ : _____ : _____ : _____ : much easier
for me to walk on a treadmill for at least 30 minutes each day

Examination of the average strength and power of the different control beliefs provides a picture of the factors that are viewed as facilitating or impeding performance of the behavior. Using an expectancy – value formulation, as shown in the following formula, it is possible to compute a belief-based measure of perceived behavioral control (PBC).

$$PBC = \sum c_i p_i$$

As was true of attitudes and subjective norms, an optimal scaling analysis should be done to determine the proper scoring of control belief strength and power prior to finalizing the belief-based measure of perceived behavioral control.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179-211.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.