

# The Development of the Reversal Theory State Measure

Mitzi L. Desselles, Stephanie L. Murphy, and Evan R. Theys  
Louisiana Tech University

Metamotivational states of mind are the cornerstones of reversal theory, yet a comprehensive measure covering all eight states and suitable across a range of research scenarios has yet to be developed. Most previous measures have been limited to a subset of states (typically the telic-paratelic pair) or applicable to specific populations or situations (e.g., athletes, dieters). The present research reports on the development of a comprehensive motivational state measure, leveraging the work on previous instruments and following a traditional psychometric protocol. The objective was to develop a condensed state measure suitable for multiple administrations per day via smartphone or similar technologies with minimal interruption. After several rounds of data collection, analysis and revision, three versions of the Reversal Theory State Measure (RTSM) emerged: a short, 3-item “bundled” version, a 12-item “branched” logic version, and a “long” 18-item version. All versions measure the four domains in reversal theory. The three RTSM versions and their respective scoring logic are being shared with the reversal theory community as an open-source instrument to encourage state-based research. Much of the previous research on the theory has focused on dominance. Perhaps the availability of an instrument to assess moment-to-moment changes within individuals will stimulate more research on the role of states.

*Keywords:* state measure, metamotivational state, reversal theory

The essential feature of reversal theory is the proposition that individuals can feel and behave differently at different times, even in the same situation (Apter, 1982, 2001a, 2007). Intra-individual changes may occur within a week, a day, or even minutes. The theory asserts that human experience is fluid and changeable. However, rather than being erratic and enigmatic, the theory asserts that experience has an interpretable, dynamic structure. Such an underlying structural model of phenomenology acknowledges human inconsistency while providing a framework for understanding intra-individual changes in how the world is experienced (Apter, 2001b).

The present article describes the development of a state-based measure of motivation that captures intra-individual changes. The Reversal Theory State Measure (RTSM) assesses the four metamotivational domains described in the theory: means/ends, rules, orientation, and interaction domains (e.g., Apter, 2001a). The aspiration is that this measure (or its successors) becomes the ‘gold standard’ for use in state-based research testing the dynamic hypotheses outlined by Apter (2001b, 2013). For example, empirical stud-

ies of constructs such as motivational richness, reversibility, and leadership microclimate require measurement at the state level.

## Previous State Measures

While a number of state measures may be found in the existing literature research, all have limitations for testing reversal theory-based hypotheses. Instruments designed outside the reversal theory framework tend to measure different constructs from those needed to test hypotheses from the theory. For example, Heatherton and Polivy (1991) developed the State Self-Esteem Scale (SSES), which consists of 20 items measuring performance, social, and appearance self-esteem. Vealey (1986) developed the State Sport-Confidence Inventory to assess athlete’s confidence levels surrounding competitions.

Some state measures reported in previous research were not originally designed to measure states. For example, Timmermans, Mechelen, and Kuppens (2010) studied intra-individual variability in core affect by having participants repeatedly fill out the Affect Grid (Russell, Weiss, & Mendelsohn, 1989). Such research starts from a confusing premise: why is there *change* in what we assume to be *constant*? Other studies report on measures that may be used as state or trait measures. For example, The State-Trait Cheerfulness Index (STCI; Ruch, Köhler, & Van Thriel, 1996, 1997) has a state version as well as a trait version. The only distinction lies in the instructions; the state measure reads “refer to your

---

Correspondence concerning this article should be addressed to Mitzi Desselles, Ph.D., Department of Psychology and Behavioral Sciences, Louisiana Tech University, P.O. Box 10048, Ruston, LA 71270. Telephone: (318) 257-2361; Fax: (318) 257-3442; E-mail: mdessel@latech.edu

current mood and mental state,” whereas the trait measure reads “refer to your mood and mental state in general” (p. 480). Similarly, the widely used Positive and Negative Affect Scales (PANAS; Watson, Clark, & Tellegen, 1988) may be used to assess various timeframes, based on differences in the instructions. Respondents may be asked to describe how they feel “right now (that is, at the present moment)” or “in general, that is, on the average” (p. 1070). Thus, the PANAS may be used as a measure of state-affectivity as well as trait-affectivity. Arguably, simply shifting the time referent in the instructions while keeping items the same may not sufficiently reflect the conceptual distinctions between a transitory intra-individual variable (state) and a more enduring inter-individual difference (trait).

In contrast, The State-Trait Personality Inventory (STPI; Spielberger, 1986, 1989; Spielberger & Reheiser, 2004) maintains the conceptual distinction between states and traits. The items and response options used to measure states are different from those used to assess traits. A STPI *state* item such as “I am jittery” is measured on a 4-point scale from “not at all” to “very much so.” A sample STPI *trait* item is “I am quick tempered,” measured on a 4-point scale from “almost never” to “almost always” (Spielberger, 1986).

A number of reversal theory-based state measures have been developed and have been useful for their intended purposes (see reviews by Apter & Desselles, 2001; O’Connell, & Cook, 2001). However, previous measures of metamotivational states have been limited in various ways. Some measure only a subset of reversal theory (RT) states: the Telic State Measure (TSM; Svebak & Murgatroyd, 1985), the Somatic State Questionnaire (SSQ; Cook, Gerkovich, Potocky, & O’Connell, 1993), the Telic/Paratelic State Instrument (T/PSI; Calhoun, 1995; O’Connell & Calhoun, 2001), the Negativism State Measure (O’Connor, 1992), the Autic Mastery-Sympathy State Measure (O’Connell & Brooks, 1997), and the Ecological Momentary Sampling Tool developed by O’Connell, Gerkovich, Bott, Cook, and Shiffman (2000). Other measures are intended for specific audiences. For example, the State of Mind Indicator for Athletes (SOMIFA) measures all eight states but is most relevant and applicable to sports research (Kerr & Apter, 1999). Two additional instruments have been designed to measure stress over a period of time: the Tension and Effort Stress Inventory (TESI; Svebak, 1991, 1993) and the Tension Stress Scale (TSS; Popkess-Vawter, Gerkovich, & Wendel, 2000; Popkess-Vawter, 1998; Wendel, 1999). The TESI assesses the reversal theory concepts of tension- and effort-stress which arise in all eight motivational states. Although not originally intended as a state measure, a modified form of the TESI has been used for that purpose (e.g., Kerr & Svebak, 1994).

The Motivational Style Profile (MSP; Apter, Mallows, & Williams, 1998) and instruments derived from it (e.g.,

the Apter Motivational Style Profile; AMSP; Apter International, 1999a) ask respondents to recall how often they experience each of the eight motivational states. The MSP measures *frequency* of states aggregated over some period of time and is therefore more closely aligned with reversal theory tenets than trait measures. However, the MSP was never intended to be a measure of state or sensitive to moment-to-moment changes.

### Considerations in Designing the State Measure

Prior research by O’Connell and her colleagues (see O’Connell & Cook, 2001 for a review), Apter and Larsen (1993), Young, Desselles, Lee, and Apter (2005), and Murphy and Desselles (2011) provide an important empirical foundation for the present work. These studies demonstrated it was possible to measure intra-individual variation in reversal theory states via repeated measures. In addition, some of the studies were able to connect intra-individual variation with meaningful constructs such as relapse among smokers and work/life balance. Different approaches to measuring state were used in these studies (e.g., item content, format, and data collection method varied by study). In reviewing previous research, we identified a number of considerations to be addressed in the design of a broadly applicable state measure, including length, instructions, item stem, response format, criteria for effectiveness, and whether to combine the two pairs of transactional states.

### Length

A major consideration in developing the state measure was length. Some would argue that the measure should be long to conform to established psychometric expectations (e.g., longer scales should be more reliable, all things being equal; Furr, 2011; Nunnally & Bernstein, 1994; Thorndike, 1982). Conversely, a short instrument may be more practical and less likely to interfere with on-going activities (Robins, Trzesniewski, Tracy, Gosling, & Potter, 2002). Robins, Hendin, and Trzesniewski (2001) reported on the development of a single-item measure of self-esteem, citing advantages such as elimination of redundant items and less “fatigue, frustration, and boredom associated with answering highly similar questions repeatedly” (p. 152). Burisch (1984, 1997) reported that long and sophisticated measures of depression are no more valid than short and simple ones. Gosling, Rentfrow, and Swann (2003) developed a 10-item inventory of the Big-Five personality domains and provided evidence on its validity and reliability. The assumed psychometric superiority of longer scales does not always materialize, and some theorists have advocated for single-item scales (Bergkvist & Rossiter, 2007, 2009; Rossiter, 2002; Drolet & Morrison, 2001).

Shorter measures may also be less prone to test reactivity, an important feature of any state instrument. The longer the

state measure, the more likely respondents may be to change states while responding. One can easily imagine respondents switching from the conforming to rebellious state when faced with very long surveys. We decided to explore both options: a longer version developed through standard test development procedures and a shorter version derived from the long version by selecting the items with highest factor loadings and inter-item correlations.

### Instructions

Instructions were written to assess the individual's motivational state at a particular point in time, as opposed to frequency or general tendency of experiencing each state (cf. Apter et al., 1998). Because answering a survey may trigger a particular state, respondents were asked to describe their motivation immediately prior to taking the survey. To anchor their responses, they were asked what they were doing just prior to the survey (e.g., working, shopping, studying, watching television, spending time with friends).

Instructions for the state measure also included statements intended to give respondents permission to be changeable (see *Psychological State Measure - Branched Version (RTSM-Br)* in the supplemental materials for this article). Self-report surveys often ask respondents to describe themselves in terms of traits or general tendencies (e.g., McCrae, & Costa, 1987). Including statements about the changeability of human nature was intended to focus respondents on their immediate state of mind, without feeling constrained to be consistent over time or responding with what they "ought" to be like.

### Stem

The distinction between being in a state (what one wants at that moment) versus being satisfied in a state (extent to which one is getting what one wants at that moment) is an important one. Because the measure was intended to assess motivation, the stem for all items was "I wanted to..." as opposed to "I was..." which would have measured the experienced outcome in the state.

### Response Format

The format of the items on the instrument presented us with another choice: should items have a forced-choice format with the states within a pair anchoring each end, or should there be separate items for each state with Likert scale response options? In alignment with the reversal theory principle that states are organized in opposing pairs, a forced-choice format was used. Each item was anchored at each end with one state within each pair. However, in order to obtain internal level data and maximize measurement variance, the forced-choice items were not purely dichotomous. Both characteristics would be important in the analysis of

the scale's psychometric properties, especially multivariate statistical analyses. Items included six scale points between anchors ranging from "a little more" to "a lot more" on each end. As a result, items could be scored to produce either nominal or interval data.

### Criteria

Several criteria were established for the reversal theory state measure: it must: a) assess all four pairs of states, b) be suitable for different sample populations, c) be applicable across research settings, d) be well-grounded conceptually, e) be bias-free across demographic subgroups, f) be sensitive enough to measure individual differences, as well as g) detect changes over time. In order to meet all seven criteria, two different versions of the scale were originally planned: a long version to conform to traditional psychometric expectations and a shorter version for use in experience sampling methodologies that require minimal interference with on-going activities.

### Combining Pairs of Transactional States

In preparing to write items, the question arose as to whether items should represent each of the transactional pairs separately ("uncrossed") or the combination of the two pairs ("crossed"). The rationale for combining states was based on the observation that the meaning of the mastery or sympathy state changes depending whether it is paired with alloic or autic (and vice versa). For example, wanting to be in control (autic-mastery) is qualitatively different from wanting someone else to be in control (alloic-mastery). While both are concerned with power (mastery), combining pairs of states enhances the precision with which state of mind is measured. Furthermore, wanting to receive affection (autic-sympathy) is a qualitatively different experience from wanting to give it (alloic-sympathy). This methodology has been previously used in constructing the Motivational Style Profile. Apter et al. (1998) chose to cross the transactional pairs in their items and created subscale scores for both the crossed and uncrossed transactional states based solely on the crossed items.

Uncrossed transactional items would contrast autic versus alloic states, with different items contrasting the mastery versus sympathy states. Writing crossed items in a forced-choice format required a slightly more complex construction, since there are four alternatives (autic-mastery, autic-sympathy, alloic-mastery and alloic-sympathy). Items were written to reflect the following contrasts: a) autic-mastery versus autic-sympathy, b) alloic-mastery versus alloic-sympathy, c) autic-sympathy versus alloic-sympathy, and d) autic-mastery versus alloic-mastery.

Writing items that purely contrast autic versus alloic states, or purely contrast mastery versus sympathy, proved challenging. Depending on how the item is written, items

Table 1  
*Examples of the Original 56 Items*

State-Pairs	Sample Item Anchors
Telic / Paratelic	Do something serious / Do something playful
Conforming / Negativistic	Do my duty / Be defiant
Mastery / Sympathy	Experience power / Experience caring
Autic / Alloic	Do something for myself / Do something for others
Autic-Mastery / Alloic-Mastery	Be strong / Strengthen others
Autic-Sympathy / Alloic-Sympathy	Be cared for / Care for others
Autic-Mastery / Autic-Sympathy	Be strong / Be loved
Alloic-Mastery / Alloic-Sympathy	Help others to be powerful / Show consideration for others

written to contrast mastery versus sympathy may appear to have an implied referent object. One example is the mastery/sympathy item “I want to experience power” versus “I want to experience caring.” While not explicitly stating who might hold power or receive caring, the item implies the autic state (i.e., desire for personal power versus caring directed toward oneself).

Despite difficulties in writing “pure” transactional items, both crossed and uncrossed items were constructed. The research team elected to pursue both approaches and let the data determine which approach, crossed or uncrossed, yielded more interpretable findings.

### Study 1: Development of the RTSM Long Version

#### Item Generation

A committee of five researchers knowledgeable in reversal theory generated items for all state pairs, including crossed and uncrossed transactional states. As a result, eight sets of items were generated, each reflecting a different set of contrasts between states, as shown in Table 1. Items were subjected to critical appraisal, especially concerning face validity. Criticism was obtained from a larger committee of referees, and, after a number of iterations, a 56-item scale was produced. Six to nine items represented each pair, and several items were reverse scored. The anchors for 26 of 56 items (46%) were reversed left-to-right, selected at random. Instructions and survey format were as described above.

Participants were asked to describe the activity in which they were engaged immediately before the survey. The following options were given: eating, socializing, talking/listening, watching TV, shopping, exercising, doing homework, working, and other (specify). The list was generated by the research team, based upon previous research conducted by Murphy and Desselles (2011). Demographic information was then collected (gender, race, ethnicity, class, college), followed by the state measure.

Table 2  
*Activities Prior to the Survey*

Activity	<i>N</i>	%
Eating	104	20.3%
Doing homework	103	20.1%
Talking/Listening to others	87	17.0%
Socializing	83	16.2%
Watching TV	33	6.4%
Working	31	6.1%
Taking an exam	26	5.1%
In class	15	2.9%
On Internet (browsing, email, social network)	13	2.5%
Shopping	2	0.4%
Exercising	1	0.2%
Other	21	4.1%

*Note.* Percentages sum to more than 100% due to multiple responses from each participant.

#### Sample

The 56-item scale was administered to a sample recruited on a university campus in the southeastern United States. Participants were recruited through various sources, such as classroom announcements, announcements on social networks, and announcements made at high-traffic areas of campus (e.g., student union, dining hall). Administration was via online survey or by paper and pencil questionnaire. Two versions of the paper version were used, differing only in the order of items. The online survey was available for approximately two weeks, and participants were able to take the survey at any time of the day.

Five hundred and twelve participants voluntarily completed the survey. Just over half of the surveys (54.9%) were taken online. Undergraduate students comprised 79.7% of the sample, graduate students 19.5%, and the remaining 0.8% were non-students (i.e., faculty and staff). The mean age of participants was 22.5 years, with a range from 18 to 57 ( $SD = 6.0$ ). The gender mix in the sample was 58.2% female and 41.6% male; 0.2% did not indicate gender. In terms of race and ethnic background, the sample was 73.8% white,

Table 3  
*Rotated Factor Loadings from Principal Component Analysis with Varimax Rotation*

State Pair	Item Anchors	F1	F2	F3	F4	F5	F6	F7	F8
CN (r)	Do what I'm not supposed to do   Do what I'm supposed to do	0.749							
CN (r)	Do the opposite of what's expected of me   Do what's expected of me	0.732							
CN (r)	Be defiant   Do my duty	0.697							
CN	Follow the rules   Break the rules	0.687							
TP (r)	Do something for its own sake   Do something significant	0.531							
CN	Be accepting   Be rejecting	0.484							
SSOS (r)	Care for others   Be cared for		0.715						
SSOS (r)	Help someone else   Be helped		0.68						
SSOS (r)	Indulge someone else   Be indulged		0.677						
SSOS (r)	Show others I like them   Be liked		0.569						
OMOS (r)	Look after others   Follow the lead of others		0.554						
SMSS	Be assertive   Be indulged		-0.429						
SMOM	Be personally successful   Help others to succeed			0.762					
SMOM	Seek personal development   Help others to develop			0.745					
SMOM	Be capable   Enhance the capability of others			0.652					
SMOM (r)	Create opportunities for others   Seize opportunities for myself			0.64					
SMOM (r)	Help others to have power   Be powerful			0.418					
SMSS (r)	Receive affection   Be masterful				0.74				
SMSS (r)	Be loved   Be strong				0.661				
SSOS	Be loved   Be loving towards others				-0.51				
OMOS (r)	Be affectionate to others   Enhance the capability of others			0.415					
SSOS	Receive affection   Be affectionate to others				0.423				
SMSS (r)	Be looked after   Dominate				-0.406				
SMSS	Be powerful   Be cared for					0.627			
SMSS	Be in control   Be helped					0.62			
TP	Do something crucial   Do something of no great concern					0.551			
TP	Do something serious   Do something playful						0.701		
TP	Accomplish something for the future   Enjoy myself at the moment						0.666		
TP	Do something that has broad consequences   Do something with few consequences						0.587		
OMOS	Enjoy someone else being in control   Indulge someone else						0.545		
SMOM (r)	Enjoy someone else being in control   Be in control							-0.732	
SMOM (r)	Follow the lead of others   Dominate							0.551	
OMOS	Strengthen others   Be loving towards others							0.466	
OMOS	Help others to be powerful   Show consideration for others					0.462			0.68
OMOS	Help others to succeed   Care for others								0.66
									0.614

*Note.* (r) denotes reverse coded item. TP = telic/paratelic, CN = conforming/negativistic, SMOM = autic-mastery/alloic-mastery, SSOS = autic-sympathy/alloic-sympathy, SMSS = autic-mastery/altoic-sympathy, and OMOS = alloic-mastery/altoic-sympathy. Loadings less than .40 have been omitted.  $N = 512$ .

Table 4  
*Items on the Long Version of the Reversal Theory State Measure (RTSM-L)*

State-Pairs	Sample Item Anchors
Telic / Paratelic	Do something crucial   Do something of no great concern Do something serious   Do something playful Accomplish something for the future   Enjoy myself at the moment
Conforming / Negativistic	Do what I'm not supposed to do   Do what I'm supposed to do ( <i>r</i> ) Do the opposite of what's expected of me   Do what's expected of me ( <i>r</i> ) Be defiant   Do my duty ( <i>r</i> )
Autic-Sympathy / Alloic-Sympathy	Care for others   Be cared for ( <i>r</i> ) Help someone else   Be helped ( <i>r</i> ) Indulge someone else   Be indulged ( <i>r</i> )
Autic-Mastery / Alloic-Mastery	Be personally successful   Help others to succeed Seek personal development   Help others to develop Be capable   Enhance the capability of others
Autic-Mastery / Autic-Sympathy	Be looked after   Dominate ( <i>r</i> ) Be powerful   Be cared for Be in control   Be helped
Alloic-Mastery / Alloic-Sympathy	Strengthen others   Be loving towards others Help others to be powerful   Show consideration for others Help others to succeed   Care for others

*Note.* (*r*) denotes reverse coded item.

18.0% African-American, 2.3% Asian, 1.8% Hispanic, 1.2% bi- or multi-racial, 0.8% Native American, 1.6% other, and 0.6% opted not to respond. The activities participants reported doing just prior to the survey are shown in Table 2.

## Results

Principal component analysis with varimax rotation was conducted on all 56 items, including both crossed and uncrossed transactional items. An orthogonal rotation method was chosen because the factors are independent of one another theoretically. Because the 56 items represented eight different contrasts between states, an 8-factor solution was forced, explaining 46.5% of the variance between items. Somatic state items (telic/paratelic and conforming/negativistic) showed the clearest factor groupings. The crossed transactional items loaded fairly well together as intended. The uncrossed autic/alloic items tended to load with crossed state items on two factors interpretable as: a) autic- vs. alloic-mastery and b) autic- vs. alloic-sympathy. The uncrossed autic/alloic items seemed to have an implied mastery or sympathy "flavor." The uncrossed mastery/sympathy items did not load highly on any factor. The mastery/sympathy distinction, without specifying whether autic or alloic, apparently did not reflect a meaningful latent construct.

All 12 uncrossed transactional items were dropped, and the principle component analysis with a varimax rotation was run on the 44 remaining items. An 8-factor solution emerged based on examination of the scree plot, with 49.1% of variance explained. Thirty-five of the forty-four total items had loadings above .40 on the eight factors. Only three items loaded on more than one factor. Rotated factor loadings are reported in Table 3. The factors, in order of extraction, were interpreted as: a) conforming/negativistic (CN), b) autic-sympathy/alloic-sympathy (SSOS), c) autic-mastery/alloic-mastery (SMOM), d) autic-sympathy combined with either autic-mastery (SSSM) or alloic-sympathy (SSOS), e) autic-mastery/alloic-sympathy (SMSS), f) telic/paratelic (TP), g) autic-mastery/alloic-mastery (SMOM), and h) alloic-mastery/alloic-sympathy (OMOS).

## RTSM Long Version

A long version of the Reversal Theory State Measure (RTSM-L) was created by taking the three highest loading items from the most clearly interpretable factors to form six subscales (see Table 4). An 18-item scale was deemed appropriately long to conform to established psychometric standards, but several concerns were noted. The major concern was whether the scale was too long, thereby increasing the

likelihood that an individual's state could change during the course of answering the items.

An additional concern arose in classifying an individual into one state within each crossed transactional pair. Each combination of transactional states (e.g., autic-mastery) was measured on two subscales (e.g., the autic-mastery/autic-sympathy subscale as well as the autic-mastery/alloic-mastery subscale). As a result, it was possible to classify some individuals into different combinations of transactional states on the two subscales. For example, an individual may be classified as autic-mastery on the autic-mastery/autic-sympathy subscale but as alloic-mastery on the autic-mastery/alloic-mastery subscale. This problem arose regardless of whether scores on individual items on each subscale were summed or whether each item on a subscale was scored dichotomously and then tallied. Examination of the original data ( $n = 512$ ) indicated that using the two scoring methods, as many as 48.9% of respondents would not be classified into the same one of four crossed transactional states (i.e., autic-mastery, autic-sympathy, alloic-mastery, or alloic-sympathy). In contrast, classification of participants into conforming or negativistic and telic or paratelic states was very consistent across scoring methods (96.5% and 94.7% respectively). Thus, the scoring dilemma for the Long Version appears to be limited to the transactional states.

### RTSM Branched Version

One proposed solution to the dilemma of scoring transactional states in the Long Version was to develop a version of the RTSM employing adaptive questioning. In this version, the individual's state on one *uncrossed* transactional domain (e.g., autic or alloic) is determined, then based on that result, the individual is presented with only the *crossed* mastery and sympathy items in that state. For example, an individual may be classified as being in the autic state based on the uncrossed, "pure" autic/alloic items. This person would then only be shown items to determine which autic state combination (i.e., autic-mastery or autic-sympathy). Based on the findings from the factor analysis that only the uncrossed autic/alloic items formed an interpretable factor, (the uncrossed mastery/sympathy items did not), we began the tailored logic chain with the uncrossed autic/alloic items. Thus, the RTSM Branched Version (RTSM-Br) consisted of 15 possible items, although each person responded to only 12. Everyone answered three conforming/negativistic items, three telic/paratelic items, and three autic/alloic items (in randomized order); these nine items were answered by all respondents. The final three items were either autic-mastery/autic-sympathy or alloic-mastery/alloic-sympathy, depending on the individual's responses to the autic/alloic items. There were no pure, uncrossed mastery/sympathy items in the RTSM-Br. The branching logic required dynamic programming, typically accomplished through computerized admin-

istration. The complete instrument and scoring instructions for the RTSM-Br are included as a supplemental document on the web page for this article.

Scoring the subscales of the RTSM-Br may be done by summing item scores but this may lead to split decisions. When these ties are present, dichotomous scoring of each item would resolve them. The disadvantage of dichotomous scoring is the loss of variance and lower level of measurement (nominal as opposed to interval). A decision to score interval-scaled items dichotomously leads one to ask, why not build the scale using dichotomous (or categorical) items from the outset?

### RTSM Bundled Version

A "bundled" version of the state measure was created in which all anchors for a state are shown at once in bundles. The respondent chooses which bundle is most descriptive of his state of mind. Researchers at the Midwest Research Institute in the 1990s (e.g., Kakolewski, Goings, O'Connell, Gerkovich, & Cook, 1996) sometimes referred to this approach as "splurging" (M. J. Apter, personal communication, December 8, 2010). A bundled approach has also been used in previous exploratory research (e.g., Murphy & Desselles, 2011; Young et al., 2005).

Forcing a choice between bundles is closely aligned with the reversal theory proposition that states exist in pairs of mutually exclusive opposites. A potential advantage of bundling is that respondents are given a more comprehensive definition of each state, and a clearer sense of the opposite nature of pairs. Conversely, it is conceivable that some respondents will fail to see how the anchors in each bundle fit together and become confused.

The RTSM Bundled Version (RTSM-B) consisted of one conforming versus negativistic item (two bundles of opposing anchors), one telic versus paratelic item (two bundles of opposing anchors), and one crossed transactional pairs item (four bundles of contracting anchors). Each bundle consisted of three anchors drawn from items on the RTSM-L. See *Psychological State Measure - Bundled Version (RTSM-B)* in the supplemental material on the web page for this article for the complete RTSM-B and scoring instructions.

### Study 2: Comparing Branched and Bundled Versions

To assess the extent of agreement between the RTSM-Br and RTSM-B, an additional 89 respondents completed both the branched and bundled versions in rotated order. All respondents were students at a university in the southeastern U.S., recruited through various sources (i.e., classroom announcements, social networks, and high-traffic areas of campus). All surveys were completed online.

Table 5  
*Frequency of States by Instrument and Scoring Method*

State	<i>RTSM-Br</i> (%)	<i>RTSM-B</i> (%)
Telic	68.5	70.8
Paratelic	31.5	29.2
Conforming	85.4	84.3
Negativistic	14.6	15.7
Autic	42.7	56.2 <sup>a</sup>
Alloic	57.3	43.8 <sup>a</sup>
Autic-Mastery	44.9	39.3
Autic-Sympathy	12.4	16.9
Alloic-Mastery	13.5	15.7
Alloic-Sympathy	29.2	28.1

*Note.* Data from the *RTSM-Br* and *RTSM-B* versions were obtained in rotated order from the same participants.  $N = 89$ .  
<sup>a</sup>Direct calculation of the state was not possible with the instrument. The figure reported is the net combined frequency of the applicable cross transactional pairs.

## Sample

Undergraduate students comprised 89.9% of the sample, while the remaining were graduate students (7.9%) or non-students (2.2%). The mean age of participants was 22.8 years, with a range from 18 to 49 ( $SD = 6.5$ ). The gender mix in the sample was 33.0% male and 67.0% female. Caucasians made up 64.0% of the participants, African-Americans 23.6%, Asians 3.4%, bi- or multi-racial 3.4%, Hispanics 1.1%, and 2.2% "other." An additional 2.2% chose not to respond to the question on race and ethnicity.

## Results

The percentage of participants classified into each state using the Branched (*RTSM-Br*) versus Bundled (*RTSM-B*) Versions is shown in Table 5. By inspection, the percentages for the telic/paratelic and conforming/negativistic states are nearly identical (1.1 to 2.3 percentage points different) across versions of the state measure. Differences between versions are slightly larger for the transactional states (1.1 to 5.6 percentage points).

Given the categorical nature of the state data, a non-parametric statistic, kappa ( $K$ ), was used to assess agreement between the *RTSM-Br* and *RTSM-B* in classifying participants into states, beyond that expected by chance (Siegel & Castellan, 1988). For the telic/paratelic pair of states, 77 out of 89 respondents (86.5%) were classified identically in both versions ( $K = .681$ ,  $p < .001$ ). For the conforming/negativistic pair of states, 82 out of 89 respondents (92.1%) were classified identically in both versions ( $K = .694$ ,  $p < .001$ ). For the crossed transactional states, 49 out of

89 respondents (55.1%) were classified identically into one of four possible states using the two versions ( $K = .552$ ,  $p < .001$ ). Consistency between versions of the state instrument (i.e., identical classification into one of four transactional states) was highest for the autic-mastery (63% to 71%) state and lowest for the alloic-mastery state (29% to 33%; see Table 6).

## Discussion

The present work describes the development of three versions of a *RTSM*: Long, Branched and Bundled. So, which is the best version? This question may be best answered by reviewing the criteria established at the outset and assessing how well each version performs on them. Table 7 summarizes key points from the discussion that follows.

The first criterion is whether the version covers all four pairs of states. All three versions cover all four pairs, albeit in slightly different fashions. The Branched Version is the most different from the other two, and the differences are most apparent in how the transactional pairs are assessed. Using adaptive logic, the *RTSM-Br* first assesses whether the respondent is in the autic or alloic state and, based on that determination, proceeds with questions of mastery versus sympathy. All three versions measure the somatic pairs similarly; the telic state is contrasted against the paratelic state, and the conforming state is contrasted against the negativistic state. The intended output from all three versions, however, is the same; each respondent may be classified into one of the four combinations of the transactional pairs (autic-mastery, autic-sympathy, alloic-mastery, alloic-sympathy), as well as one state from each of the two pairs of somatic states (telic versus paratelic and conforming versus negativistic). However, the previously described scoring difficulty in the Long Version is a significant weakness. As a result, both Branched and Bundled Versions are preferred over the Long Version on the first criterion.

Are the measures applicable across sample populations and across research settings? All versions appear equally suitable to adults, based on our sample. The Long Version may be limited to a narrower range of research settings than the Bundled and Branched Versions, simply because it is longer. The Bundled Version is shorter than the Long or Branched Versions, thus making it more respondent-friendly and potentially resulting in higher completion rates and less respondent reactivity in time sampling research protocols. A second disadvantage to the Branched Version is the need for advanced branching logic, thus limiting its use to dynamically-programmed surveys, such as those administered online. Thus, the Bundled Version appears preferable across a wider set of research settings than the other two versions.

Are the versions well-grounded conceptually? All versions were based on the same original set of 56 items and

Table 6  
*Frequency of States by Instrument*

		Bundled Version ( <i>RTSM-B</i> )				
		Telic		Paratelic	Total	
	Telic	56		5	61	
	Paratelic	7		21	28	
	Total	63		26	89	
		Conforming		Negativistic	Total	
Branched Version ( <i>RTSM-Br</i> )	Conforming	72		3	75	
	Negativistic	4		10	14	
	Total	76		13	89	
		Autic-Mastery	Autic-Sympathy	Alloic-Mastery	Alloic-Sympathy	Total
	Autic-Mastery	25	4	5	6	40
	Autic-Sympathy	2	6	1	2	11
	Alloic-Mastery	5	0	4	3	12
	Alloic-Sympathy	3	5	4	14	26
	Total	35	15	14	25	89

*Note.* Data from the *RTSM-Br* and *RTSM-B* versions were obtained from the same participants in rotated order.  $N = 89$ .

their underlying constructs. One of the core elements of the theory is that an individual cannot be partially in both states within a domain. The scoring indeterminacy observed using the Long and Branched Versions runs counter to this central principle of reversal theory. In contrast, this issue does not arise with the Bundled Version. Only the Bundled Version employs a true forced-choice format in clear fidelity with the oppositional structure of states in reversal theory. The other versions allow respondents to make incremental judgments using a 6-point scale, which may be then converted into a dichotomous variable. The Bundled Version has the additional advantage of providing respondents with a more holistic representation of each state using multiple descriptors at each endpoint. The purpose of bundling descriptors is to provide a clearer definition of each state and sharpen the distinction between alternatives. Although some may argue that the Bundled Version compromises measurement variance, we point to the fact that bi-stability is a core feature of reversal theory. Once in a given state, an individual cannot be in that state to a greater or lesser extent (e.g., one cannot be more or less telic at a given moment). States are either-or propositions (e.g., Apter, 2001a).

The fifth criterion (to be bias-free across demographic subgroups) requires more data before a firm conclusion may be drawn about the superiority of any version. Unfortunately, the demographic mix in the current dataset was too skewed to test for differences by age, gender, or race. We envision a series of studies with relatively large and demographically diverse samples. The various versions of the state measure would be compared for similar incidences of each state as well as psychometric properties (e.g., internal consistency;

only applicable to Long and Branched Versions) to determine whether they work similarly across groups.

Are the versions sensitive enough to measure intra- and inter-individual differences (the final two criteria)? The evidence reported in this research that a sizable number of individuals may not be classifiable into states by the Long Version is a serious concern regarding its sensitivity to assess inter-individual differences. More experimental evidence (for example, using a between-groups design in which states are manipulated) is necessary to arrive at robust conclusions regarding the scales' ability to differentiate between individuals in different states. We had hoped to compare the states of respondents engaging in different activities just prior to completing the state measure, but the incidence of each activity in the Branched and Bundled Versions dataset was too low. A repeated-measures design is recommended, in combination with either a descriptive within-person time-sampling approach or a more powerful experimental manipulation of states. Ideally, a series of manipulations would be identified targeting specific states, combinations of states, and none of the states. We would hope to see movement in the hypothesized direction on specific states following a strong manipulation and no movement when manipulation has no theoretical link to the state(s). The research protocol should also include dependent measures with no hypothesized relationship to the state manipulation. In this way, we will begin to assemble evidence regarding the validity of the various versions, via a multi-state, multi-method approach. Similarly, we would advocate for additional research to establish criterion-related validity. Examples of planned research include using the state measure to predict performance on spe-

Table 7  
*Summary of Preferred Versions by Criteria*

Criteria	Preferred Version(s)	Explanation
Coverage of all four state pairs	• Bundled or Branched	<ul style="list-style-type: none"> <li>• Branched and Bundled Versions allow for simple scoring and measurement of all pairs.</li> <li>• Long Version has scoring and categorization issues on the transactional pairs.</li> </ul>
Applicable across sample populations	• Any	<ul style="list-style-type: none"> <li>• All versions appear applicable across adult populations.</li> </ul>
Suitable across research settings	• Bundled	<ul style="list-style-type: none"> <li>• The Branched Version is not suitable for paper-and-pencil administration as transactional states must be dynamically scored for adaptive questioning to occur.</li> <li>• Long Version may be problematic in repeated measures designs due to overall length.</li> <li>• Bundled Version may be administered online as well as via paper or smartphone.</li> </ul>
Well-grounded conceptually	• Bundled	<ul style="list-style-type: none"> <li>• All versions mirror the conceptual and theoretical basis of the theory by forcing requiring individuals to choose between the states pairs within the given domain(s).</li> <li>• Bundled Version has the additional advantage of providing respondents with a more holistic representation of each state using multiple anchors for each alternative.</li> </ul>
Bias-free across demographic groups	• No data available	<ul style="list-style-type: none"> <li>• Additional research is required to assess potential bias.</li> </ul>
Sensitive to inter-individual differences	• Bundled or Branched	<ul style="list-style-type: none"> <li>• Long Version may be problematic due to scoring and categorization issues with the transactional pairs.</li> </ul>
Sensitive to intra-individual changes	• Bundled or Branched	<ul style="list-style-type: none"> <li>• Long Version may be problematic due to overall length.</li> </ul>

cific types of job tasks (e.g., tasks requiring different levels of creativity and perseverance).

Regarding sensitivity to capture intra-individual changes, we see the Long Version as less appropriate than the other two versions, given its overall length. As mentioned earlier, we have serious concerns that using an 18-item survey increases the possibility that the focal constructs (states) will have altered during the time needed to answer all items on each subscale. This may be especially true when items on a particular subscale are randomly dispersed throughout the instrument. For example, one could easily imagine a respondent starting the instrument in the paratelic state (e.g., curious as to what's being asked for its own sake) and reversing to the telic state if boredom sets in as they work their way through 18 items (i.e., reversal due to frustration). In this scenario, the Long Version may be temporally out of sync with the states being measured; it may not have the im-

mediacy to measure rapidly-changing states. In addition, a specific frame of mind may be induced by a particular item (test reactivity; Fischer & Corcoran, 2007). One might argue that longer state scales are more susceptible to reactivity than shorter ones, since longer scales may be more psychologically demanding.

Whether respondents change state as they move through the Long Version is a testable proposition. If survey length triggers reversals, then items later in the survey would tend to classify the respondent into a different state than earlier items, as long as item content may be ruled out as an alternative explanation (e.g., by randomly varying the order in which items on each subscale appear). One might also vary respondent reversibility by manipulating the experimental situation in order to determine if the problem is more pronounced in situations in which the likelihood of reversibility is high.

In summary, it is our view that the Bundled Version (RTSM-B) is the strongest version of the three and the most likely to prove useful in research on motivational states. It would be difficult to recommend using the Long Version given the scoring issues already discussed. However, researchers investigating only the somatic pairs may find three items per subscale an acceptable length and elect to use just those subscales of the Long Version. Either the Branched or the Bundled version may prove useful for different purposes under different conditions. In any case, our belief is that we have the foundation of a useful pair of instruments which may prove invaluable in future reversal theory research to assess the impact of moment-to-moment changes within individuals.

### References

- Apter, M. J. (1982). *The experience of motivation: A theory of psychological reversals*. London and New York: Academic Press.
- Apter, M. J. (Ed.). (2001a). *Motivational styles in everyday life: A guide to reversal theory*. Washington, D.C.: American Psychological Association.
- Apter, M. J. (2001b). Reversal theory as a set of propositions. In M. J. Apter (Ed.), *Motivational styles in everyday life: A guide to reversal theory* (pp. 37-51). Washington, D.C.: American Psychological Association.
- Apter, M. J. (2007). *Reversal Theory: The Dynamics of Motivation, Emotion and Personality* (2nd ed.). Oxford: Oneworld Publications.
- Apter, M. J. (2013). Developing Reversal Theory: Some suggestions for future research. *Journal of Motivation, Emotion, and Personality*, 1, 1-8.
- Apter, M. J., & Desselles, M. (2001). Reversal theory measures. In M. J. Apter (Ed.), *Motivational styles in everyday life: A guide to reversal theory* (pp. 55-76). Washington, D.C.: American Psychological Association.
- Apter, M. J., & Larsen, R. (1993). Sixty consecutive days : Telic and paratelic states in everyday life. In J. H. Kerr, S. Murgatroyd, & M. J. Apter (Eds.), *Advances in reversal theory* (pp. 107-122). Amsterdam: Swets & Zeitlinger.
- Apter, M. J., Mallows, R. & Williams, S. (1998). The development of the Motivational Style Profile. *Personality and Individual Differences*, 24, 7-18.
- Apter International. (1999a). *The Apter Motivational Style Profile*. Uppingham, England: Author.
- Apter International. (1999b). *Manual for the Apter Motivational Style Profile*. Uppingham, England: Author.
- Bergkvist, L., & Rossiter, J. R. (2007). The predictive validity of multiple-item versus single-item measures of the same constructs. *Journal of Marketing Research*, 44(2), 175-184.
- Bergkvist, L., & Rossiter, J. R. (2009). Tailor-made single-item measures of doubly concrete constructs. *International Journal of Advertising*, 28(4), 607-621.
- Burisch, M. (1984). You don't always get what you pay for: Measuring depression with short and simple versus long and sophisticated scales. *Journal of Research in Personality*, 18, 81-98.
- Burisch, M. (1997). Test length and validity revisited. *European Journal of Personality*, 11, 303-315.
- Calhoun, J. E. (1995). Construct validity of the Telic/Paratelic State Instrument: A measure of reversal theory constructs. Unpublished doctoral dissertation, University of Kansas School of Nursing, Lawrence.
- Cook, M. R., Gerkovich, M. M., Potocky, M., & O'Connell, K. A. (1993). Instruments for the assessment of reversal theory states. *Patient Education and Counseling*, 22, 99-106.
- Drolet, A. L., & Morrison, D. G. (2001). Do we really need multiple-item measures in service research? *Journal of Service Research*, 3(3), 196-204.
- Fischer, J., & Corcoran, K. (2007). *Measures for clinical practice and research: A sourcebook* (Vol. 1). New York, NY: Oxford University Press.
- Furr, R. M. (2011). *Scale construction and psychometrics for social and personality psychology*. London: SAGE Publications Ltd.
- Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6), 504-528.
- Heatherston, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology*, 60(6), 895.
- Kakolewski, K. E., Goings, K., O'Connell, K. A., Gerkovich, M. M., & Cook, M. R. (1996, August). A laboratory test of the Global Assessment Scale as a measure of telic/paratelic states. *Synergy: The Reversal Theory Newsletter*, 3-4.
- Kerr, J. H., & Apter, M. J. (1999). The State of Mind Indicator for Athletes. In J. H. Kerr (Ed.), *Experiencing sport: Reversal theory*. (pp. 239-244). Chichester, England: Wiley.
- Kerr, J. H., & Svebak, S. (1994). The acute effects of participation in sports on mood: The importance of level of "antagonistic physical interaction." *Personality and Individual Differences*, 16, 159-166.
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1), 81.
- Murphy, S. L., & Desselles, M. (2011, July). Liability and affect: An ecological momentary assessment approach. Paper presented at the Fifteenth International Reversal Theory Conference, Washington, DC.

- Nunnally, J. C., & Bernstein, I. R. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- O'Connor, P. R. (1992). Reversal theory and mother-child compatibility. Unpublished doctoral dissertation, University of Tasmania, Hobart.
- O'Connell, K. A., & Calhoun, J. E. (2001). The telic/paratelic state instrument (T/PSI): Validating a reversal theory measure. *Personality and Individual Differences, 30*, 193-204.
- O'Connell, K. A., & Cook, M. R. (2001). Smoking and smoking cessation. In M. J. Apter (Ed.), *Motivational styles in everyday life: A guide to reversal theory* (pp. 139-153). Washington, D.C.: American Psychological Association.
- O'Connell, K. A., & Brooks, E. (1997). Resisting urges and adopting new behaviors. In S. Svebak & M. J. Apter (Eds.), *Stress and health: A reversal theory perspective* (pp. 157-171). Washington, DC: Taylor & Francis.
- O'Connell, K. A., Gerkovich, M. M., Bott, M., Cook, M. R., & Shiffman, S. (2000). Playfulness, arousal-seeking, and rebelliousness during smoking cessation. *Personality and Individual Differences, 29*, 671-683.
- Popkess-Vawter, S. (1998). Reversal theory and overeating: A new paradigm to study weight control. *Western Journal of Nursing Research, 20*(1), 67-83.
- Popkess-Vawter, S., Gerkovich, M. M., & Wendel, S. (2000). Reliability and validity of the Overeating Tension Scale. *Journal of Nursing Measurement, 8*(2), 145-160.
- Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem scale. *Personality and Social Psychology Bulletin, 27*, 151-161.
- Robins, R. W., Trzesniewski, K. H., Tracy, J. L., Gosling, S. D., & Potter, J. (2002). Self-esteem across the lifespan. *Psychology and Aging, 17*, 423-434.
- Rossiter, J. R. (2002). The C-OAR-SE procedure for scale development in marketing. *International Journal of Research in Marketing, 19*(4), 305-335.
- Ruch, W., Köhler, G., & Van Thriel, C. (1996). Assessing the "humorous temperament": Construction of the facet and standard trait forms of the State-Trait-Cheerfulness-Inventory-STCI. *Humor, 9*, 303-340.
- Ruch, W., Köhler, G., & Van Thriel, C. (1997). To be in good or bad humor: Construction of the state form of the State-Trait-Cheerfulness-Inventory—STCI. *Personality and Individual Differences, 22*, 477-491.
- Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect grid: a single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology, 57*(3), 493.
- Siegel, S., & Castellan, N. J. J. (1988). *Nonparametric Statistics for the Behavioral Sciences* (2nd ed.). New York: McGraw-Hill.
- Spielberger, C. D. (1986). Preliminary manual for the State-Trait Personality Inventory (STPI). Select Press.
- Spielberger, C. D. (1989). State-trait anxiety inventory: a comprehensive bibliography. Palo Alto, CA: Consulting Psychologists Press.
- Spielberger, C. D., & Reheiser, E. C. (2004). Measuring anxiety, anger, depression, and curiosity as emotional states and personality traits with the STAI, STAXI, and STPI. *Comprehensive Handbook of Psychological Assessment, 2*, 70-86.
- Svebak, S. (1991). One state's agony, another state's delight: Perspectives in coping and musculoskeletal complaints. In C. D. Spielberger, I. D. Swanson, J. Strelau, & M. T. Brebner (Eds.), *Stress and anxiety* (Vol. 13, pp. 215-229). New York: Hemisphere.
- Svebak, S. (1993). The development of the Tension and Effort Stress Inventory (TESI). In J. H. Kerr, S. Murgatroyd, & M. J. Apter (Eds.), *Advances in reversal theory* (pp. 189-204). Amsterdam: Swets & Zeitlinger.
- Svebak, S., & Murgatroyd, S. (1985). Metamotivational dominance: A multimethod validation of reversal theory constructs. *Journal of Personality and Social Psychology, 48*, 107-116.
- Thorndike, R. L. (1982). *Applied psychometrics*. Boston: Houghton Mifflin.
- Timmermans, T., Van Mechelen, I., & Kuppens, P. (2010). The relationship between individual differences in intraindividual variability in core affect and interpersonal behaviour. *European Journal of Personality, 24*(8), 623-638.
- Vealey, R. S. (1986). Conceptualization of sport-confidence and competitive orientation: Preliminary investigation and instrument development. *Journal of Sport Psychology, 8*, 221-246.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063.
- Wendel, S. (1999). Reversal theory: Motivation for overeating in obese dieting individuals. Unpublished doctoral dissertation, University of Kansas School of Nursing, Lawrence.
- Young, T., Desselles, M., Lee, J., Apter, M. J. (2005, July). Measuring reversal theory states in real time: Developing the Reversal Theory State Inventory. Paper presented at the Twelfth International Reversal Theory Conference, Winnipeg, Canada.