Article

Team Role Experience and Orientation: A Measure and Tests of Construct Validity

Group & Organization Management
2015, Vol. 40(1) 6–34

DOI: 10.1177/1059601114562000
gom.sagepub.com
SAGE

John E. Mathieu[1], Scott I. Tannenbaum[2], Michael R. Kukenberger[3], Jamie S. Donsbach[2], and George M. Alliger[2]

Abstract

We review and synthesize previous team research and suggest that individuals' previous experiences and orientations combine to yield predispositions to occupy six different team roles, which we refer to as Team Role Experience and Orientation (TREO) dimensions. We report the development of a survey measure of TREO dimensions and establish its content validity using a sample of subject matter experts' item classifications. Furthermore, we provide evidence that TREO dimensions are distinguishable from, but related to, measures of the "Big 5" personality constructs. We also illustrate the temporal stability of the measures. Moreover, we test the predictive validity of TREO scores as related to peer ratings of members' behaviors during team activities. We discuss future theoretical and research implications of TREO dimensions, and potential future applications of the measure.

Keywords

group or team composition or diversity, personality, roles or role behavior, teams or teamwork

[1]University of Connecticut, Storrs, USA
[2]The Group for Organizational Effectiveness, Albany, NY, USA
[3]University of New Hampshire, Durham, USA

Corresponding Author:
John E. Mathieu, School of Business, University of Connecticut, 2100 Hillside Road, Unit 1041MG, Storrs, CT 06269-1041, USA.
Email: Jmathieu@business.uconn.edu

Teams are widely recognized as the basic building blocks of most modern-day organizations (Ilgen, Hollenbeck, Johnson, & Jundt, 2005; Kozlowski & Ilgen, 2006; Mathieu, Maynard, Rapp, & Gilson, 2008). Team-based designs enable organizations to quickly align their human resources with the multitude of changing work demands and competitive pressures. Enhancing team effectiveness offers a powerful means by which organizations can gain and maintain competitive advantage. Team effectiveness can be driven by a number of factors such as a supportive organizational environment, team-oriented external leadership, design features, dynamic processes and emergent states, and a host of other variables (Mathieu et al., 2008). However, research and practice have suggested that the best teams are well designed up-front. Teams that have an optimal mix of members' knowledge, skills, abilities, and other characteristics (KSAOs) are better positioned to work well together and to perform effectively than are teams composed of a less-optimal combination of members (Bell, 2007; Ilgen, 1999). In short, team composition serves as the foundation upon which other team factors are built, and represents a key enabling feature of teams. Moreover, understanding how a team is "composed" can provide insights for targeted team development activities.

Numerous characteristics have been used to index team composition, including personality, functional expertise, competencies, goal orientations, teamwork orientations, and a host of other attributes (Klimoski & Zukin, 1999; Mathieu, Tannenbaum, Donsbach, & Alliger, 2014). Importantly, these individual attributes motivate and enable individuals to occupy different team roles (Stewart, Fulmer, & Barrick, 2005). A role is generally defined as a cluster of related and goal-directed behaviors taken on by a person within a specific situation (Stewart, Manz, & Sims, 1999). Teams rely on different members to fulfill different critical needs such as organizing work, maintaining group harmony, and aligning their efforts with those of others in an organization (Aritzeta, Swailes, & Senior, 2007; Stewart et al., 2005). Accordingly, both research and practice will benefit from a greater understanding of individual differences that are associated with team role fulfillment, and from tools to assess those differences.

Although the existing literature has advanced numerous team role taxonomies, we believe that there have been a number of major barriers to a productive discourse on the topic. Team roles are largely considered a critical part of effective teaming (Ancona & Caldwell, 1988; Belbin, 1993) and viewed as a basic feature of work teams (Sundstrom, De Meuse, & Futrell, 1990). Yet, the topic has received relatively little empirical attention beyond the introduction of team role classifications. Although a number of taxonomies of team roles exist in the literature (Bales, 1950; Belbin, 1981, 1993; Benne & Sheats, 1948; Parker, 1994; Humphrey, Morgeson, & Mannor, 2009), work in this area is

stifled due to a lack of integration and the absence of publicly available and construct valid measures of role propensities. Accordingly, we have several goals for this article. First, we briefly review and synthesize previous work that has focused on team roles. From this, we distill six different roles that have potential widespread applicability. Second, we argue that individuals' previous experiences, on one hand, and personal orientations and predilections, on the other hand, combine to yield predispositions to occupy these different types of team roles. We suggest that this synthesis of preferences and experiences represents individual differences that are predictive of individuals' future role-related behaviors in teams. We refer to these individual differences as Team Role Experience and Orientation (TREO) dimensions. Third, we describe the potential relations between TREO dimensions and other variables in a nomological network, and suggest that TREO dimensions are distinguishable from, but related to, measures of traditional "Big 5" personality dimensions.

Fourth, we describe the development and validation of a survey measure of TREO dimensions. This effort includes qualitative grounding and content validity work by subject matter experts (SMEs) to generate and classify survey items. It also involves gathering data from two development samples, evaluating and refining the construct validity of the items, and then administering the revised measures to two validation samples. Fifth, we evaluate the discriminant validity of the TREO measures from measures of Big 5 personality dimensions, and test their pattern of relations using meta-analyses. Sixth, we demonstrate the stability of TREO scores over time using a fifth sample. And finally, we test the predictive validity of TREO measures as related to peers' ratings of role-related behaviors exhibited over the course of a 2-month long team business simulation exercise. We conclude with a discussion of our findings in terms of the nomological network associated with team's role-related behaviors and potential applications of TREO measures in the future.

In sum, the primary contributions of this work are threefold. We synthesize the diverse prior team role taxonomies and offer a six-dimensional framework designed to balance comprehensiveness with parsimony. Second, we develop measures of those six dimensions and provide evidence of their psychometric properties, construct validity, and predictive validity. The items are presented in the appendix and are freely available for research purposes. Finally, we outline an agenda for future research incorporating the six-dimensional role taxonomy.

Team Role Theories

Roles are often considered to be one of the fundamental and defining features of both organizations (Ilgen & Hollenbeck, 1991; Sluss, van Dick, &

Thompson, 2011) and teams (Hackman, 1990). Aritzeta et al. (2007) noted that there are two heritages in the team role literature. One approach, which we could term *role as position*, equates roles with expected behavior associated with the particular position that a team member occupies (e.g., Katz & Kahn, 1978). Essentially, this view *focuses on the characteristics and demands of jobs* and how they give rise to certain expected role behaviors of occupants. A second approach, "role as person," suggests that roles can be defined as a *combination of the values, attitudes, and behaviors of individuals who occupy particular locations in a social network*. From this perspective, roles emerge from a combination of members' natural inclinations or preferences, as well as the social-psychological dynamics of the group (Ilgen & Hollenbeck, 1991). We adopt this latter approach because we are interested in developing indices of individual differences that may predispose people to fulfill particular roles in teams.

Early research attempted to outline typologies of roles that team members take when interacting with others (e.g., Bales, 1950; Benne & Sheats, 1948). Benne and Sheats (1948) examined small discussion groups that were engaged in problem-solving activities. They observed the appearance of two types of roles: *Task* and *Maintenance*. Behaviors such as facilitating and coordinating group activities, suggesting new ideas and ways of solving problems fell under the "task role." In contrast, behaviors that encouraged praising, agreeing, and accepting the contribution of others within the group were considered part of the "maintenance role." Bales (1950, 1970) built on this team role research by analyzing the interactions between members of small groups and categorized their behaviors into *task-oriented and socio-emotional* categories. Task roles are held by those group members who engage in behaviors that are designed to facilitate task completion. Task activities involve behaviors such as delegation and the coordination of group communication. Socio-emotional leaders are those group members who engage in behaviors that are designed to facilitate positive relations within the group. Socio-emotional activities involve behaviors such as encouragement and compromising. Interestingly, Bales and colleagues argued that task and relationship roles emerge as a result of one another in that when task leaders work toward the completion of goals, they engage in behaviors that do not promote positive group relations, such as criticizing and giving orders. As such, most successful groups would have some members specializing in task activities and other members specializing in socio-emotional activities to balance the group dynamics. These early studies centered on the individual's behaviors within a group and the classification of these behaviors into broader roles. Benne and Sheats (1948) and Bales (1950) provided insight into two critical functions that roles are instrumental in performing: task execution

and the encouragement of strong interpersonal relations. Another primary contribution made by these researchers is that their work demonstrated the utility of using role constructs to cluster team members' behaviors.

The interest in team roles gained momentum in the 1980s with the publication of Belbin's (1981) work on successful management teams. Belbin's (1981) theory advanced eight distinct team role types: (a) idea generator, (b) resource investigator, (c) chairman, (d) shaper, (e) monitor evaluator, (f) team worker, (g) company worker, and (h) completer–finisher. In later editions, he changed various names (i.e., chairman to coordinator, company worker to implementer) and introduced a new role called specialist. Belbin (1981, 1993) examined management teams playing executive simulations (e.g., computerized management and business exercises) during training courses where team performance was measured in terms of winning or losing.

Belbin's team role model has been associated with both team behaviors and performance (for a recent review, see Aritzeta et al., 2007). Belbin and others have argued that successful teams that have all nine roles present, although individuals may display multiple roles, will be more balanced, which will result in higher levels of success. Although a number of studies have reported nonsignificant results (Anderson & Sleap, 2004; Jackson, 2002; Partington & Harris, 1999; Sommerville & Dalziel, 1998; Rushmer, 1996), other research has yielded positive relationships between such balance and management styles (Lessem & Baruch, 2000), team performance (Aritzeta & Ayestaran, 2003; Senior, 1998), cognitive styles (Aritzeta, Senior, & Swailes, 2005), and to the exercise of power and control (Fisher, Hunter, & Macrosson, 2001). Specifically related to team performance, in a small sample ($n = 11$), Senior (1997) provides some support for the link Belbin makes between team role balance and team performance. Aritzeta and Ayestaran (2003) demonstrate positive evidence for the team role balance argument as the majority of their mostly female work teams were balanced and showed high performance. However, using the same criteria, Park and Bang (2002) found that less than 5% of their largely male work teams were balanced, and did not find evidence to support that team role balance results in higher performance. Gender differences between both studies could explain differences observed with respect to team role balance and some of the inconsistent results across Belbin's team role studies.

In addition to the possible gender basis, the lack of consistent results across studies may be related to the type of team being assessed and the original sample and context. Belbin's work focused extensively on management teams, and while certainly of interest, others have argued that top management teams are qualitatively different from other types of teams (Hollenbeck, Beersma, & Schouten, 2012). Hambrick (1994) went so far as to refer to them

as top management *groups*. Suffice it to say that a taxonomy of management team roles may have limited applicability to other types of teams.

Furthermore, the psychometric properties of the scale may contribute to the inconsistent results and, importantly, question the validity of the scale. The team role model has predominately been measured through the Team Role Self Perception Inventory (TRSPI-8R; TRSPI-9R), in some cases alongside an Observer Assessment Sheet (OAS) where teammates could rate each other. As originally conceived and implemented, Belbin's inventory utilizes ipsative responses, which have demonstrated weak internal consistencies (Furnham, Steele, & Pendleton, 1993) and yield a contaminated covariance matrix, undermining efforts to evaluate its factor structure (Cheung & Chan, 2002). Given the popularity of the Belbin role theory, many studies have attempted to validate the scale (cf. Aritzeta et al., 2007; Broucek & Randell, 1996; Dierendonck & Groen, 2011; Fisher et al., 2001; Furnham et al., 1993; Senior, 1998; Smith, Polglase, & Parry, 2012). In their review of the research examining Belbin's scale, Aritzeta et al. (2007) notes the following:

> Having reviewed psychometric studies it is clear that neither the eight role nor the nine role version of the TRSPI has unequivocal psychometric support and most of the studies show low or at best average effect sizes indicating only partial psychometric support. (p. 110)

This lack of psychometric support is rather surprising given the widespread use and popularity of the Belbin scale.

Margerison and McCann (1985) developed a model containing eight roles located on four dimensions, which the researchers referred to as (a) relationships, (b) information, (c) decision making, and (d) organization (for a detailed review of Margerison and McCann's work, see Sadler-Smith, 2001). McCann and Margerison (1989) also developed a team role measure that has eight types: (a) explorer–promoter, (b) assessor–developer, (c) thruster–organizer, (d) concluder–producer, (e) checker–inspector, (f) upholder–maintainer, (g) reporter–advisor, and (h) creator–innovator. While recognizing team behavior as being related to the task and the sociological factors (including the roles that people play, and the norms and values operating), Margerison and McCann argue that individual preferences that people bring to teamwork are key psychological factors influencing individual and team performance. Like Belbin, they suggest that a team should have a balance of required behaviors and role preferences. Although this measure has also received practitioner attention, there seems not to be rigorous empirical evidence to support the balance hypotheses related to these roles. This measure is heavily influenced by the Jungian theories underlying the Myers-Briggs test (Myers,

McCaulley, & Most, 1985). It not only has norms but also has evidence of internal reliability and concurrent validity; however, there appears to be little or no evidence of the factorial structure of the measure (to confirm the classification or taxonomic scheme), or any evidence of the predictive or construct validity of the test.

Parker (1994, 1996) developed a set of team player styles, which fall into four types: (a) contributor, (b) collaborator, (c) communicator, and (d) challenger. In contrast to Belbin's emphasis on preferred team roles, Parker focused on personal style—a function or process may be performed in a variety of ways depending on the team members' personal style. In her review of team roles, Sadler-Smith (2000) note that according to Parker, an individual who approaches a function such as planning with a Contributor style, may adopt a tactical, statistical, specific, measurable, and conservative approach, whereas an individual with a Collaborator style facing the same process would use a strategic, visionary, open, and involving approach. While Margerison and McCann's and Belbin's team roles are based on Jungian tradition (Jung, 1923), Parker argued that his style-based team roles provide more latitude for different members taking different team roles and feedback for improvement. There is little research exploring the applicability or validity of Parker's work, with the exception of Kirnan and Woodruff's (1994) study that found evidence of construct reliability and validity of peer to self-ratings.

Barry (1991) classified leadership behavior approaches within self-managed teams. He described the following four types of leadership roles: (a) envisioning leadership (innovative and vision fostering), (b) organizing leadership (giving orders on missions), (c) spanning leadership (facilitating the activities that connect the team to the organization), and (d) social leadership (developing and maintaining the psychology and sociability of the team). DuBrin (1995) advanced 10 team roles: (a) knowledge contributor, (b) process observer, (c) collaborator, (d) people supporter, (e) challenger, (f) listener, (g) summarizer, (h) conciliator, (i) mediator, and (j) gate keeper. No validation evidence has been provided in support of Barry's or DuBrin's largely practitioner-based leadership roles.

Mumford, Van Iddekinge, Morgeson, and Campion (2006) attempted to address the fragmentation in the team role literature by sorting existing roles into categories they suggested captured the essence of each role. Their process fit 120 roles into 10 unique team knowledge roles that they in turn clustered into three broader categories of task, social, and boundary-spanning roles. *Task roles* share the common function of carrying out the work that constitutes the team's objective and included (a) contractor, (b) creator, (c) contributor, (d) completer, and (e) critic. *Social roles* involve maintaining the

social environment in which teams function and included sub-dimensions of (a) communicator, (b) cooperator, and (c) calibrator. The communicator role, for example, encompasses behaviors that create a social environment that is positive, open, and conducive to collaboration. One of the key theoretical contributions of Mumford et al. (2006) is that they included Ancona's (Ancona & Caldwell, 1988, 1992) work by integrating roles that interact both internally and externally to the team into their typology. *Boundary-spanning roles* address members' effort to align their team's actions with those of individuals outside of the team and include a counselor and a coordinator subrole. T. V. Mumford, Van Iddekinge, Morgeson, and Campion (2008) then developed situational judgment measures of members' team role *knowledge* and provided evidence of their validity. However, it is not clear the extent to which having knowledge of certain team roles relates to whether individuals will occupy or perform such roles in practice. In addition, Mumford created role-category level variables by collapsing roles into three categories: task, social, and boundary-spanning roles, which raises questions regarding the validity of their 10 specific roles.

TREO Classification and Dimensions

We believe that there is a long and valuable history exploring team roles. Yet, in a recent review of the literature, Aritzeta and colleagues (2007) questioned the existence of the team roles: "…we raise the question about the real existence of nine well-differentiated team roles and whether these team roles, in fact, are better differentiated using some other grouping suggested in the literature" (p. 105). Accordingly, in Table 1, we list and compare the types of roles that have been discussed by the different authors noted above and offer a synthesis in terms of six unifying themes or orientations. We submit that an individual's role propensities or predispositions are likely to be a function of their previous experiences and orientations. This view rests on the assumption that past behavior is a good indicator of future behavior (cf. Brown, 1978; Mumford & Owens, 1984). Work history measures also tend to yield consistent factor structures over time and across samples (cf. Davis, 1984; Neiner & Owens, 1982). By "orientations," we mean individual, natural inclinations that may well reflect personality or other individual differences (Stewart et al., 2005). All else being equal, we would anticipate these orientations or preferences to guide individuals' behaviors. However, we recognize that different people may have different opportunities to occupy various team roles as a function of the positions to which they have been assigned and of the mix of other team members' and their respective inclinations. Thus, we believe that individuals' predispositions to occupy various roles in future

Table 1. Previous Literature Integrated Within the TREO.

	Roles studied within previous team role research								
TREO roles	Benne and Sheats (1948)	Bales (1950)	Belbin (1981, 1993)	Ancona and Caldwell (1988)	McCann and Margerison (1989)	Barry (1991)	DuBrin (1995)	Parker (1994)	Mumford et al. (2006)
Organizer	Coordinator; group observer; gate keeper	Gives and asks for orientation opinion, suggestion	Coordinator, shaper		Assessor thruster	Organizing	Collaborator, summarizer	Collaborator	Contractor, calibrator
Doer	Procedural, technician, recorder		Completer–finisher, specialist, implementer		Concluder, controller, reporter			Contributor	Completer, creator, contributor
Challenger	Evaluator-critic	Disagrees	Monitor, evaluator				Challenger	Challenger	Critic
Innovator			Plant/idea generator		Creator	Envisioning	Knowledge contributor	Innovator	
Team Builder	Encourager, follower, compromiser, harmonizer	Shows solidarity, tension release, agrees	Team/company worker			Upholder, social	People supporter, listener		Communicator, cooperator
Connector			Resource investigator	Task coordinator, scout, ambassador, guard	Explorer	Spanning			Collector, consult

Note. TREO = Team Role Experience and Orientation.

teams will be a joint function of what they have done in past teams (past experience), as well as roles that they would like to occupy (behavioral predispositions). Taking both an orientation and experience approach allows us to consider both work history and natural inclinations providing a more holistic method to measuring individual team roles. We believe that the result will increase the likelihood of clearly validated team roles.

Table 1 lists the roles identified by previous theorists clustered into six TREO categories. As illustrated, the various sources list up to 10 specific types of roles, yet empirical evidence has yet to demonstrate their discriminant validity (see Aritzeta et al., 2007, for a detailed review). In short, we doubt whether there are 10+ empirically distinguishable team roles that members may occupy. Alternatively, many approaches have reduced their role profiles to two or three clusters (e.g., task vs. maintenance) that may obscure important differences. We endeavored to strike a balance between comprehensiveness and inclusion versus discriminant validity and parsimony, and propose a six-dimensional framework. Admittedly, this synthesis represents our beliefs about the number and nature of important team roles, based on our collective experience working with hundreds of different teams across a wide variety of settings. Yet, this framework is also grounded in the extant literature as detailed below. Table 2 lists the TREO dimensions and their associated definitions. Notably, the TREO dimensions are designed to represent individuals' general propensities to fulfill different team roles independent of any particular team context. In other words, we envision these dimensions as reflecting individual differences in propensities, not situated preferences or self-reports of behavior in a given context.

We distilled two team roles that are primarily directed at getting task work accomplished. First, we define "*Organizer*" as someone who acts to structure what the team is doing. An Organizer also keeps track of accomplishments and how the team is progressing relative to team and individual goals and timelines. This dimension is prominent in Benne and Sheats' (1948) coordinator and gate keeper dimensions, Barry's (1991) organizer, and DuBrin's (1995) and Parker's (1994) collaborator facet. Team success rests not only on the extent to which members can process information and coordinate their actions but also on the sheer fact that work gets accomplished. In other words, getting work done is a prime role feature in teams. This may be manifested differently in different teams working on different tasks, but the common element is that some members need to devote themselves to getting task work done. Accordingly, we define the second task work oriented dimension as "*Doer*" as someone who willingly takes on work and gets things done. A Doer can be counted on to complete work, meet deadlines, and take on tasks to ensure the team's success. This theme was evident in Benne and Sheats's

Table 2. Team Role Definitions.

Role	Definition
Organizer	Someone who acts to structure what the team is doing. An Organizer also keeps track of accomplishments and how the team is progressing relative to goals and timelines.
Doer	Someone who willingly takes on work and gets things done. A "Doer" can be counted on to complete work, meet deadlines, and take on tasks to ensure the team's success.
Challenger	Someone who will push the team to explore all aspects of a situation and to consider alternative assumptions, explanations, and solutions. A Challenger often asks "why" and is comfortable debating and critiquing.
Innovator	Someone who regularly generates new and creative ideas, strategies, and approaches for how the team can handle various situations and challenges. An Innovator often offers original and imaginative suggestions.
Team Builder	Someone who helps establish norms, supports decisions, and maintains a positive work atmosphere within the team. A Team Builder calms members when they are stressed, and motivates them when they are down.
Connector	Someone who helps bridge and connect the team with people, groups, or other stakeholders outside of the team. Connectors ensure good working relationships between the team and "outsiders," whereas Team Builders work to ensure good relationship within the team.

(1948) cluster of procedural, technician, and recorder dimensions; McCann and Margerison's (1989) controller, concluder, and reporter bundle; and Mumford et al.'s (2006) and Belbin's (1993) completer–finisher themes.

Based on our review, we also distilled two socio-emotional types of roles. Effectively managing members' interpersonal processes is critical to team success (Marks et al., 2001), and teams are social mechanisms for integrating the expertise and views held by diverse individuals (Larson & LaFasto, 1989). Team members often need to socially integrate opposing views and complex expertise. Accordingly, we define "*Team Builders*" as people who help to establish norms, support decisions, and maintain a positive work atmosphere within the team. This is a common theme in previous taxonomies and can been seen in Benne and Sheats's (1948) encourager, follower, compromiser, and harmonizer set; Bales's (1950) show solidarity, tension release, and agree dimensions; and DuBrin's (1995) people supporter and listener bracket. Whereas Team Builders focus their attention within the team, in

contrast, it can also be important to have one or more members coordinating team actions with external constituencies. We define a "*Connector*" as someone who helps bridge and connect the team with people, groups, or other stakeholders outside of the team. Connectors ensure good working relationships between the team and outsiders. This external linking role was the focal feature of Ancona and Caldwell's (1988) work, but can also be seen in Belbin's (1993) resource investigator, McCann & Margerison's (1989) explorer, Barry's (1991) spanning, and Mumford et al.'s (2006) collector and consult dimensions.

Finally, we also identified two destabilizing or change-oriented team roles. Various authors have described a challenger (e.g., DuBrin, 1995; Parker, 1994), critic (Benne & Sheats, 1948; Mumford et al., 2006), or disagrees (Bales, 1950) team role. This type of behavior offers value by mixing things up and getting members to consider alternative conceptions or work modes and helps them to prevent premature closure on decisions. Accordingly, we define a "*Challenger*" as someone who will push the team to explore all aspects of a situation and to consider alternative assumptions, explanations, and solutions. A Challenger often asks "why" and is comfortable debating and critiquing ideas. Whereas Challengers question what is being done, we define "*Innovators*" as people who regularly generate new and creative ideas, strategies, and approaches for how the team can handle various situations and challenges. An Innovator often offers original and imaginative suggestions. This theme was articulated by Belbin's (1993) plant/idea generator, McCann and Margerison's (1989) creator, Barry's (1991) envisioning, DuBrin's (1995) knowledge contributor, and Parker's (1994) innovator dimensions.

Psychometric and Nomological Network Analyses

The online supplement to this article provides extensive details concerning the development of Likert-type scales for measuring these six TREO dimensions. Table 3 summarizes the samples that we used and the purpose(s) for each. We first drafted preliminary items for each of the six TREO dimensions. Within each role, we generated two sub-sets of items: role orientations and specific behavioral experiences. We then had those items classified into the six substantive areas by 12 SMEs. From this classification, we identified 48 items, 4 orientations, and 4 behavioral experiences for each dimension that were consistently classified into their respective TREO categories. We then administered the 48 TREO items, along with measures of the Big 5 personality dimensions, to 317 Military officers and 266 upper-level undergraduate business students. Analyses using those data, with a few exceptions, confirmed the internal consistencies of the different scales, their convergent

Table 3. Summary of Study Samples and Their Use.

Sample population	Primary purpose(s)
1. Academic (n = 6) and practitioner (n = 6) subject matter experts	Classify 105 draft survey items into six TREO dimensions
2. Military officers (n = 317) 3. Large university capstone business students (n = 266)	Evaluate TREO scale reliabilities and factor structure, and their discriminant validity from measures of Big 5 personality constructs. Revise items.
4. Military officers (n = 341) 5. Large university capstone business students (n = 515)	Evaluate revised TREO scale reliabilities and factor structure, and their discriminant validity from measures of Big 5 personality constructs.
6. Small college business students (n = 172) in a strategy course	Evaluate the test–retest stability of the TREO measures
7. Large university capstone business students (n = 225) playing complex simulation	Evaluate the predictive validity of TREO scale as related to peer ratings of members' behaviors during a 10-week business simulation

Note. TREO = Team Role Experience and Orientation.

validity across the orientation and behavioral experience subscales, and their discriminant validity from measures of the Big 5 personality scales.

We made some minor wording changes to some items, and then administered the TREO and Big 5 items to another sample of 341 military officers and a second sample of 515 upper-level undergraduate business students for cross-validation purposes. Again, analyses of these data confirmed the internal consistencies of the different scales, their convergent validity across the orientation and behavioral experience subscales, and their discriminant validity from measures of the Big 5 personality scales. We further generated hypothesized patterns of correlations between the TREO and Big 5 measures. Then, using "bare bones" meta-analyses (Hunter & Schmidt, 2004), we synthesized the results of the four samples and concluded that the TREO scales, by and large, exhibited a pattern of correlations with the Big 5 measures that were consistent with the anticipated nomological network. Finally, we used a fifth sample of 172 upper-level business students from a small college to evaluate the stability of TREO scores over a 2.5-month long period. Each of the dimensions exhibited significant test–retest correlations (mean $r = .52$, $p < .001$).

In sum, the supplement to this article details how the 48-item TREO measurement tool (shown in the appendix) was developed and its psychometric

properties. The separate orientation and behavioral experience facets of each dimension provided unique variance and information, while also demonstrating significant convergent validity. The TREO measures are distinguishable from, yet generally showed the expected pattern of correlations with, measures of the Big 5 personality dimensions. And the TREO measures illustrated significant and reasonably high test–retest reliabilities over a 2.5-month long period. Nevertheless, it is also important to test whether the self-report TREO predisposition measures actually predict the extent to which individuals enact these different team role behaviors in group settings. The following investigation was focused on that very question.

Method

Sample

A sample of 225 students who were enrolled in seven capstone business strategy courses taught at a large northeastern public university were recruited for this study. Besides typical class activities, the students worked in four- to six-person teams on a business strategy simulation for 10 weeks. We collected TREO survey responses and demographics from students using an online survey before they began the simulation in their classes. Reliabilities for the TREO scales were all above .80 and are detailed in Table 4. We collected peer ratings of members' TREO-related behaviors exhibited during the simulation 3 months later using another online survey. We had complete information for 196 students who were members of 66 teams that competed in the simulation. This sample was 64% men with an average age of 21 (SD = 5.1), and 87% reported their ethnicity as White with the remainder distributed across other categories. They represented a spectrum of business majors (Accounting = 20%, Finance = 29%, Marketing = 14%, Management = 19%, other = 18%) and had an average overall grade point average of 3.51 (SD = .27).

Simulation

The "StratSim Management" business strategy simulation is based on the automobile industry and students function as the top management teams of virtual firms competing with one another (S. W. James & Deighan, 2008). The simulation calls for teams to make strategic decisions over time concerning a number of business functions (i.e., operations, marketing, finance, and R&D) for their firms for each weekly round. Students' grades were based, in part, on how well their team performed, as measured by a variety of objective indices (e.g., the change in firm stock price).

Table 4. Predictive Validity of TREO Scales as Related to Peer Ratings of Team Contributions.

	Peer ratings of TREO exhibited behaviors						
TREO scales	Organizer	Doer	Challenger	Innovator	Team Builder	Connector	Overall
Organizer	*.33***	.15*	.09	.09	.23**	.12	.27**
Doer	.20**	*.18***	.07	.14*	.15*	.06	.17*
Challenger	.12	.04	*.16**	.19*	.10	.08	.11
Innovator	.03	.06	.12	*.18**	.14	.05	.07
Team Builder	.21**	.11	.00	.10	*.24***	.17*	.14*
Connector	.10	.09	.06	.16*	.16*	*.10*	.10
TREO scale *M*	3.80	4.26	3.73	3.81	4.04	3.60	
Scale *SD*	.54	.41	.47	.52	.46	.59	
Alpha	.85	.83	.88	.84	.82	.85	
Peer rating *M*	3.81	4.04	3.81	3.78	3.78	3.63	4.57
Peer rating *SD*	.79	.72	.65	.68	.68	.83	.59
Median r_{wg}	.83	.83	.75	.75	.75	.65	.88

Note. n = 196. TREO = Team Role Experience and Orientation.
*p < .05. **p < .01.

The simulation lasted 10 weeks, with each week representing one virtual year of operation. During each period, teams used a decision interface to gather information about topics including virtual macroeconomic parameters, general market conditions, competitive landscape, and a rich array of demographic information on customer segments and their preferences. They made a wide variety of decisions concerning matters such as product attributes (e.g., safety, capacity, economy, interior design); operations (e.g., manufacturing capacity); sales and marketing (e.g., pricing, targeted or generic advertising, geographic dealership coverage); financial decisions (e.g., cash flow, issuing or buying back bonds, loans, and common stock); as well as other strategic decisions such as launching new models or discontinuing existing models, bidding for business-to-business contracts, and licensing to and from competitors. Team decisions were uploaded and processed by the software in a dynamic and competitive simulation environment, which incorporated all decisions of all teams in any given industry, thus leading to considerable variance between industry-level outcomes, such as competitiveness, innovativeness, rivalry, and so on. The simulation software then provided outcome figures for macroeconomy, customer preferences, industry, and each firm's performance. Detailed feedback reports were made available for each company, along with nonproprietary information on competitors, as well as the industry and the macroeconomic environment as a whole, per demand of

teams within the simulation interface. Notably, members were free to structure themselves and develop various roles and modes of operation as they desired.

Peer Ratings of TREO Behaviors

At the completion of the simulation, we collected self and peer evaluations of members' TREO-related behaviors exhibited during the simulation. Students were assured that their ratings would remain confidential. Although they did not translate directly to other students' grades, these ratings were used as input to instructors who assigned class participation grades for the course. We used just the average peer evaluations as our criteria, and study participants received ratings from 1 to 5 peers (median = 4 peer ratings).

Students rated each of their team members on the six TREO dimensions and provided an overall rating. It is important to emphasize that these ratings are about the extent to which each member exhibited TREO behaviors during their group activities over the course of the simulation. For the TREO behavior ratings, we provided the definitions of each dimension (see Table 2) and asked members to "provide ratings on six scales concerning *the extent to which your teammates' contributed effectively to each of the roles*." Responses were made on the following 5-point scale: 5 = *To a very great extent: This person exhibited a great deal of this behavior*; 4 = *To a great extent: This person exhibited more than an average amount of this behavior*; 3 = *To a moderate extent: This person exhibited about an average amount of this behavior*; 2 = *To a slight extent: This person exhibited some, but less than average, amount of this behavior*; and 1 = *Not at all: This person did not exhibit any of this behavior*. We used the following item for the *overall rating*: "For this rating, think about the total contributions that each person made to your team effort. This need not be simply an average of the other ratings." We instructed participants to select from the following the most appropriate score for each person: *Poor*: We would have been better off without this person. She or he was disruptive or simply didn't show up or do anything for the group (score 1); *Fair*: She or he really didn't pull her or his weight (e.g., didn't do a fair share of the work, missed meetings, etc.) and expected others to do the work (2); *Average*: a good team player (3); *Good*: She or he was a great help to others and did at least her or his share of the work (4); or *Excellent*: She or he was absolutely critical to our team. She or he did substantially more than most people and helped everyone understand and accomplish what we were doing (5). We used L. R. James, Demaree, and Wolf's (1984) single item r_{wg} agreement index to assess inter-rater agreement on each of these ratings. As reported in Table 4, median r_{wg} values were .75 or higher for all dimensions except for Connector, which had an r_{wg} = .65.

Results

Table 3 presents correlations between participants' TREO responses and their peers' ratings of the extent to which they exhibited TREO-related behaviors over the course of the 10-week simulation. Predictive validities are italicized in Table 4 and show significant correlations for all dimensions (*Organizer* $r = .33$, $p < .01$; *Doer* $r = .18$, $p < .01$; *Challenger* $r = .16$, $p < .05$; *Innovator* $r = .18$, $p < .05$; *Team Builder* $r = .24$, $p < .01$) except *Connector* ($r = .10$, *ns*). Moreover, the corresponding predictive validity coefficients were the highest correlations per peer rating for *Organizer, Doer, Challenger*, and *Team Builder* dimensions.

Among the non-predictive validity correlations that were evident, TREO *Doer* ($r = .20$, $p < .01$) and *Team Builder* ($r = .21$, $p < .01$) scores correlated significantly with peers' ratings of members' *Organizer behaviors*, whereas TREO *Organizer* ($r = .23$, $p < .01$) and *Doer* ($r = .15$, $p < .05$) scores also correlated with peers' ratings of *members' team building behaviors*. Interestingly, TREO *Doer* ($r = .14$, $p < .05$), *Challenger* ($r = .19$, $p < .05$), and *Connector* ($r = .16$, $p < .05$) scores all correlated significantly with members' ratings of *Innovator type behaviors*. In terms of predicting peers' overall evaluations of members' contributions to the team, the TREO dimensions of *Organizer* ($r = .27$, $p < .01$), *Doer* ($r = .17$, $p < .05$), and *Team Builder* ($r = .24$, $p < .05$) all evidenced significant correlations.

In sum, the results from this sample are encouraging. First, peers evidenced sufficient agreement on the extent to which students exhibited TREO aligned behaviors while working together on a business simulation. Second, TREO scale responses demonstrated significant predictive validity with their corresponding peer ratings of their behaviors 3 months later. Third, the exception to this pattern concerned the Connector dimension for which peers had markedly lower inter-rater agreement and the predictive validity was not significant. This is not particularly surprising, however, as the business simulation afforded virtually no opportunities for boundary-spanning activities to manifest. Finally, at least in terms of this business simulation context, the bundle of high *Organizer, Doer*, and *Team Builder* TREO scores appears to be associated with high ratings of individuals' subsequent team contributions.

Discussion

We had several goals for this article. First, we offered a review and synthesis of the team role literature. We argued that better incorporating individuals' potential to occupy various team roles holds great promise for advancing our understanding of team composition. We distilled six integrative types of team

roles and developed an accompanying TREO survey measure that considers both individuals' previous team's role-related experiences, as well as their preferences to engage in certain types of role-related behaviors.

SMEs worked to establish the content validity of the items per dimension. We demonstrated the convergent and discriminant validity of TREO measure, detailed their correlations with measures of individuals' Big 5 personality constructs, and illustrated their stability over time. Moreover, we demonstrated the predictive validity of the TREO scale as related to peers' ratings of team members' role-related behaviors. Below we consider the implications of these findings in terms of future theory and research incorporating individuals' predispositions to fulfill different team roles in the larger team composition domain. We also consider the use of the TREO measure in future work, and note some limitations of our investigations. We conclude with recommendations for practice.

Role Theory

Team role theory dates back to at least the 1940s and Kurt Lewin's Research Center for Group Dynamics. In the years since then, numerous taxonomies have been advanced, lists of team roles have proliferated, and there have been limited attempts at integration. In areas where integration has occurred (e.g., Belbin's framework), the construct validity of measurement tools has been questioned (e.g., Aritzeta et al., 2007; Broucek & Randell, 1996; Dierendonck & Groen, 2011; Furnham et al., 1993) and freely available measures are lacking. We offer an integration of this literature in terms of six team role dimensions, and have developed, validated, and made freely available an instrument (TREO) to measure individuals' propensities to occupy the roles for research purposes.

We should note that the intercorrelations of the six TREO dimensions were fairly high, averaging approximately $r = .70$, $p < .001$, across our item development and validation samples. We hasten to add, however, that those correlations are among *latent variables* that, in effect, are corrected for measurement attenuation. In other words, latent variable correlations are always higher than observed correlations if the measurement scales have less than perfect reliabilities. Therefore, although the high inter-dimensional correlations are still a source of concern, they are not as problematic as they appear using conventional thresholds. The confirmatory factor analysis results consistently illustrated that the TREO dimensions were empirically distinguishable. Moreover, the TREO scale correlations with the Big 5 personality variables differed across dimensions, indicating that they are not redundant (see supplement Table S8). That said, it does appear that there could be a

higher order latent structure underlying the six dimensions: (a) a task-oriented factor (i.e., Organizer and Doer), (b) a change-oriented factor (i.e., Challenger and Innovator), and (c) a socio-emotional or linkage dimension (i.e., Team Builder and Connector). Whether there is value in differentiating the six TREO dimensions or operating at a higher level of abstraction remains a question for future research. We suspect that the relative value of either approach would hinge on the purpose of any particular investigation. But what is clear from our and previous research is that it is not likely the case that there are 10 or more clearly distinguishable team role dimensions. However, there is evidence to support the six TREO dimensions.

We restricted our work to examining TREO as a measure of individual differences and differentiated it from measures of other individual differences. Certainly, TREO can be used to predict the likelihood that individuals will take on certain type of roles and exhibit certain types of behaviors in teams. But the TREO measure is also perhaps valuable in terms of assessing and testing team compilation models. For example, the popular Belbin (1993) framework advocates composing a team to have an ideal "balance" of styles. Individuals are categorized as one of nine types on the basis of their or observers' ratings, and presumably teams are more effective to the extent that they have a greater rather than lesser variety of team roles present. The empirical evidence on the value of such balance is inconsistent (cf. Senior, 1997; Smith et al., 2012), and there are shortcomings in the way in which the framework is implemented (e.g., classifying individuals ipsatively into one and only one type, equating different varieties of balance, etc.). Nevertheless, the general idea that different member combinations or profiles may be more or less advantageous in different situations warrants far more scrutiny (Mathieu et al., 2014).

We advocate using individuals' TREO scores in a more holistic fashion than merely to categorize members into types. For example, perhaps a team might benefit from having at least one member who is high on Organizer scores, one who is high on team building, and at least two who have high Doer scores. In this situation, there may be two different individuals who fulfill the Organizer and team building needs, or there may be a single person who fulfills both needs for a team. Naturally, this begs the question of whether it is preferable for one person to be occupying multiple team roles, or whether it is better to spread the fulfillment of different roles across members. This represents an interesting question for future theory, research, and application.

Roles may be naturally related to certain positions (Ilgen & Hollenbeck, 1991; Sluss et al., 2011). For example, an individual may occupy a formal leadership position whereby he or she would be expected to organize work, provide individuals with performance-related feedback, and take

responsibility for team functioning. However, those functions may well be performed by others in the group and are not necessarily limited to a person who occupies a particular position in the team (Manz & Sims, 1980). In other words, roles are emergent phenomena and are not tied inextricably to formal positions. Nevertheless, future scholars might consider the value of different member TREO profiles in team situations where individuals are fixed into certain positions (i.e., where there are substantial "role as position" by "role as person" interactions). For example, what are the implications of a surgeon, head nurse, or anesthesiologist scoring high on team building? What are the implications of the most junior member of a team versus a designated leader scoring the highest on the TREO Challenger dimension? In short, the team role area is ripe for development. We believe that optimal team compositions are not likely to be adequately captured by a simple average or variety (i.e., variance) index; rather, we believe that configural or compilation (Kozlowski & Ilgen, 2006) combinations are likely to better capture the complexities of team memberships. What such combinations are and whether they are a product of situational demands, members' work histories, or the extent to which teams have rigid or loosely defined positions (if positions at all) are all interesting questions worthy of pursuit. The TREO measure will help to enable those inquiries.

Limitations

We developed the TREO measures using SME input and then explored different facets of its construct validity using seven different samples. Although leveraging multiple samples is certainly advantageous, the fact that we sampled military officers and business students does raise questions about the generalizability of our findings. Additional evidence of the construct validity of the TREO measures using different populations is warranted. Moreover, the relative importance of TREO versus other individual differences is worthy of investigation. For example, are individuals with certain TREO profiles seen as more valuable team members than individuals who possess task-specific knowledge, general intelligence, or experience in a particular domain? Do TREO profiles interact with any of these other KSAOs as related to important individual or team outcomes? Additional research along these lines would be welcome.

The TREO dimensions did evidence relatively high intercorrelations. Notably, items are implicitly worded in terms of positive attributes—about one's propensity to fulfill important team roles. Thus, dimension correlations may be inflated to the extent that respondents exhibited any social desirability bias. Incorporating measures that assess such bias would be a welcome

addition in future investigations. It may also be valuable to develop force-choice or other ipsative method of measurements that would yield sharper TREO profiles, although such techniques introduce different psychometric models and considerations. In a related vein, we have discussed TREO dimensions as though greater amounts of each are always advantageous. However, there may well be value in examining whether there are tipping points or thresholds beyond which a positive attribute turns into a negative one (see Smith et al., 2012). For example, whereas a propensity to organize team activities is generally seen as a positive influence, at some point, such structuring may lapse into being overly controlling and stifling. Some challenging behavior is healthy for teams and prevents them from prematurely closing discussion or failing to consider alternatives. Yet too much challenging is likely to breed contempt and be viewed as antagonistic. There may also be problematic combinations or profiles. For example, someone who is exceedingly high on organizing and doing, yet very low on team building, may be perceived as an overly dominating force in a team. Or, someone who scores very high on challenging and connecting may be perceived as a negative element who is not invested in the team. In other words, tipping points, acidic members, and team profiles all warrant future investigation.

Applications

Improving our understanding of team role propensities and being able to measure them in a reliable manner can have several practical applications starting with initial team formation through ongoing team, leader, and personal development. When composing a team, knowledge of candidates' team role propensities could be used to avoid configurations that are more likely to result in team dysfunction. For example, Belbin (1993) recommends composing teams to maximize members' role differentiation. However, optimal (and suboptimal) team configurations may well be more complex than simply establishing diversity. For example, a team that is low on Doers or has no one who is an Organizer is apt to struggle, as is a team with too many or too few Challengers. Research is needed to better understand the types of profiles that are likely to be problematic, but a psychometrically sound measure is the starting point for such research.

Once a team is formed, understanding its members' TREO profile may provide insights as to the utility of other team interventions. Use of the TREO tool can provide a common language for discussing team members' preferences and inclinations, enhancing awareness, and perhaps enabling the team to better anticipate each other's reactions in various situations. In effect, knowing each other's predispositions may enhance the effectiveness of team

interventions such as charters, team training and development, debriefs, and leadership strategies (Mathieu et al., 2014).

Finally, the TREO can be used during leader training and as a self-awareness tool. For example, leaders could be taught about the various team role propensities and how best to lead individuals and teams with certain team role profiles. Similarly, an individual who completes a tool such as the TREO may learn to better understand their natural inclinations in team settings, increase their self-awareness, and better recognize when they need to operate in a manner that is not their natural "default" mode. Naturally, the applications described here for team composition and development should be tested in future research. All such applications are predicated on having a well-validated measure such as TREO readily available.

Appendix

TREO Scale Items

Based on my prior experiences, as a member of different teams . . .

1. I learn how to get outside resources that our team needs to be successful.
2. I'm comfortable being critical of my teammates.
3. I like it when we keep busy and get things done.
4. I like to challenge peoples' assumptions.
5. I like to be the one that sorts out the details of a team project.
6. I often volunteer new ideas and suggestions without being asked my opinion.
7. I can calm people down and get them focused on the task when things get stressful.
8. I like to be the one who decides who will do which tasks on a team.
9. I am the one who questions why we are doing things in a certain way.
10. Sometimes, I just voice a different opinion to keep my team thinking about what we should be doing.
11. I'm always ready to support a good suggestion in the common interest of the team.
12. People usually look to me when something needs to be done in the team.
13. I like to try out new ideas and approaches.
14. I question what my team should be doing to get the job done.
15. I can be counted on to follow through on any tasks which I've been assigned.

16. I can be counted on when a task needs to be done.
17. I keep my team on pace and aware of deadlines.
18. I make sure that my teammates are clear about their responsibilities.
19. I'm comfortable dealing with interpersonal conflicts and helping people work through them.
20. I enjoy coordinating team efforts with people or groups outside of the team.
21. My primary focus is on getting my assignments done for the team.
22. I can be counted on to spread ideas between my team and people outside of my team.
23. I'm comfortable being the spokesperson for a team.
24. I am the one who steps up and does whatever is necessary to make the team successful.
25. I'm often the first to volunteer for a difficult or unpopular assignment if that is what the team needs.
26. I like to be the one who keeps track of how well my team is doing.
27. I am usually the one who suggests a new idea or direction when the team gets stuck on something.
28. I bring a sense of organization to any job a team undertakes.
29. I get bored when we do the same task the same way every time.
30. I structure team activities.
31. I discover and connect with people who can help my team succeed.
32. I'm not afraid to question my teammates' authority.
33. I'm known for thinking creatively and "outside the box."
34. I typically find out what is going on outside my team and share that with my teammates.
35. I like coming up with new ways that our team can accomplish our tasks.
36. I usually suggest the appropriate steps that my team should follow to get something done.
37. I like helping different kinds of people work effectively together.
38. I'm comfortable producing and sharing new ideas with my team.
39. I often work to maintain good working relationships within my team.
40. It bothers me when I see teammates getting frustrated or depressed.
41. I'm always committed to my team tasks.
42. I often point out the potential risks or hazards of a team plan or course of action.
43. I help people move beyond their disagreements and find common ground.
44. My teammates often view my suggestions as creative or innovative.
45. I often serve as a liaison between my team and outside groups.

46. I promote my team's mission and goals with other teams or units.
47. I can typically provide a strong rationale to refute ideas that I believe are unsound.
48. I encourage my teammates when I know they have a difficult assignment or challenge.

Item Mapping

Organizer: 5O, 8O, 17E, 18E, 26O, 28O, 30E, 36E.
Doer: 3O, 12E, 15E, 16O, 21E, 24E, 25O, 41O.
Challenger: 2O, 4O, 9E, 10E, 14E, 32O, 42E, 47O.
Innovator: 6E, 13O, 27E, 29O, 33E, 35O, 38O, 44E.
Team Builder: 7E, 11O, 19O, 37O, 39E, 40O, 43E, 48E.
Connector: 1E, 20O, 22O, 23O, 31E, 34O, 45E, 46E.

Note. E = experience subscale; O = orientation subscale.

Authors' Note

All opinions expressed herein are strictly those of the authors and not necessarily those of the sponsoring organizations.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The work was supported by Grant NNX11AR22G from the National Aeronautics and Space Administration (NASA) and Contract W91WAW-08-C-0021 from the Army Research Institute awarded to the authors.

References

Ancona, D. G., & Caldwell, D. F. (1988). Beyond task and maintenance: Defining external functions in groups. *Group & Organizational Studies*, *13*, 468-494.

Ancona, D. G., & Caldwell, D. F. (1992). Bridging the boundary: External activity and performance in organizational teams. *Administrative Science Quarterly*, *37*, 634-665.

Anderson, N., & Sleap, S. (2004). An evaluation of gender differences on the Belbin Team Role Self Perception Inventory. *Journal of Occupational and Organizational Psychology*, *77*, 429-437.

Aritzeta, A., & Ayestaran, S. (2003). Applicability of Belbin's team roles theory: A longitudinal and comparative study with work teams. *General and Applied Psychology*, *56*, 61-75.

Aritzeta, A., Senior, B., & Swailes, S. (2005). Team role preference and cognitive styles: A convergent validity study. *Small Group Research*, *36*, 404-436.

Aritzeta, A., Swailes, S., & Senior, B. (2007). Belbin's team role model: Development, validity and applications for team building. *Journal of Management Studies*, *44*, 96-118.

Bales, R. F. (1950). *Interactive process analysis: A method for the study of small groups*. Cambridge, MA: Addison-Wesley.

Bales, R. (1970). *Personality and interpersonal behaviour*. New York: Holt, Rinehart and Winston.

Barry, D. (1991). Managing the bossless team: Lessons in distributed leadership. *Organizational Dynamics*, *20*, 31-47.

Belbin, R. M. (1981). *Management teams: Why they succeed or fail*. London, England: Butterworth-Heinemann.

Belbin, R. M. (1993). *Team roles at work*. London, England: Butterworth-Heinemann.

Bell, S. T. (2007). Deep-level composition variables as predictors of team performance: A meta-analysis. *Journal of Applied Psychology*, *92*, 595-615.

Benne, K. D., & Sheats, P. (1948). Functional roles of group members. *Journal of Social Issues*, *4*, 41-49.

Broucek, W. G., & Randell, G. (1996). An assessment of the construct validity of the Belbin Self-Perception Inventory and observer's assessment from the perspective of the five-factor model. *Journal of Occupational and Organizational Psychology*, *69*, 389-405.

Brown, S. H. (1978). Long-term validity of a personal history item scoring procedure. *Journal of Applied Psychology*, *63*, 673-676.

Cheung, M. W., & Chan, W. (2002). Reducing uniform response bias with ipsative measurement in multiple-group confirmatory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*, 55-77.

Davis, K. R. (1984). A longitudinal analysis of biographical sub-groups using Owen's development-integrative model. *Personnel Psychology*, *37*, 1-14.

Dierendonck, D. V., & Groen, R. (2011). Belbin revisited: A multitrait-multimethod investigation of a team role instrument. *European Journal of Work & Organizational Psychology*, *20*, 345-366.

DuBrin, A. J. (1995). *The breakthrough team player*. New York, NY: American Management Association.

Fisher, S. G., Hunter, T. A., & Macrosson, W. D. K. (2001). A validation study of Belbin's team roles. *European Journal of Work & Organizational Psychology*, *10*, 121-144.

Furnham, A., Steele, H., & Pendleton, D. (1993). A psychometric assessment of the Belbin Team-Role Self-Perception Inventory. *Journal of Occupational and Organizational Psychology*, *66*, 245-257.

Hackman, J. R. (1990). *Groups that work and those that don't*. San Francisco, CA: Jossey-Bass.

Hambrick, D. C. (1994). Top management groups: A conceptual integration and reconsideration of the "team" label. *Research in Organizational Behavior*, *16*, 171-214.

Hollenbeck, J. R., Beersma, B., & Schouten, M. E. (2012). Beyond team types and taxonomies: A dimensional scaling conceptualization for team description. *Academy of Management Review*, *37*, 82-106.

Humphrey, S. E., Morgeson, F. P., & Mannor, M. J. (2009). Developing a theory of the strategic core of teams: A role composition model of team performance. *Journal of Applied Psychology*, *94*, 48-61.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: SAGE.

Ilgen, D. R. (1999). Teams embedded in organizations: Some implications. *American Psychologist*, *54*, 129-139.

Ilgen, D. R., & Hollenbeck, J. R. (1991). The structure of work: Job design and roles. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. *2*, pp. 165-208). Palo Alto, CA: Consulting Psychologists Press.

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in organizations: From I-P-O models to IMOI models. *Annual Review of Psychology*, *56*, 517-543.

Jackson, C. (2002). Predicting team performance from a learning process model. *Journal of Managerial Psychology*, *17*, 6-13.

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, *69*, 85-98.

James, S. W., & Deighan, M. (2008). *StratSim: Administrators guide*. Charlottesville, VA: Interpretive Software.

Jung, C. G. (1923). *Psychological types*. London, England: Routledge and Kegan Paul.

Katz, D., & Kahn, R. L. (1978). *The social psychology of organizations* (2nd ed.). New York, NY: John Wiley.

Kirnan, J. P., & Woodruff, D. (1994). Reliability and validity estimates of the Parker Team Player Survey. *Educational and Psychological Measurement*, *54*, 1030-1037.

Klimoski, R. J., & Zukin, L. B. (1999). Selection and staffing for team effectiveness. In E. Sundstrom (Ed.), *Supporting work team effectiveness* (pp. 63-94). San Francisco, CA: Jossey-Bass.

Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science*, *7*, 77-124.

Larson, C. E., & LaFasto, F. M. (1989). *Teamwork: What must go right/wrong can go wrong*. Newburg Park, CA: SAGE.

Lessem, R., & Baruck, Y. (2000). Testing the SMT and Belbin inventories in top management teams. *Leadership and Organization Development Journal*, *34*, 75-83.

Manz, C. C., & Sims, H. P. J. (1980). Self-management as a substitute for leadership: A social learning theory perspective. *Academy of Management Review*, *5*, 361-367.

Margerison, M., & McCann, D. (1985). *How to lead a winning team*. Bradford, UK: MCB University Press.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review*, *26*, 356-376.

Mathieu, J. E., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future. *Journal of Management*, *34*, 410-476.

Mathieu, J. E., Tannenbaum, S. I., Donsbach, J. S., & Alliger, G. M. (2014). A Review and Intergration of Team Composition Models: Moving Towards a Dynamic and Temporal Framework. *Journal of Management*, *40*, 130-160.

McCann, D., & Margerison, C. (1989). Managing high-performance teams. *Training and Development Journal*, *43*, 52-60.

Mumford, M. D., & Owens, W. A. (1984). Individuality in a developmental context: Some empirical and theoretical considerations. *Human Development*, *27*, 84-108.

Mumford, T. V., Campion, M. A., & Morgeson, F. P. (2006). Situational judgment in work teams: A team role typology. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 319– 343). Mahwah, NJ: Erlbaum.

Mumford, T. V., Van Iddekinge, C. H., Morgeson, F. P., & Campion, M. A. (2008). The team role test: Development and validation of a team role knowledge situational judgment test. *Journal of Applied Psychology*, *93*, 250-267.

Myers, I. B., McCaulley, M. H., & Most, R. (1985). *Manual: A guide to the development and use of the Myers-Briggs type indicator*. Palo Alto, CA: Consulting Psychologists Press.

Neiner, A. G., & Owens, W. A. (1982). Relationships between two sets of biodata with 7 years separation. *Journal of Applied Psychology*, *67*, 146-150.

Park, W. W., & Bang, H. (2002). *Team role balance and team performance*. Paper presented at the Biannual Conference on Belbin's Team Roles, Cambridge, UK.July.

Parker, G. M. (1994). *Cross-functional teams*. San Francisco, CA: Jossey-Bass.

Parker, G. M. (1996). *Team players and teamwork*. San Francisco, CA: Jossey-Bass.

Partington, D., & Harris, H. (1999). Team role balance and team performance: An empirical study. *Journal of Management Development*, *18*, 694-705.

Ronnie, L., & Yehuda, B. (2000). Testing the SMT and Belbin inventories in top management teams. *Leadership & Organization Development Journal*, *21*, 75-84.

Rushmer, R. (1996). Is Belbin's shaper really TMS's thruster-organizer? An empirical investigation into the correspondence between the Belbin and TMS team role models. *Leadership & Organizational Development Journal*, *17*, 20-26.

Sadler-Smith, E. (2001). Self-perception of team-roles: Some of the implications for business and management. In R. Riding & S. G. Rayner (Eds.), *International perspectives on individual differences* (Vol. 2, pp. 101-130). Westport, CT: Ablex.

Senior, B. (1997). Team roles and team performance: Is there "really" a link? *Journal of Occupational and Organizational Psychology*, *70*, 241-258.

Senior, B. (1998). An empirically-based assessment of Belbin's team roles. *Human Resource Management Journal*, *8*, 54-60.

Sluss, D. M., van Dick, R., & Thompson, B. S. (2011). Role theory in organizations: A relational perspective. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology* (Vol. 1, pp. 505-534). Washington, DC: American Psychological Association.

Smith, M., Polglase, G., & Parry, C. (2012). Construction of student groups using Belbin: Supporting group work in environmental management. *Journal of Geography in Higher Education*, *36*, 585-601.

Sommerville, J., & Dalziel, S. (1998). Project teambuilding—The applicability of Belbin's Team-Role Self-Perception Inventory. *International Journal of Project Management*, *16*, 165-171.

Stewart, G. L., Fulmer, I. S., & Barrick, M. R. (2005). An exploration of member roles as a multilevel linking mechanism for individual traits and team outcomes. *Personnel Psychology*, *58*, 343-365.

Stewart, G. L., Manz, C. C., & Sims, H. P. J. (1999). *Team work and group dynamics*. New York, NY: Wiley.

Sundstrom, E., De Meuse, K. P., & Futrell, D. (1990). Work teams: Applications and effectiveness. *American Psychologist*, *45*, 120-133.

Author Biographies

John E. Mathieu is a Professor of Management at the University of Connecticut, and holds the Friar Chair in Leadership and Teams at UConn. His primary areas of interest include models of team and multi-team effectiveness, leadership, training effectiveness, and cross-level models of organizational behavior.

Scott Tannenbaum is President of the Group for Organizational Effectiveness (gOE), a New York-based consulting and research firm. Formerly a tenured professor, his research has been cited over 8000 times and he is a Fellow of SIOP and APS.

Michael R. Kukenberger is an assistant professor in the Department of Management at the Peter T. Paul School of Business and Economics at University of New Hampshire. He received his Ph.D. in Organizational Behavior at the University of Connecticut and his primary areas of research interest include shared and team leadership, team effectiveness, team and individual learning and multilevel theory building.

Jamie S. Donsbach is a Senior Consultant with The Group for Organizational Effectiveness, Inc., a consulting and research firm. Throughout her career, Dr. Donsbach has supported corporate and public sector customers with a variety of organizational needs such as enhancing team effectiveness, training, identifying competency gaps, managing performance, and developing/validating assessment tools. Dr. Donsbach received her doctorate in Industrial/Organizational Psychology from the University at Albany, State University of New York.

George M. Alliger, PhD, is VP of Solutions for the Group for Organizational Effectiveness. Dr. Alliger has over many years conducted research in and written about job analysis, training, and assessment in organizations. He recently co-edited the Handbook of Work Analysis.

Team Role Experiences and Orientations: A Measure and Tests of Construct Validity

[On-line supplement]

In this supplement to Mathieu, Tannenbaum, Kukenberger, Donsbach and Alliger (2015, *Group & Organization Management, 40*(1), 6 - 34), we describe the development of Likert-type scales for measuring the six Team Role Experiences and Orientations (TREO) dimensions. We first outline anticipated relationships between the TREO measures and measures of the Big 5 personality dimensions. We then detail the generation of scale items, the judgment of their content validity by SMEs, and then report the results of four empirical investigations of their convergent and discriminant validity. We further test whether the experience versus orientation subscales yield comparable role profiles. Finally, we report the stability of the measures over a two and half month period using a fifth sample.

As for relationships in the nomological network, Table S1 outlines the pattern of correlations that we anticipate between the TREO dimensions and the Big 5 personality dimensions of Conscientiousness, Extraversion, Agreeableness, Emotional Stability, and Openness to Experience (Stewart, Fulmer & Barrick, 2005). The Big 5 is perhaps the most widely used taxonomy of personality variables in applied psychology and management. Definitions of each dimension appear in Table S2. Although various dimensions may exhibit significant correlations in a nomological network, we anticipate that the TREO scales will evidence a particular pattern of predominant correlations with measures of the Big 5. For example, Organizers and Doers have a task orientation and focus on structuring work for the team and getting it done. Therefore, we hypothesize that those two dimensions will correlate positively and primarily with Conscientiousness and Emotional Stability. We further anticipate that the TREO Doer dimension will correlate highly with Agreeableness from the Big 5. In contrast, TREO Challengers and Innovators are interested in questioning the status quo and proposing new and different ideas and work procedures. We hypothesize that Challengers will correlate positively with Openness and Extraversion, but perhaps negatively with Agreeableness from the Big 5. This follows from the fact that challenging behaviors can be seen as antagonist and disruptive to the social fabric of a team. However, we hypothesize that Innovators will correlate positively with Conscientiousness, Extraversion, and Openness. Team Builders and Connectors are about facilitating linkages between people, and we hypothesize that both will correlate positively with Agreeableness, Extraversion and Openness. We further anticipate that Team Builders will correlate positively with Extraversion.

CONTENT VALIDATION

Method

TREO Design and Content Validation

Content validity is defined by Haynes, Richard, & Kubany (1995) as the extent to which "elements of an assessment instrument are relevant to and representative of the targeted construct for a particular assessment purpose (p. 238)." In establishing the content validity of the TREO instrument we followed the general recommendations set out by McKenzie, Wood and Kotecki (1999) and Rubio, Berg-Weger, Tebb, Lee and Rauch (2003). Specifically, we conducted a

review of the literature as described above to identify and define the content domain. We then examined and distilled the various roles described in the literature into the six TREO dimensions.

Item Generation. On the basis of the literature review, we generated 120 items with the goal of having at least fifteen items per TREO dimension. The goal of generating so many items was to adequately represent the content domain of each of the six constructs and have a large base to choose from. Within each role, we generated two sub-sets of items: role orientations and specific behavioral experiences. The authors then collectively reviewed and combined redundant items to yield a set of 105 which served as the basis for our content validation effort.

Subject Matter Expert (SME) Review & Categorization. We next employed a pretest assessment of item substantive (i.e., content) validities methodology outlined by Anderson and Gerbing (1991). Specifically, we secured a panel of 12 SMEs, six of whom were academics within the areas of organizational behavior/psychology with a focus on team/group research, and six of whom were industry management professionals whose daily work involved the extensive use of teams. These SMEs have published widely in the team research domain and actively consult with many leading organizations on team-related matters. Each SME was sent an electronic copy of the TREO dimensions and definitions and the 105 randomly ordered items. Their task was to classify each one of the items into one of six TREO dimensions. In order to ensure consistency across our two expert sub-groups we examined differences between the academics' and practitioners classifications using a chi square difference tests and found no significant differences.

Anderson and Gerbing (1991) describe the use of a *substantive-validity coefficient* (SVC) which ranges from -1 to 1, with larger values indicating greater item substantive validity. In other words, using the SMEs' item categorization ratings, items have a higher SVC value to the extent that they are classified consistently as belonging to their intended construct and not consistently to other constructs. Eighty-one of the 105 initial items had significant SVC values (i.e., SVC >.50, $p < .05$). From those, we selected a total of 48 items guided by their relative SVC values, but with the goal of including four orientation-focused items and four behavioral-focused items for each of the six TREO dimensions. The 48 TREO items were then administered to two item development samples and later to two validation samples as described below.

Development and Validation Samples

Military Samples. Two field samples were collected from U.S. Army Officers preparing to operate as transition team members. A U.S. Military Transition Team is a 10-15 Soldier team that lives with and trains Foreign Security Forces (FSF). The primary mission of Transition Teams is to advise the security forces in places such as Iraq and Afghanistan in the areas of intelligence, communications, fire support, logistics, and infantry tactics. The aim is to make the FSF capable of conducting independent counterinsurgency operations, tactically, operationally, and logistically (KMI Media Group, 2008). The two survey administrations were separated by approximately a year, with the first (Military A) being used for measurement development and the second (Military B) for validation. The Military A sample (N= 317) was 94% men and averaged approximately 12.14 (SD= 6.72) years of military service. Approximately 37% of the sample were Captains, 42% Sergeants, and the remainder distributed over other officer ranks.

The Military B sample (N= 341) was 93% men and averaged approximately 12.06 (SD= 6.95) years of military service. Approximately 26% were Captains, 33% Sergeants, and the remainder again distributed over other officer ranks. We should note that because of a typographical error, one Connector orientation item (Item 23, see article Appendix) was inadvertently left off of the Military Sample A survey.

Business Student Samples. Participants were undergraduate business majors and Masters of Business Administration (MBA) students enrolled in different sections of Management Capstone Strategy courses at a large northeastern public university. Business Student Sample A (N = 266) was 53% male with an average age of 23.47 (SD = 5.31). Their average overall GPA was 3.50 (SD = . 32) and they had a range of majors (20% Accounting, 28% Finance, 17% Marketing, 15% Management, and 20% other). Business Student Sample B (N = 515) was 63% male with an average age of 21.36 (SD = .85). Their average overall GPA was 3.45 (SD = . 29) and they too had a range of majors (31% Accounting, 30% Finance, 14% Marketing, 12% Management, and 13% other).

Measures

Besides reporting demographic information, all four samples completed the TREO (see article Appendix for the final instrument). The Military samples completed paper surveys whereas the Business students responded to on-line surveys. Respondents were asked to evaluate how well each TREO item described themselves using 1-5 Likert-type scales that ranged from "*Not at All*" to "*To a very great extent*." Details concerning the psychometric properties of the TREO measures are provided below. All participants also completed a 20-item "mini-markers" measure of Big 5 personality dimensions (Donnellan, Oswald, Baird & Lucas, 2006). Each scale contains four items and has demonstrated reasonable internal consistencies for research purposes (Donnellan et al., 2006). The scale alphas for our four samples were: *Extraversion*: .71, .77, .70, .73; *Agreeableness*: .61, .71, .61, .70; *Conscientiousness*: .64, .70, .60, .71; *Emotional Stability*: .62, .61; .68, .55; and *Openness to Experience*: .59, .63, .70, .65) for the Military A and B, and Business Student A and B samples, respectively.

ITEM DEVELOPMENT SAMPLES

Results

Sub-Scale Psychometrics & Mean Scores

Given the ratio of TREO items to sample size, we first examined the psychometric properties of the orientation and experience subscales with the intention of using the subscales as indicators in the construct validity Confirmatory Factor Analysis (CFA) models. As summarized in Table S3, all subscales exhibited alphas that ranged from .63 to .83, which are reasonable for 4-item measures. We then fit a single-factor CFA model to each set of items using MPlus 7.0 (Muthen & Muthen, 2012), which are also summarized in Table S3. To gauge model fit, we report the Comparative Fit Index (CFI; Bentler, 1990) and the Standardized Root Mean Squared Residual (SRMSR; Hu & Bentler, 1999). Following recommendations from Mathieu and Taylor (2006), we consider models with CFI values < .90 and SRMSR values > .10 as deficient, those

with CFI $\geq$ .90 to < . 95 and SRMSR > .08 to $\leq$ .10 ranges as acceptable, and ones with CFI $\geq$ .95 and SRMSR $\leq$ .08 ranges as excellent. The single factor CFIs ranged from .89 to 1.00 whereas the SRMRs ranged from .001 to .067. The only instance that was deemed as less than acceptable was the *Connector* experience subscale for the Military A sample. Given this pattern of results, we computed orientation and experience subscales by averaging their respective items per TREO dimension for use in further analyses.

We then computed a two-factor, within-subjects Analysis of Variance (ANOVA) using the subscales scores as the criterion. Although we anticipate similar patterns across TREO dimensions for the orientation and experience subscales, this analysis reveals whether the two facets are redundant or reflect different aspects of the latent constructs. The within-subjects factors were the (6) TREO *scales* crossed with the (2) orientation versus experience *foci*. Using the Military A data, we found significant main effects for both *scales* [$F(5,312) = 138.01$, $p <$ .001; partial eta^2 = .69] and *foci* [$F(1,316) = 63.32$, $p < .001$; partial eta^2 = .17], as well as a *scale by foci* interaction [$F(5,312) = 53.78$, $p < .001$; partial eta^2 = .46]. The scales effect was attributable to people reporting significantly ($p < .05$) higher *Doer* scores (M = 3.81), and relatively lower *Organizer* (M = 3.16), and *Challenger* Scores (M = 3.17), as compared to the other scales. In general, higher scores were obtained for the *orientation* foci (M = 3.47), as compared to the *experience* foci (M = 3.37). The interaction stemmed from three of the scales. For the *Innovator* (M = 3.48 versus M = 3.13) and *Doer* Scales (M = 3.97 versus M = 3.65), respondents reported significantly higher *orientation* as compared to their corresponding *experience* scores. The reverse was true for the *Organizer* scores where the reported *experiences* were significantly higher (M = 3.03 versus M = 3.30) than their reported *orientation*.

We computed a second 6X2 ANOVA using the Business Students A subscales scores as the criterion. We again found significant main effects for both *scales* [$F(5,261) = 37.47$, $p <$ .001; partial eta^2 = .42] and *foci* [$F(1,265) = 32.11$, $p < .001$; partial eta^2 = .11], and a significant *scale by foci* interaction [$F(5,261) = 45.51$, $p < .001$; partial eta^2 = .46]. The scales effect was attributable to individuals reporting relatively higher *Doer* scores (M = 4.50), and relatively lower *Connector* Scores (M = 3.95), as compared to the other scales. In general, higher scores were again seen for the *orientation* foci (M = 4.26), as compared to the *experience* foci (M = 4.17). Finally, the interaction stemmed from two of the scales. As observed with the Military A scores, the *orientation* subscales were significantly higher for both the *Innovator* (M = 4.43 versus M = 3.88), and *Doer* (M = 4.63 versus M = 3.38) scales.

In sum, these analyses indicate that the TREO subscales are sufficiently reliable and unidimensional to use as indicators. Moreover, the mean patterns suggest that the two foci represent somewhat different facets of the underlying dimensions. Combined, these findings are supportive of using both subscales as indictors of the corresponding TREO latent variables.

Convergent and Discriminant Validity

Using the 12 subscales as indicators, we fit a 6-factor TREO model to both samples. Table S4 summarizes the factor loadings on their respective latent variables, per sample. As shown, the CFA loadings ranged from .757 to .897 and all were significant $p < .001$. The 6-factor CFA model evidenced excellent fit indices for both the Military A [$\chi^2(39) = 124.07$, $p < .001$;

CFI = .968; SRMR = .031] and for the Business Students A [$\chi^2(39) = 128.68$, $p < .001$; CFI = .957; SRMR = .038] samples.

We then added the Big 5 measures and fit an 11 factor model to each sample. For these analyses, we randomly paired the four Big 5 items per dimension to form two parcels (Landis, Beal & Tesluk, 2000). Consequently, these models are testing the fit of 22 parcels to 11 underlying latent variables (i.e., six TREO and the Big 5 dimensions). The 11-factor CFA model evidenced acceptable fit indices for both the Military A [$\chi^2(154) = 390.36$, $p < 001$; CFI = .943; SRMR = .054] and for the Business Students A [$\chi^2(154) = 337.92$, $p < .001$; CFI = .937; SRMR = .046] samples. The correlations among the *latent variables* are reported in Table S5 for both samples. The Military A results appear in the lower left triangle, while the Business Students A results appear in the upper right triangle. Note that *the magnitudes of the reported correlations are among latent variables* which, in effect, have been adjusted for measurement attenuation. Naturally such correlations are larger than typical uncorrected correlations.

For the Military A sample, the largest correlation among the latent TREO dimensions was r = .91, p < .01 between the *Challenger* and *Innovator* dimensions. Consequently, we collapsed their indicators to a single latent variable and fit a 10-factor CFA model. While the resulting model evidenced acceptable overall fit indices [$\chi^2(155) = 401.72$, $p < .001$; CFI = .941; SRMR = .054], using a nested model contrast it exhibited a significantly worse [$\Delta\chi^2(1) = 11.36$, $p < .001$] model, as compared to the 11-factor model. These results indicated that, while highly correlated, the *Challenger* and *Innovator* TREO dimensions are empirically distinguishable.

Still using the Military A results, we observed an expected pattern of correlations. For example, both *Team Builder* and *Connector* TREO dimensions correlated significantly with Big 5 dimensions of *Extraversion* (*r*s = .41 and .46, $p < .01$), *Agreeableness* (*r*s = .50 and .40, $p < .01$), and *Openness to Experience* (*r*s = .36 and .34, $p < .01$), respectively. TREO *Innovator* correlated significantly with *Openness to Experience* from the Big 5 ($r = .36$, $p < .01$), and the TREO *Organizer* correlated with both *Extraversion* and *Agreeableness* personality dimensions (*r*s = .35 and .25, $p < .01$), respectively. Surprisingly, though, *Conscientiousness* exhibited weak correlations with TREO *Doer* ($r = .22$, $p < .05$), and *Organizer* dimensions ($r = .13$, *ns*).The largest correlation between a TREO and a Big 5 dimension was the *Team Builder – Agreeableness* pair ($r = .50$, $p < .001$). Accordingly, we collapsed their indicators to a single latent variable and fit a 10-factor CFA model. This model also evidenced acceptable overall fit indices [$\chi^2(155) = 489.83$, $p < .001$; CFI = .920; SRMR = .068], but was significantly worse [$\Delta\chi^2(1) = 99.47$, $p < .001$] as compared to the 11-factor model. In sum, these results indicated that the TREO dimensions evidenced the expected pattern of correlations with, but were empirically distinguishable from, the Big 5 personality dimensions.

We next conducted the same nested model tests using the Business Students A sample. The largest correlation among the TREO dimensions for this sample was $r = .87$, $p < .01$, again between the *Challenger* and *Innovator* dimensions. The 10-factor CFA model [$\chi^2(155) = 356.63$, $p < .001$; CFI = .931; SRMR = .047], again exhibited a significantly worse fit [$\Delta\chi^2(1) = 18.71$, $p < .001$] as compared to the 11-factor model.

The pattern of correlations between the TREO and Big 5 dimensions were encouraging with this sample. For example, the TREO *Team Builder* dimension correlated significantly with the *Extraversion*, *Agreeableness*, and *Conscientiousness* Big 5 dimensions (*r*s = .36, .36 and .30, $p < .01$), respectively. Both the TREO Organizer and *Doer* dimensions correlated significantly with *Conscientiousness* (*r*s = .59 and .63, $p < .001$), whereas the TREO *Innovator* dimension correlated significantly with both *Extraversion* and *Openness to Experience* from the Big 5 (*r*s = .35 and .54, $p < .01$), respectively. The largest correlation between a TREO and Big 5 dimension was the *Doer – Conscientiousness* pair ($r = .63$, $p < .001$). Accordingly, we collapsed their indicators to a single latent variable and fit another 10-factor CFA model. This 10-factor model [$\chi^2(155) = 382.41$, $p < .001$; CFI = .922; SRMR = .050] was significantly worse [$\Delta\chi^2(1) = 44.49$, $p < .001$] than the 11-factor model. In sum, the findings from the Business Students A sample paralleled those of the Military A sample, in that the TREO subscales mapped to six distinguishable latent variables as anticipated. The TREO dimensions were also shown to be distinguishable from measures of the Big 5 personality variables, but did exhibit the anticipated pattern of correlations that would be expected in the nomological network. Nevertheless, these results and user reactions suggested some areas of ambiguity and room for improvement. Consequently, we adjusted the wording on several items, most notably the *Connector* experience set, and administered them to another sample of Military Team (B) members, and another semester of Business Students Team (B) students to evaluate the refined scales and to further explore the construct validity of the TREO scales.[1]

VALIDATION SAMPLES

Results

Sub-Scale Psychometrics & Mean Scores

We administered the revised 48 item set to two additional samples for validation purposes. Table S6 provides a summary of the scale reliabilities observed with the two validation samples. Notably, two of the reliability coefficients for the Military B sample subscales were below .60, yet they were > .70 for all 8 item measures. We next fit a single-factor CFA model to each set of items which are also summarized in Table S6. The CFIs ranged from .893 to 1.00 whereas the SRMRs ranged from .003 to .066. The only instance that was deemed less than acceptable was the Connector experience subscale for the Military B sample. We then computed *orientation* and *experience* subscales, per TREO dimension, for further evaluation.

We again computed the two-factor, within-subjects Analysis of Variance (ANOVA) using the subscales scores as the criterion. The within-subjects factors were the (6) TREO *scales* crossed with the (2) orientation versus experience *foci*. Using the Military B data, we found significant main effects for both *scales* [$F(5,336) = 114.38$, $p < .001$; partial eta^2 = .63] and *foci* [$F(1,340) = 44.09$, $p < .001$; partial eta^2 = .12], as well as a *scale by foci* interaction [$F(5,336) = 79.56$, $p < .001$; partial eta^2 = .54]. The scales effect was attributable to participants reporting significantly ($p < .05$) higher *Doer* (M = 4.11) and *Team Builder* scores (M = 4.09), and lower *Organizer* (M = 3.63), and *Challenger* Scores (M = 3.62), as compared to the remaining scales. The foci effect was attributable to higher *orientation* (M = 3.86) than *experience* (M = 3.78) scores. The interaction stemmed from three of the scales as respondents reported significantly higher *orientation* than *experience* scores for the *Innovator* (M = 3.89 versus M = 3.59) and

Doer Scales (M = 4.26 versus M = 3.96), and relatively higher *experience* than *orientation Organizer* scores (M = 3.79 versus M = 3.46).

We computed a second 6X2 ANOVA using the Business Students A subscales scores as the criterion. We again found significant main effects for both *scales* [$F(5,510) = 70.76$, $p < .001$; partial $eta^2 = .40$] and *foci* [$F(1,514) = 82.40$, $p < .001$; partial $eta^2 = .14$], and a significant *scale by foci* interaction [$F(5,510) = 82.94$, $p < .001$; partial $eta^2 = .45$]. The scales effect was attributable to individuals reporting relatively higher *Doer* (M = 4.46) and *Team Builder* (M = 4.13), and relatively lower *Connector* (M = 3.94) and *Challenger* (M = 4.08) scores, as compared to the other scales. In general, higher scores were again seen for the *orientation* foci (M = 4.24), as compared to the *experience* foci (M = 4.14). Finally, the interaction stemmed from four scales as the *Organizer* (M = 4.29 versus M = 4.17) and *Team Builder* (M = 4.37 versus M = 4.26) scales had relatively higher *experience* than *orientation* scores. In contrast, the *Doer* (M = 4.57 versus M = 4.34) and *Innovator* (M = 4.40 versus M = 3.86) scales had relatively higher *orientation* than *experience* scores. In sum, these analyses again indicate that the TREO subscales are sufficiently reliable and unidimensional to use as indicators. Moreover, the mean patterns suggest that the two foci represent somewhat different facets of the underlying dimensions. These findings are supportive of using both subscales as indictors of the corresponding TREO latent variables.

Convergent and Discriminant Validity

Using the 12 subscales as indicators, we again fit a 6-factor TREO model to each sample. Table S4 lists the factor loadings on their respective latent variables, per sample. As shown, the CFA loadings ranged from .708 to .914 and all were significant p < .001. The 6-factor CFA model evidenced excellent fit indices for both the Military B [$\chi^2(39) = 122.61$, $p < .001$; CFI = .967; SRMR = .031] and for the Business Students B [$\chi^2(39) = 198.41$, $p < .001$; CFI = .965; SRMR = .030] samples.

We then added the Big 5 measures and fit an 11 factor model to each sample, again randomly pairing the four Big 5 items per dimension to form two parcels. The 11-factor CFA model evidenced acceptable fit indices for both the Military B [$\chi^2(154) = 387.22$, $p < .001$; CFI = .938; SRMR = .048] and for the Business Students B [$\chi^2(154) = 784.50$, $p < .001$; CFI = .900; SRMR = .059] samples. The correlations among the *latent variables* are reported in Table S7, with the Military B results in the lower left triangle, and the Business Students B findings in the upper right triangle.

The largest correlation among the TREO dimensions was $r = .88$, $p < .01$ between the *Team Builder* and *Connector* dimensions in the Military B sample. Collapsing their indicators to a single latent variable yielded a 10-factor CFA model that evidenced acceptable overall fit indices [$\chi^2(155) = 411.05$, $p < .001$; CFI = .932; SRMR = .048], but was significantly worse [$\Delta\chi^2(1) = 23.83$, $p < .001$] than the 11-factor model. These results indicated that, while highly correlated, the *Team Builder* and *Connector* TREO dimensions are empirically distinguishable. Keeping our focus on the Military B results, the TREO *Team Builder* dimension correlated significantly with all of the Big 5 dimensions (*r*s = .35 to .54, $p < .01$). The TREO *Connector* dimension correlated significantly with both the *Extraversion* and *Agreeableness* dimensions (*r*s

= .45 and .31, $p < .001$), respectively. And both the TREO *Organizer* and *Doer* dimensions correlated significantly with the Big 5 *Extraversion* (*r*s = .40 and .34, $p < .001$), and *Conscientiousness* (*r*s = .34 and .36, $p < .001$), dimensions, respectively. *Extraversion* also correlated significantly with TREO *Challenger* ($r = .34$, $p < .001$), and *Innovator* ($r = .41$, $p < .001$), dimensions. The largest correlation between a TREO and a Big 5 dimension was the *Team Builder – Agreeableness* pair ($r = .54$, $p < .001$). Collapsing their indicators yielded a 10-factor CFA model with acceptable overall fit indices [$\chi^2(155) = 472.28$, $p < .001$; CFI = .916; SRMR = .058], but it was significantly worse [$\Delta\chi^2(1) = 85.06$, $p < .001$] than the 11-factor model. These results indicated that the TREO dimensions are empirically distinguishable from the Big 5 dimensions.

Turning to the Business Students B sample, the largest correlation among the TREO dimensions was $r = .87$, $p < .01$, between the *Doer* and *Organizer* dimensions. The 10-factor CFA model collapsing their indicators [$\chi^2(155) = 844.64$, $p < .001$; CFI = .889; SRMR = .060] yielded an unacceptable model that was significantly worse [$\Delta\chi^2(1) = 60.14$, $p < .001$] than the 11-factor model. Examining the Business Students B correlations, the TREO *Team Builder* dimension again correlated significantly with all of the Big 5 dimensions (*rs* = .30 to .59, $p < .01$), and the TREO *Connector* dimension correlated significant with both *Extraversion* and *Agreeableness* (*rs* = .44 and .29, $p < .01$), respectively. Again both the TREO *Organizer* and *Doer* dimensions correlated significantly with *Conscientiousness* (*rs* = .58 and .57, $p < .001$), respectively. *Openness to Experience* correlated significantly with TREO *Challenger* ($r = .39$, $p < .001$), and *Innovator* ($r = .53$, $p < .001$), dimensions. The largest correlation between a TREO and Big 5 dimension was the *Team Builder – Agreeableness* combination ($r = .59$, $p < .001$). Accordingly, the 10-factor model combining their indicators produced an unacceptable [$\chi^2(155) = 824.72$, $p < .001$; CFI = .893; SRMR = .062] model that was significantly worse [$\Delta\chi^2(1) = 40.22$, $p < .001$] than the 11-factor model.

In sum, the findings from the Military B and Business Students B validation samples paralleled those of the two development samples. Specifically, the TREO subscales were generally reliable, exhibited significant patterned differences, mapped to six distinguishable latent variables, and evidenced meaningful correlations with, but were distinguishable from, measures of the Big 5 personality dimensions. As shown in Table S8, we used meta-analyses to summarize the nomological correlations between the TREO and Big 5 dimensions across the four samples chronicled thus far. Specifically, using the latent correlations (which have accounted for measurement unreliability) reported in Tables S5 and S7, we performed "bare bones" meta-analyses which account for differences across the four sample sizes to derive an estimated population correlation (ρ) and variability for each TREO - Big 5 pairing (Hunter & Schmidt, 2004). For these analyses, we maintain that hypothesized relationships should not only be significant (i.e., their 80% credibility interval not include zero), but their ρ should at least of medium value (i.e., > .20) to be not be rejected.

Reviewing the meta-analytic results, *Organizer* correlated significantly with *Conscientiousness* (ρ = .43) but not with *Emotional Stability* (ρ = .14). Moreover, the *Organizer* dimension exhibited positive correlations with Agreeableness (ρ = .22) and Extraversion (ρ = .31). As expected, *Doer* correlated positively with Conscientiousness (ρ = .45), and Emotional Stability (ρ = .25) but not with Agreeableness (ρ = .19). It also demonstrated a positive

correlation with Extraversion (ρ = .26). Consistent with our hypotheses, *Team Builder* correlated positively with Agreeableness, Extraversion, Emotional Stability, and Openness (ρs = .52, .40, .29, .29, respectively), as well as with Conscientiousness (ρ = .32). *Connector* illustrated positive correlations with *Agreeableness* (ρ = .28), *Extraversion* (ρ = .44), and *Openness* (ρ = .26), as hypothesized. As further hypothesized, *Innovator* correlated positively with Extraversion (ρ = .40), and with Openness (ρ = .45), but not with *Conscientiousness* (ρ = .14). Finally, *Challenger* correlated positively with *Extraversion* (ρ = .28), and *Openness* (ρ = .30), but the anticipated negative correlation with *Agreeableness* (ρ = -.00), was not evident. In total, 78% of the anticipated correlations were significant and > .20.

There were five unanticipated significant correlations. Interestingly, *Conscientiousness* correlated with all of the TREO dimensions except *Innovator* and *Challenger*. *Organizers* had more pervasive correlations than anticipated which accounted for two of the other non-hypothesized effects. Clearly the TREO dimensions were distinguishable from the Big 5 dimensions, as only one meta-analytic correlation reached the .50s (*Team Builder – Agreeableness*, ρ = .52). On balance, we consider these findings to be consistent with the anticipated nature of the nomological network with the Big 5 personality measures. Nevertheless, evidence of the stability of the TREO measures over time is warranted, which we consider next.

TREO STABILITY

To be valuable as a measure of individual differences, it is important to assess the temporal stability of TREO scores. The purpose of this sample and analyses was to evaluate TREO test-retest stability coefficients.

Sample

Data were collected from 172 business students enrolled in nine class sections of a Strategic Management course in a small northeastern university. On average, participants were 21.6 years old, 91% Caucasian, and 48% were women. Their average academic grade point average was 3.06 (SD = .38). The representations of majors were: 25% Accounting, 20% Finance, 31% Marketing, 21% Management, and 3% Economics. Students work in teams on a number of assignments (e.g., case study reports, presentations) over the last 10 weeks of the semester. We collected paper TREO surveys before and after the group activities, and are therefore able to examine the stability of the measures over approximately a two and a half month period. All scales exhibited acceptable internal consistencies, which are summarized in Table S9.

Results

Table S9 presents TREO scale descriptive statistics and correlations within and across time. We computed pair t-tests to examine whether any significant mean changes occurred over time, and none were evident: (*Organizer*: $t(172) = -.42$, *ns*; Doer: $t(172) = .72$, *ns*; *Challenger*: $t(172) = .30$, *ns*; *Innovator:* $t(172) = .47$, *ns*; *Team Builder*: $t(172) = .02$, *ns*; and *Connector*: $t(172) = -.64$, *ns*). Notably the correlations within-time across-dimensions (i.e., triangles in Table S9) were high and averaged $r = .64$, $p < .001$ at time 1 and $r = .69$, $p < .001$ at time 2. As highlighted in the small boxes in Table S9, the time 1 – time 2 stability coefficients were

significant and averaged $r = .52$, $p < .001$, which are within the range observed for measures of personality variables (Ardelt, 2000). Moreover, the across-time across-dimension correlations were markedly lower, with the test-retest correlations exhibiting the highest correlations within both the rows and columns of that portion of the matrix (i.e., large square in Table S9). These results are encouraging as they demonstrate that the TREO measures did not show any significant mean changes over time, while also exhibiting significant stability coefficients over time.

Supplement Footnotes

[1] The original item set and details concerning these item revisions are available from the authors.

Supplement References

Anderson, J. C. & Gerbing, D. W. (1991). Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. *Journal of Applied Psychology, 76,* 732-740.

Ardelt, M. (2000). Still stable after all these years? Personality stability theory revisited. Sopciological Social Psychology, 63, 4, 392-405.

Barrick, M. R. & Mount, M . K. (1991). The big five personality dimensions and job performance: A meta analysis. *Personnel Psychology, 44,* 1-26.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*(107), 238-246.

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment, 18*(2), 192-203.

Haynes, S. N., Richard, D. R., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*, 238-247.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings.* (2 ed.). Thousand Oaks, CA: Sage.

KMI Media Group. (2008). Transition team's role in Iraq. *Military Training Technology, 12*(1). Retrieved from http://www.kmimediagroup.com/mt2-home/75-mt2-2008-volume-12-issue-1/1466-transition-teams-role-in-iraq.pdf

Landis, R. S., Beal, D. J., & Tesluk, P. E. (2000). A comparison of approaches to forming composites measures in structural equations models. *Organizational Research Methods, 3*, 186-207.

McKenzie, J. F., Wood, M. L., & Kotecki, J. E. (1999). Establishing content validity: Using qualitative and quantitative steps. *American Journal of Health Behavior, 23*, 311-318.

Mathieu, J. E., & Taylor, S. R. (2006). Clarifying conditions and decision points for mediational type inferences in Organizational Behavior. *Journal of Organizational Behavior, 27*, 1031-1056.

Muthen, L. K., & Muthen, B. O. (2012). *Mplus user's guide*. Los Angeles, CA: Muthen & Muthen.

Rubio, D. M., Berg-Weger, M., Tebb, S. T., Lee, E. S., & Rauch, S. (2003). Objectifying content validity study in social work research *Social Work Research, 27*(2), 94-104.

Stewart, G. L., Fulmer, I. S., & Barrick, M. R. (2005). An exploration of member roles as a multilevel linking mechanism for individual traits and team outcomes. *Personnel Psychology, 58*(2), 343-365.

Table S1. Hypothesized Correlational Pattern between TREO Dimensions and Big 5 Personality Dimensions

Big 5 Dimensions	***TREO Dimensions***					
	Organizer	Doer	Team Builder	Connecter	Innovator	Challenger
Conscientiousness	+	+			+	
Agreeableness		+	+	+		-
Extraversion			+	+	+	+
Emotional Stability	+	+	+			
Openness to Experience			+	+	+	+

Table S2. Big 5 Personality Dimensions and Definitions

Personality traits refer to stable psychological characteristics of individuals' that guide their interactions with others and the environment. Although numerous taxonomies have been advanced, the Big 5 framework is among the most popular (Barrick & Mount, 1991). It includes the following 5 dimensions.

Big 5 Dimension	Definition
Conscientiousness:	includes being dependable, careful, thorough, responsible, organized, and planful. Also includes hardworking, achievement-oriented, and preserving.
Extraversion:	includes being sociable, gregarious, assertive, talkative, and active. Also includes initiative, ambition, impetuous, sociable, exhibitionist, and expressive.
Agreeableness:	includes being courteous, flexible, trusting, good-natured, cooperative, forgiving, soft-hearted, and tolerant.
Emotional Stability:	includes being not anxious, depressed, angry, embarrassed, emotional, worried or insecure. Also includes not being moody, irritable, or sad.
Openness to Experience:	includes being imaginative, cultured, curious, original, broad-minded, intelligent, and artistically sensitive.

Table S3. Subscale Reliabilities and Fit Indices per Development Samples

	Military A (N = 317)			***Business Students A*** (N = 266)		
Scale	α	CFI	SRMR	α	CFI	SMMR
Organizer	.85			.85		
Orientation	.74	1.00	.001	.73	.900	.053
Experience	.79	.966	.029	.73	1.00	.012
Doer	.82			.83		
Orientation	.70	1.00	.008	.78	.997	.015
Experience	.65	.998	.017	.65	1.00	.004
Challenger	.81			.88		
Orientation	.63	.976	.023	.82	.971	.029
Experience	.72	1.00	.013	.77	.952	.035
Innovator	.82			.83		
Orientation	.71	.995	.021	.77	1.00	.011
Experience	.70	1.00	.002	.71	.996	.017
Team Builder	.85			.88		
Orientation	.66	.930	.036	.82	1.00	.015
Experience	.78	.974	.027	.76	.958	.033
Connector	.87[1]			.85		
Orientation	.83[1]	.987	.023	.71	1.00	.005
Experience	.75	.887	.067	.77	.994	.017

CFI = Comparative Fit Index, SRMR = Standardized Root Mean Squared Residual

[1] Missing Item 23 due to a typographical error on survey

Table S4. Subscale Corresponding 6-Factor Standardized Loadings per Sample

Scale	***Military A***	***Business Students A***	***Military B***	***Business Students B***
Organizer				
Orientation	.867	.831	.851	.866
Experience	.855	.894	.863	.911
Doer				
Orientation	.828	.850	.756	.837
Experience	.835	.865	.827	.869
Challenger				
Orientation	.826	.841	.726	.860
Experience	.813	.897	.708	.889
Innovator				
Orientation	.771	.836	.819	.866
Experience	.888	.757	.815	.763
Team Builder				
Orientation	.867	.817	.800	.841
Experience	.871	.875	.858	.914
Connector				
Orientation	.823[1]	.864	.877	.900
Experience	.850	.851	.792	.862
N	317	266	341	515

[1] Missing Item 23 due to a typographical error on survey. All loadings $p < .001$

Table S5. Correlations between TREO and Big 5 Latent Variables per Development Sample

Dimensions	1	2	3	4	5	6	7	8	9	10	11
1. Organizer	---	.86	.56	.59	.65	.61	.24	.09	.59	.10	-.13
2. Doer	.77	---	.53	.65	.61	.54	.18	.07	.63	.19	.05
3. Challenger	.80	.63	---	.87	.59	.64	.18	-.21	.10	.18	.30
4. Innovator	.86	.70	.91	---	.62	.79	.35	-.09	.16	.20	.54
5. Team Builder	.76	.72	.65	.78	---	.69	.36	.36	.30	.29	.10
6. Connector	.87	.76	.72	.85	.85	---	.40	.05	.12	.03	.16
7. Extraversion	.35	.28	.28	.39	.41	.46	---	.27	.12	.21	.21
8. Agreeableness	.25	.16	.17	.25	.50	.40	.60	---	.30	-.07	.22
9. Conscientiousness	.13	.22	.09	.00	.18	.11	.45	.51	---	.17	-.01
10. Emotional Stability	.13	.20	.06	.07	.19	.17	.40	.59	.85	---	.13
11. Openness to Experience	.19	.16	.26	.36	.36	.34	.62	.82	.55	.59	---

Note. Military A sample (N = 317) correlations appear in lower left triangle: $r \geq |.11|$, $p < .05$; $r \geq |.15|$, $p < .01$

Business Students A sample (N = 266) correlations appear in lower left triangle: $r \geq |.12|$, $p < .05$; $r \geq |.16|$, $p < .01$

Table S6. Subscale Reliabilities and Fit Indices per Validation Samples

	Military B (N = 341)			***Business Students B*** (N = 515)		
Scale	α	CFI	SRMR	α	CFI	SRMR
Organizer	.84			.85		
Orientation	.70	.942	.037	.76	.944	.038
Experience	.77	.934	.049	.77	1.00	.004
Doer	.78			.85		
Orientation	.66	.992	.019	.79	.998	.012
Experience	.59	1.00	.014	.68	1.00	.003
Challenger	.74			.88		
Orientation	.57	.982	.022	.82	.974	.025
Experience	.65	.986	.023	.75	.950	.034
Innovator	.78			.84		
Orientation	.69	.981	.026	.76	.988	.020
Experience	.60	.989	.022	.73	.983	.022
Team Builder	.81			.85		
Orientation	.65	.985	.025	.67	.995	.018
Experience	.71	.952	.034	.78	.982	.022
Connector	.83			.87		
Orientation	.72	1.00	.012	.74	.999	.013
Experience	.71	.893	.066	.79	.993	.015

CFI = Comparative Fit Index, SRMR = Standardized Root Mean Squared Residual

Table S7. Correlations between TREO and Big 5 Latent Variables per Validation Sample

Dimensions	1	2	3	4	5	6	7	8	9	10	11
1. Organizer	---	.87	.60	.66	.68	.66	.27	.30	.58	.15	.10
2. Doer	.82	---	.65	.75	.69	.64	.23	.32	.57	.31	.19
3. Challenger	.80	.70	---	.84	.63	.63	.28	.02	.19	.15	.39
4. Innovator	.79	.76	.86	---	.72	.76	.43	.19	.22	.23	.53
5. Team Builder	.74	.80	.61	.73	---	.73	.43	.59	.36	.30	.30
6. Connector	.88	.75	.76	.82	.88	---	.44	.29	.25	.11	.23
7. Extraversion	.40	.34	.36	.41	.37	.45	---	.28	.05	.06	.09
8. Agreeableness	.15	.10	-.03	.07	.54	.31	.28	---	.39	-.11	.25
9. Conscientiousness	.34	.36	.08	.12	.41	.28	.15	.50	---	-.03	-.08
10. Emotional Stability	.15	.27	.06	.09	.35	.24	.19	.36	.56	---	.22
11. Openness to Experience	.18	.23	.21	.34	.36	.29	.20	.43	.59	.46	---

Note. Military B sample (N = 341) correlations appear in lower left triangle: $r \geq |.11|$, $p < .05$; $r \geq |.15|$, $p < .01$

Business Students B sample (N = 515) correlations appear in lower left triangle: $r \geq |.09|$, $p < .05$; $r \geq |.12|$, $p < .01$

Table S8. Meta-Analytic Summaries of TREO – Big 5 Correlations Across Four Samples

Big 5 Dimensions	***TREO Dimensions***					
	Organizer	Doer	Team Builder	Connecter	Innovator	Challenger
Conscientiousness	ρ= .43, SD=.18 .20 - .66 5%	ρ= .45, SD=.15 .26 - .65 7%	ρ= .32, SD=.07 .23 - .41 33%	ρ= .20, SD=.05 .13 - .27 48%	ρ= .14, SD=.06 .03 - .23 25%	ρ= .13, SD=.000 .13 - .13 100%
Agreeableness	ρ= .22, SD=.07 .13 - .30 38%	ρ= .19, SD=.09 .07 - .30 24%	ρ= .52, SD=.07 .42 - .61 23%	ρ= .28, SD=.10 .14 - .41 18%	ρ= .12 , SD=.11 -.01 - .26 19%	ρ= -.00,SD=.10 -.14 - .14 19%
Extraversion	ρ= .31, SD=.04 .26 - .36 60%	ρ= .26, SD=.03 .23 - .29 78%	ρ= .40, SD=.00 .40 - .40 100%	ρ= .44, SD=.00 .44 - .44 100%	ρ= .40, SD=.00 .40 - .40 100%	ρ= .28, SD=.03 .24 - .32 70%
Emotional Stability	ρ= .14, SD=.00 .14 - .14 100%	ρ= .25, SD=.01 .24 - .27 94%	ρ= .29, SD=.03 .25 - .32 76%	ρ= .14, SD=.05 .08 - .20 52%	ρ= .16, SD=.05 .09 - .22 53%	ρ= .11, SD=.00 .11 - .11 100%
Openness to Experience	ρ= .14, SD=.00 .14 - .14 100%	ρ= .17, SD=.03 .13 - .21 72%	ρ= .29, SD=.08 .19 - .40 26%	ρ= .26, SD=.04 .21 - .30 65%	ρ= .45, SD=.08 .34 - .55 21%	ρ= .30, SD=.05 .23 - .37 45%

Note. Correlations cumulated across the two development (Ns = 317 and 266) and two validation (Ns = 341 and 515) samples.
P = estimated population correlations, SD = estimated standard deviation of population correlations. Lower values per cell are the 80% credibility interval and the percentage of between sample correlation variance accounted for by sampling error.

Table S9. TREO Test-Retest Correlations

Scales	1	2	3	4	5	6	7	8	9	10	11	12
Time 1												
1. Organizer	(.82)											
2. Doer	.75	(.77)										
3. Challenger	.59	.59	(.69)									
4. Innovator	.58	.61	.64	(.83)								
5. Team Builder	.65	.62	.65	.71	(.84)							
6. Connector	.68	.54	.61	.61	.73	(.82)						
Time 2												
7. Organizer	.53	.43	.29	.33	.38	.31	(.89)					
8. Doer	.38	.54	.32	.39	.36	.23	.76	(.84)				
9. Challenger	.28	.32	.51	.36	.36	.31	.64	.59	(.84)			
10. Innovator	.22	.29	.34	.50	.39	.33	.65	.72	.67	(.87)		
11. Team Builder	.37	.39	.38	.41	.55	.36	.71	.68	.66	.73	(.87)	
12. Connector	.41	.37	.40	.41	.45	.46	.76	.65	.64	.72	.75	(.82)
Mean	3.91	4.15	3.78	4.00	3.93	3.74	3.93	4.12	3.77	3.98	3.93	3.78
SD	.62	.54	.56	.58	.65	.71	.69	.61	.70	.62	.69	.69

Notes. N = 172 Values on diagonal are scale alphas. Correlations > |.15|, $p < .05$; > |.20|, $p < .05$.