

# Six Criteria of a Viable Theory: Putting Reversal Theory to the Test

Kenneth M. Cramer

Psychology Department, University of Windsor, Windsor, Ontario, Canada

In the wake of differential quality of theories, where some outshine others by way of, for instance, scope of explained phenomena or track record of empirical support, I review a series of criteria to judge the merit of multiple theories. The six criteria include comprehensiveness, precision and testability, parsimony, empirical validity, and both heuristic and applied value. I then utilize these criteria to review various leading personality theories, and include among them – for comparison – meta-motivational Reversal Theory, which challenges the universality of stable traits and instead embraces the notion of changeable states according to circumstances. Finally, I offer the reader a graphic representation (via cluster analysis) of the partnerships and families that unite together various theories based on similar criteria of evaluation.

*Keywords:* theory evaluation, formal criteria, personality, meta-motivation, cluster analysis

*Nothing is more dangerous than an idea  
when it's the only one we have.*

Emile-Auguste Chartier (1868-1951)  
French philosopher/journalist

With a plethora of explanations to account for the behavioral, motivational, and emotional phenomena of the human experience, the relative efficiency of theories becomes more evident. That is, as theories are not created equal, it is important to have available to us the tools – much like a Consumer's Reports Buyers Guide – to scrutinize theories and separate those that are superior from those beset with flaws. Psychology for example is populated with an abundance of theories of varying quality, and herein I explore the relative efficiency of a relatively complete set of personality theories (e.g., Freud and Maslow), but include among them a theory based not on stability, but on change. As a motivational theory, Reversal Theory (Apter, 2007) posits human behavior as a ballet among various changing states and goals, so initiated by the circumstances (satiation) and needs of the individual. Long ago, the Greek Philosopher Heraclitus (c. 535 - c.475 BCE) suggested that the entire world is in a state of flux; today, Reversal Theory challenges the static vision of the individual as governed by a relatively permanent collection of internal dispositions or traits (Lafreniere & Cramer, 2006).

The present paper is divided then into four key parts: first, I explore the six criteria (alone and in combination) to highlight both the individual relevance to theory building and the

relative weight or salience to judge the strength and veracity of a theory. Second, I will briefly review how these criteria fare for several prominent personality theories, including Reversal Theory. Finally, I offer some thoughts for future directions in an effort to build a stronger theory.

## Six Criteria of Evaluation

In no particular order, I present the six criteria of evaluation that may be used to judge the veracity and utility of a given scientific theory; upon reviewing each separately, I will consider the combination of various pairs of criteria under common nomenclature (Ryckman, 2013).

### Comprehensiveness

Comprehensive theories encompass a greater scope or range of explanation for various phenomena. Whereas some theories may be narrow in focus (explaining a limited number of observations, e.g., the facial feedback theory of emotion, Andréasson & Dimberg, 2008), other theories may cast a greater net over a wider range of phenomena (e.g., Evolutionary Theory). Freud's Psychodynamic Theory is likely the most comprehensive personality system ever devised (Hergenhahn, Olson, & Cramer, 2003), explaining love and war, marriage and religion, normal and abnormal developmental, humor, incest, slips of the tongue – the list is endless. Given that good theories should describe, explain, predict, and control phenomena and behavior, comprehensive theories aim to accomplish all these goals. Still, many seemingly viable theories may fall short in their scope. Despite being the dominant theory of personality today, Trait Theory is chiefly one of description rather than any of explanation (the hallmark of Freudian depth psychology), prediction, or control. By this, trait theorists may describe *that* an individual is extraverted, but not account for *why*. This feature may well highlight the

---

Correspondence concerning this article should be addressed to Kenneth M. Cramer, Psychology Department, University of Windsor, 401 Sunset Ave., Windsor, ON, N9B 3P4, Canada. E-mail: kramer@uwindsor.ca

danger of ad hoc explanations, wherein a theory appears to offer a suitable explanation only after the data are collected (with no prior predictions being made). For instance, Freud's Psychodynamic Theory (while heralded for its comprehensiveness) preferred an ad hoc interpretation into how an individual's childhood likely unfolded in order to yield the person's unique behavior protocol. Without a theory or without *a priori* hypotheses and predictions, patterns may be observed in any series of data, even if random. Mathematicians are keen to ask: what number is next in the series? Having already offered us:

1 2 3 5 8 . . .

our job is to provide justification for the number that suitably comes next. But given that various answers and explanations can be offered to the problem above, we are no closer to the solution. Is the answer 13 (as the sum of the previous two digits in the series,  $1+2=3$ ,  $2+3=5$ ,  $3+5=8$ ,  $5+8=13$ )? Or is the solution more complex? Supposing we take the digit sum + 2, so that  $1+2+(2)=5$ , and  $1+2+3+5+(2)=13$ . In this way, the solution "13" is arrived at by two different routes – so which is correct? The struggle over whose explanation of the same phenomenon is correct is widespread in science. For instance, two geological theories will attempt to account for the presence of mountain ranges by either catastrophe or painfully slow geological morphism (Stein & Wysession, 2009). Whereas each theory will surely offer an account for the staggering heights of Mount Everest (thus enhancing their comprehensiveness), but which is correct? Both? Neither? It will be the supporting data that determines our judgment; as such, comprehensiveness is a solely insufficient criterion by which to judge the feasibility of a theory; other criteria are needed.

### Precision and Testability

As the second and often most rigorous criterion, precision and testability demands that a good theory consist of constructs that are clearly defined, tightly interrelated, and readily open to reliable and valid measurement through falsifiable hypotheses (Popper, 1963). Plainly we ask: How do you define and identify and measure a psychological property like a motive or a need or even a reversal? What about an id, identity, peak experience, archetype, trait, condition of worth, etc.

Good theories should also expose themselves to rigorous hypothesis testing. Indeed, Meehl (1978) differentiated between point predictions and directional predictions. For example, Eysenck's personality theory of cortical arousal (1965) will predict (directionally) that extroverts will smoke more, drink more, take more Valium, enjoy more cups of coffee in the morning, require more classical conditioning trials to acquire a given skill, even salivate more to the presentation of lemon juice to the tongue – but will it predict how much more nicotine and alcohol and Valium and coffee

and conditioning trials and lemon juice will be necessary? No. "Harder" sciences like physics and chemistry are more typically held to the taller feats of point predictions, and this speaks to the rigor of precision. In these theories, components and constructs are neatly interrelated, tied conceptually to each other, so that velocity is the combination of speed and direction, acceleration is the combination of velocity and direction, and momentum is the combination of mass and velocity.

That a theory is empirically testable extends beyond its precision of interrelated concepts to the rigor of the instruments used to measure those concepts. For instance, many instruments today struggle to reliably measure many worthwhile personality traits, like self-actualization (Leclerc, Lefrançois, Dubé, Hébert, & Gaulin, 1999); this yields poor indices of internal consistency or reliability. Unreliable measures – those that render sometimes vastly disparate values upon repeat assessment – can never be valid, because one cannot be certain which value rendered is the correct one. Imagine an individual taking an IQ test three times within a matter of days with wildly different values. Reliability is the cornerstone of measurement; without it, we cannot confidently be certain of the properties in the world around us. We could add 10 pounds to a gymnasium scale without anyone's knowing and every weigh-in would render a comparable, reliable, but unhappy result – reliable, yes; but invalid, because it does not render the true weight.

Many theories appear on the surface at least to conquer this criterion but woefully do not. Evolution theory is largely unfalsifiable (Gould, 1978; Hergenhahn et al., 2003), resistant to competing evidence and alternative hypotheses (e.g., how can homosexuality be an adaptive feature of human behavior?). So too, the cornerstone of operant or instrumental conditioning is founded on Thorndike's Law of Effect (Ryckman, 2013), which states that behaviors that bring about positive outcomes are likely to be strengthened (i.e., if it feels good, do it). From this law, we derive the definition of the positive reinforcer, or any stimulus in the environment that when presented following a given behavior will increase the likelihood of that behavior; but how could you test this theory? If you offered a person a stick of black licorice (my favorite) after each production of a factory widget, the frequency of widget production should increase – but what if it does not? Was the theory disconfirmed? Or was black licorice not a positive reinforcer?

### Parsimony

English philosopher Sir William of Occam (Hergenhahn et al., 2003) explained that much like a butcher's cut of meat beset with excess fat, a trimmed cut would be more attractive to the shop patron (not so the butcher); so too a theory trimmed of excess concepts and needless explanation would likely manifest the correct explanation of the world. In short,

all things being equal, the simpler or more parsimonious theory is likely the correct one. Indeed, whereas there is no reason to believe this to be true, we do tend to prefer simple explanations to complex ones, still acknowledging that the simpler theory may be wrong. Did God create the universe? It would be a simple explanation to be sure if He did and all the cosmologists could go home early. . . Does the Earth revolve around the Sun or does it merely rest on the back of an elephant, which rests on the back of a turtle, and that turtle on another turtle, and so on and so on – with turtles all the way down. Simple explanations truly, but doubtfully tenable. In fact, complex theories are certainly abundant, as seen in physics' string theory (Smolin, 2006). Consider two electrons sent off at the speed of light in opposite directions, and through magnets we can influence the directional spin of one of these electrons; this manipulation will perfectly and oppositely predict the spin of the other electron – but it takes 12 dimensions (not the standard 3) to do it; hardly simplistic, but extraordinarily accurate.

Returning to psychology, learning theory takes notable exception to the term “personality” in general (Ryckman, 2013), since ultimately we are still talking about a defined collection of behaviors that embody the notion of (for instance) extraversion. If an extraverted individual talks often on the phone, visits frequently with friends, and strikes up with ease a conversation with a stranger, what use is the trait “extraversion” since it adds nothing to the description? In short, to a behaviorist, personality does not exist – trimmed free by Occam's Razor.

### Empirical Validity

I always hated sports fans who would change their jerseys and hats and franchise affiliations based on the final championship game, so these “fans” would forever align themselves with a new winning team (i.e., the team with the best win/loss record). Given my loyalties to the same team year after year (regardless of their win/loss record; and yes, there were some very bad years in there), I find the flip-flop habit of team replacement reprehensible, but this is exactly how science judges the empirical validity of a theory. Beyond the descriptive and explanatory scope of a theory (comprehensiveness), our confidence in a theory's viability is strengthened by its ability to correctly predict and control phenomena. We grow disenchanted with a theory that fails to predict and control behavior, brandishing a losing record – it may be time to change teams.

Empirical prowess may also reflect the extent to which a theory manages disconfirming evidence, since studies with negative results carry more weight than those with positive results. Suppose you reviewed 100 studies evaluating the existence of reversals in human behavior, and 85 of those studies confirmed the hypothesis, and 15 failed to find support. While 85 to 15 is hardly an even scorecard, those 15

studies are far more salient, because a theory is easier to disprove (to reveal as incorrect) rather than prove to be correct. For instance, we could purport to say that a human being needs two arms to live, and find example after example of a living breathing two-armed person. Although each case is warmly encouraging and supports the hypothesis, it takes the presentation of a single one-armed living person to fully dispel the theory. As such, disconfirming evidence carries more weight; we must demand that a theory offers us an explanation about any and all disconfirming evidence – and we must then scrutinize its treatment of those instances: does the theory get revised in light of the evidence, does it simply ignore those unfortunate studies, or does it create a just-so story that sounds plausible but is not supported by any evidence (Gould, 1978).

### Heuristic Value

A theory's heuristic value involves its ability to generate unique thoughts and perspectives and directions in other fields. For example, capitalism spurred a cost/benefit analysis among thinkers considering interpersonal relationships (viz. social exchange theory; Stanford, 2008), whereas the invention of the computer generated a new conceptualization of the brain. Freud's Psychodynamic Theory has been transformed (or sublimated) into many diverse fields, including education, art, literature, cinema, sexuality, music, architecture, history, etc. (Ryckman, 2013). But some theories remain dormant, instilling little inspiration in other fields (cf. Raymond Cattell and George Kelly).

### Applied Value

Finally, a theory's applied value can be measured by the extent to which it offers effective solutions to life's problems. Freud's free-association “talking cure” was used as a treatment for a new form of war neurosis called “shell-shock” (Rivers, 1918a, 1918b); the theory of planned action (Ajzen, 1991) has been applied to smoking cessation and drug rehabilitation, and cognitive dissonance theory has offered solutions to guard against wartime abuses like those at the Abu Ghraib prison (Zimbardo, 2007). Many theories often thrive on this one component, and generate a devout following (especially in professional circles) on the heels of the theory's success in relieving real-world problems. It remains difficult, in solitary consideration, to applaud the salience of a theory's applied value when we realize there could be many reasons why a particular solution may be successful in treating a problem.

### Criteria Pairing

It is worthwhile to note that many of these criteria can be seen in partnership. Comprehensiveness often couples with

Applied Value to determine a theory's utility. Another common pairing is seen with Comprehensiveness and Parsimony – a theory that accounts for a wide range of phenomena is often more complex, requiring more component parts. Consider the search for the most comprehensive number of cardinal personality traits to explain the full scope of behaviors – Cattell and colleagues (Cattell, Saunders, & Stice, 1950) originally said there were 16, Big Five proponents suggest that was triple the correct number (McCrae & Costa, 1990), Eysenck's (1991) PEN model suggested only three, and an interpersonal circumplex model of love and dominance offers only two (Wiggins, 2003). One final combination worth mentioning is Precision/Testability and Empirical Validity – often inseparable since concepts that are clearly defined and precisely measured will pave the way for more definitive evidence. As one final note, the six criteria will differ in their salience, and are so divided into three pairs: the least stringent criteria are Comprehensiveness and Applied Value. More stringent criteria are found in a theory's Parsimony and Heuristic Value. Finally, the ultimate test of a theory is precise and valid measurement coupled with supportive evidence.

### **Evaluating Personality Theories**

I will now outline several prominent personality theories according to their management of the six aforementioned theoretical criteria before doing the same for Reversal Theory. Readers may conduct a similar analysis for other categories of theories as an exercise in theory-building. The following six personality theories were selected either because (in selected cases) of their commonality, and (in other cases) their dissimilarity. They are the theories of Freud, Jung, Erikson, Eysenck, Rogers, and Maslow (see Hergenhahn et al. 2003; and Ryckman, 2013 for review).

#### **Freud's Psychodynamic Theory**

Freud's theory fares evenly among the six criteria, with strengths and weaknesses (common to theories of the era; Ryckman, 2013). The theory earns very high marks for comprehensiveness, with an outstanding range and diversity of explained phenomena. Ryckman writes that "Freud developed a theoretical system that explicitly sought to explain virtually all of human behavior. His theory remains, to the present day, the most comprehensive conceptual system ever created by a personality investigator" (p. 64). Especially high marks are also earned for both heuristic and applied value, influencing thinkers well outside personality psychology who readily made use of Freudian concepts. Conversely, Freud's theory fails with respect to precision and testability (through the use of fuzzy concepts, impressive though misinterpretable and ambiguous metaphors, and post-hoc interpretations). The theory fails at parsimony because it is overly simplistic, assuming sex and aggression as the master motive

behind all human behavior; it further held an especially pessimistic view of the human condition, holding humans to be irrationally driven amoral creatures. Finally, though many believe Freud's theory to be unsupported largely in the literature, given thousands of studies to date, more studies stand in favor of the theory; however, we should remember that disconfirming evidence carries more weight, and so Freud receives moderate to low marks for empirical validity.

#### **Jung's Analytic Theory**

Jung's theory (Ryckman, 2013) renders a similar scorecard to Freud's – especially comprehensive and rich in both heuristic and applied value. Similar to Freud too is the lack of parsimony (far too complex); imprecise, inconsistent, and vague constructs; and difficulty with empirical testing resulting from imprecision.

#### **Erikson's Psychosocial Theory**

As another Freudian-style evaluation, Erikson's theory earns high marks for comprehensiveness and both heuristic and applied value; but low marks for parsimony (i.e., limited explanatory mechanism, placing too much emphasis [often post-hoc] on identity formation and integration), and vague and poorly defined constructs that are difficult to test empirically.

#### **Eysenck's Biological Theory**

We wanted to introduce Eysenck's theory to illustrate that several personality theories do not all conform to the Freudian profile, and this is a good example. Eysenck receives very high marks for comprehensiveness but only moderate marks for parsimony, since, although reasonably efficient, the theory fails to account for personality functioning. Although more precise than typical Freudian and neo-Freudian constructs, Eysenck encountered some difficulty outlining key terms such as arousal and typological categories. Empirical support is strong but domain-specific, inviting considerably more work in the area of psychopathology and biology. Although the theory's heuristic value is not as strong as seen in other personality theories, limited chiefly to the United Kingdom; the application to real-world problems has been extensive.

#### **Rogers' Person-Centered Theory**

Although initially restricted in its scope, Rogers labored to explain a greater range of domains, so that to date the theory is more comprehensive than originally drafted. Constructs such as "genuineness," "empathy," and "unconditional positive regard" are difficult to define clearly and so the theory receives moderate ratings for precision/testability. Given that Rogers' theory appears wrought with feel-good terms, it appears to be overly simplistic in its assumption that people

are perpetual do-gooders; even his understanding toward the solution to real world historical and political problems takes a decidedly optimistic perspective, when in fact people can often act in truly selfish ways. As such, the theory is not parsimonious. Positive results have emerged from the empirical arena, and both the heuristic and applied values are impressive.

### **Maslow's Self-Actualization Theory**

Maslow receives moderate ratings for comprehensiveness, since – although he incorporated many Freudian concepts with greater clarity into his humanistic theory – the explanatory base remains limited in its scope and diversity; as such, the theory fails to be parsimonious, appearing too optimistic and dismissive of situational influences. Low marks are also earned for precision and testability, with many fuzzy concepts and a motivational pyramid beset by multiple exceptions (martyrs die for a greater cause, Michelangelo starved while painting the chapel ceiling; must we love ourselves before we love others).

## **Evaluating Reversal Theory**

### **Comprehensiveness and Applied Value**

A review of previous studies ([www.reversaltheory.org](http://www.reversaltheory.org)) and conference proceedings uncovers a broad range and diversity of relevant topics, and so Reversal Theory receives very high marks for comprehensiveness. As a metamotivational theory, it accounts for a host and wide diversity of phenomena. Indeed, it includes many features highlighted in other motivational theories (e.g., incentive, drive, arousal, instinct, etc.), so that Reversal Theory acts much like a multi-tool; while often preferred, multi-tooled theories are not necessarily valid, so I urge caution. Although one may grant high marks to the theory based on multiple applications in the fields of business, health, education, and counseling; I caution the reader to further probe why applications derived from Reversal Theory may be effective.

### **Parsimony and Heuristic Value**

Because the theory does not appear bogged down with an excess of constructs or conceptual framework, and because a sufficient number appear in place to explain the relevant phenomena, the theory appears to be parsimonious. The heuristic value however is not especially high, because it has yet to stimulate new directions or initiatives in other fields, though this may change over time.

### **Empirical Validity and Precision/Testability**

Although the empirical validity is compelling, it is difficult to evaluate the merit of this criterion since it is largely

based on the precision and testability of the component constructs. There are problems in the realm of precision and testability. First, the notion of a reversible state within the individual's daily life may be confused with other transitory properties, like mood, needs, attitudes, thoughts, attributions, etc. Given that reversals constitute a change in state to the opposing pole, the notion of state is left somewhat ambiguous as to what change is to be measured. In short, it remains elusive as to what state is precisely changing and what state is precisely measured, be it one or several or all. Secondly, various inconsistencies have been observed between the frequency of opposing states (e.g., telic vs. paratelic), wherein – though hypothesized to be negatively correlated – they are often positively correlated (Lafreniere & Cramer, 2006; Sit, Lindner, Apter, Michel, & Mallows, 2010).

Moreover, the measures available to assess both the existence of a state and subsequently any reversal of that state need more psychometric rigor, since reliable measures have yet to be universally available. Whereas it is acknowledged that retest reliability estimates of reversal instruments will be less useful given the transitory state of the properties being measured (since the state is changing), said instruments still need to be internally consistent (i.e., the derivative items all measure the same construct), with Cronbach alpha coefficients of 0.70 or greater (Anastasi, 1988; Nunnally, 1982; Nunnally & Bernstein, 1994). By this, the Motivational Style Profile (MSP; Apter, Mallows, & Williams, 1998) and measures like it (Apter & Desselles, 2001; Sit et al., 2010) struggle in many key dimensions (i.e., conformity vs. negativity) to render those necessary psychometric benchmarks. Without reliable measures, correlations to other hypothesized constructs remain suspect, and we cannot draw reliable conclusions with any confidence. Certain state measures have been developed in the areas of anxiety, loneliness, depression to name a few (see Spielberger, Gorsuch, & Lushene, 1970; Spielberger, Reheiser, Ritterband, Sydeman, & Unger, 1995) to be sure, and typically scale developers will transform an existing trait measure to reflect how the respondent feels “at this moment.” However, even with a modest change in wording, measures must still be internally consistent so we can be sure the constituent items are measuring the same construct.

Furthermore, an additional assumption belies all scientific measurement wherein we believe we are not altering or rendering the property by the mere measurement of it. This is not a problem strictly for state measures, but for all measures across the science. Physicists will explain the difficulty in measuring the speed and rotation of an electron without influencing its behavior through the bombardment of photons, arguably necessary for the measurements. So while the matter besets all of science, it is especially pronounced in the social sciences. To what extent then does a person's state become more goal-directed and telic after being asked (in a

questionnaire, for instance) about their serious-mindedness? It is difficult to be sure how to advise researchers regarding the solution to this problem. Finally, with solid and sound psychometric instruments available for widespread use, researchers may then make and test their hypotheses, which should necessarily grow from directional to point predictions as the theory evolves. Many studies to date have mapped Reversal Theory constructs into the everyday lives of people – be it in popular music or arctic exploration – but these studies, while compelling, offer little support as to the veracity of the theory, since many models of personality and motivation could well explain the same behavior. The time for more rigorous and theory-driven hypothesis testing is nigh.

### Combining the Evaluations through Cluster Analysis

In an effort to determine which theories (of variable quality, based on the six criteria of evaluation) will align with other comparable theories, a cluster analysis was conducted based on 16 personality theories and Reversal Theory. Each of the 17 theories was scored on a 5-point Likert scale according to the six criteria of evaluation, where 1 = very poor and 5 = very good. Unlike factor analysis, which groups similar variables based on participants' scores (e.g., identifying similar verbal intelligence subtests such as vocabulary and arithmetic), cluster analysis will group similar participants or subjects or entities based on variable scores. In other words, a stew recipe consisting of perhaps 25 unique ingredients can be grouped according to vegetables, spices, liquids, meats, etc. So too, members of a biological or botanical category or genus or family will share similar characteristics according to various measures: laying eggs, suckling their young, breathing air, angle of spinal entry to the skull, etc. It is through the method of cluster analysis that categorical specialists may identify a subtype of depression or a unique species of giraffe.

The 17 theoretical profiles were entered into a cluster analysis (Everitt, 1993; Gordon, 1987), which designates unique families or clusters of theorists by examining their profiles for similarities that maximize intercorrelations among cluster members. Inter-theory similarity can be quantified using squared Euclidian distances, graphically represented in a dendrogram (from the Greek root for "tree") with theories on the vertical y-axis and linkage distances on the horizontal x-axis (see Figure 1; dendrograms are often used by biologists to identify closest ancestors in a family tree, so that based on their morphological measures, dogs and dolphins will share a common ancestor). The sooner that an entity in the dendrogram joins with another entity, the greater the likelihood that those entities belong in the same family. More distant or dissimilar entities will join later in the dendrogram. The more similar the theories, the shorter the linkage distance.

Based on algorithms outlined by Ward (1963) and the visual inspection of the dendrogram, two main clusters

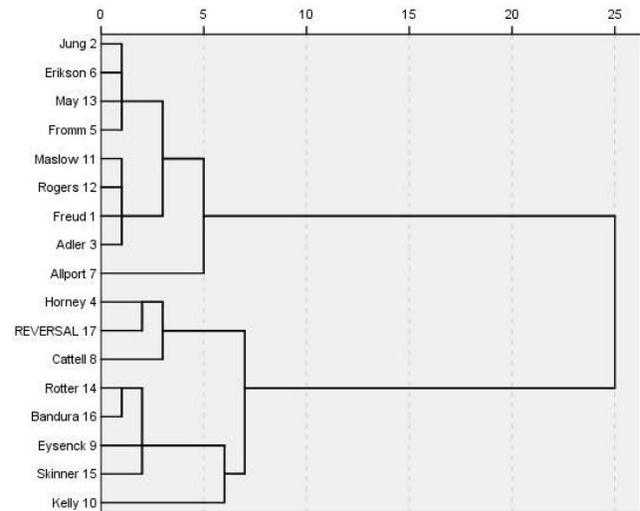


Figure 1. Cluster Analysis Dendrogram using Ward Linkages

emerged, each further divided into two subclusters. By interpretation, we can see two main families emerging among the 17 theories, noting again that it is a grouping according to quality rather than constituent features. The first main cluster consists of nine personality theories, subdivided into three subclusters. The first subtype consists of Jung, Erikson, May, and Fromm (high in comprehensiveness and both heuristic and applied value, but low in precision, parsimony, and empirical support); the second consists of Maslow, Rogers, Freud, and Adler (like subcluster-1 but with greater empirical support); the third consists of Allport only (like subcluster-1 but with fewer practical applications).

The second main cluster consists of seven personality theories plus Reversal Theory; these 8 theories also divide according to 3 subclusters. The first subcluster consists of Horney, Cattell, and Reversal Theory (higher comprehensiveness and parsimony, but lower precision and heuristic value); the second consists of Rotter, Bandura, Eysenck, and Skinner (relatively high in all categories); the third consists of Kelly only (low comprehensiveness and parsimony, but high precision and empirical support).

### Conclusions

With a set of efficient criteria by which to judge the quality of a scientific theory, both everyday consumers of the theory as well as those who authored it would benefit from the universal scorecard that would pinpoint strengths and weaknesses, paving the route to future work. That was the exercise pursued presently, juxtaposing a variety of formal personality theories alongside meta-motivational Reversal Theory. Although other branches of scientific theory could well have been selected for comparison, researchers and theoreticians

are encouraged to conduct this exercise for various theories in an effort to highlight needed comparisons and future directions.

This exercise yielded a profile for Reversal Theory that builders may find fruitful. The cluster analysis helps to highlight related theories that share a similar scorecard; indeed, Reversal Theory shares with both Horney and Cattell the features of a comprehensive and parsimonious model with low precision and heuristic value. Note as well that Reversal Theory joins later, as more distant cousins, with still other empirically-based theories of Rotter, Bandura, Eysenck and Skinner. The theory demonstrated high scores for comprehensiveness with an impressive scope and diversity of phenomena explained. In addition, the theory appears to be parsimonious, with neither an abundance nor paucity of component parts and theoretical concepts. Although the heuristic value or “beyond one’s own yard” influence was limited to date, the applied value remained impressive, with applications and evaluated solutions to many current problems. However, the areas in need of more work, in order to shore up the confidence in the empirical findings, reside chiefly in both the theory’s precision and testability. The issue of reliability remains troubling, with five key areas that need to be addressed before consumers offer greater confidence in the usefulness of Reversal Theory: (a) exclusive measurement of reversals and no other transitory state, (b) addressing inconsistencies in theoretically opposing processes that are empirically positively correlated, (c) adequate internal consistencies of Reversal Theory measures, (d) avoiding state change through state measurement, and (e) the advocacy for point vs. directional predictions. Until these matters are individually addressed, the theory remains in doubt among scientific scrutinizers and watchdogs.

### References

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179–211.
- Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.
- Andréasson, P., & Dimberg, U. (2008). Emotional empathy and facial feedback. *Journal of Nonverbal Behavior*, 32, 215–224.
- Apter, M. J. (2007). *Reversal Theory: The dynamics of motivation, emotion, and personality* (2nd ed.). Oxford, England: Oneworld Publications.
- 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, DC: American Psychological Association.
- Apter, M. J., Mallows, R., & Williams, S. (1998). The development of the Motivational Style Profile. *Personality and Individual Differences*, 24, 7–18.
- Cattell, R. B., Saunders, D. R., & Stice, G. F. (1950). *The 16 Personality Factor Questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.
- Everitt, B. S. (1993). *Cluster analysis* (3rd ed.). New York: Wiley.
- Eysenck, H. J. (1965). Extraversion and the acquisition of eyeblink and GSR conditioned responses. *Psychological Bulletin*, 63, 258–279.
- Eysenck, H. J. (1991). Dimensions of personality: 16, 5, or 3? – Criteria for a taxonomic paradigm. *Personality and Individual Differences*, 12, 773–790.
- Gordon, A. D. (1987). A review of hierarchical classification. *Journal of the Royal Statistical Society*, 150, 119–137.
- Gould, S. J. (1978). Sociobiology: The art of story telling. *New Scientist*, 80, 530–533.
- Hergenhahn, B. R., Olson, M. H., & Cramer, K. M. (2003). *An introduction to theories of personality*. Toronto: Prentice Hall.
- Kerr, J. H. (2001). *Counselling athletes: Applying Reversal Theory*. London: Routledge.
- Lafreniere, K. D., & Cramer, K. M. (2006). Examining Reversal Theory measures in relation to NEO personality dimensions and consideration of future consequences. *Personality and Individual Differences*, 40, 1387–1397.
- Leclerc, G., Lefrançois, R., Dubé, M., Hébert, R., & Gaulin, P. (1999). Criterion validity of a new measure of self-actualization. *Psychological Reports*, 85, 1167–1176.
- McCrae, R. R., & Costa, P. T. Jr. (1990). *Personality in adulthood*. New York: The Guildford Press.
- Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, 46, 806–834.
- Nunnally, J. C. (1982). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Popper, K. (1963). *Conjectures and refutations*. New York: Basic.
- Rivers, W. H. R. (1918a). An address on the repression of war experience. *Lancet*, 1, 173–177.
- Rivers, W. H. R. (1918b). War neurosis and military training. *Military Hygiene*, 2, 513–533.
- Ryckman, R. M. (2013). *Theories of personality* (10th ed.). Toronto: Nelson.
- Sit, C. H. P., Lindner, K. J., Apter, M. J., Michel, G., & Mallows, R. (2010). The development of the Motivational Style Profile for Children (MSP-C). *Current Psychology*, 29, 71–87.
- Smolin, L. (2006). *The trouble with physics: The rise of String Theory, the fall of a science, and what comes next*. New York: Houghton Mifflin.

- Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Spielberger, C. D., Reheiser, E.C., Ritterband, L.M., Sydeman, S.J., and Unger, K.K. (1995). Assessment of emotional states and personality traits: Measuring psychological vital signs. In Butcher, J.N. (Ed.) *Clinical Personality Assessment: Practical Approaches*. New York: Oxford University Press.
- (not cited) Standford, L. (2008). Social Exchange theories. In L.A. Baxter & D.O. Braithwaite (Eds.), *Engaging theories in interpersonal communication: Multiple perspectives* (pp.377-389). New York: Thousand Oaks.
- Stein, S., & Wysession, M. (2009). *An introduction to seismology, earthquakes, and earth structure*. Chichester: John Wiley & Sons).
- Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58, 236-244.
- Wiggins, J. S. (2003). *Paradigms of personality assessment*. New York, NY: Guilford Press.
- Zimbardo, P. (2007). *The Lucifer Effect: Understanding how good people turn evil*. New York: Random House.