The Cognitive Consequences of Using Categorical versus Dimensional Classification Systems: The Case of Personality Disorder Experts

Benjamin M. Rottman (benjamin.rottman@yale.edu)
Dept. of Psychology, Yale U., 2 Hillhouse Ave., New Haven, CT 06520

Nancy S. Kim (n.kim@neu.edu)
Dept. of Psychology, Northeastern U., 125 Nightingale Hall, 360 Huntington Ave., Boston, MA 02115

Woo-kyoung Ahn (woo-kyoung.ahn@yale.edu)
Dept. of Psychology, Yale U., 2 Hillhouse Ave., New Haven, CT 06520

Charles A. Sanislow (charles.sanislow@yale.edu)
Dept. of Psychiatry, Yale U., 301 Cedar Street, 2nd Fl., New Haven, CT 06520

Abstract

Features are inherently ambiguous in that their meanings depend on the categories they describe (e.g., *small* for planets vs. molecules; Murphy, 1988). However, a new proposal for the next version of the DSM (*DSM-IV-TR*, Diagnostic and Statistical Manual of Mental Disorders, 4th Ed., text revision; American Psychiatric Association, 2000) advocates eliminating personality disorder categories, instead describing patients using only dimensions with the well-known Five-Factor Model. We investigated whether experts in personality pathology are able to translate dimensional patient descriptions into their corresponding diagnostic categories in the current version of the DSM. The results showed that even experts had considerable difficulty disambiguating the meaning of the dimensions to determine correct diagnoses and found the utility of the dimensional system to be lacking. Implications for categorization research are discussed.

Keywords: Concepts; Categorization; Expertise; DSM; FFM

Introduction

Categories offer many cognitive benefits (e.g., Murphy, 2002). They facilitate communication and cognitive economy; instead of "large, gray, with a trunk," one can simply say or remember "elephant." Importantly, they also *instantiate* the meaning of *features* (e.g., Kamp, 1975; Murphy & Medin, 1985; Rosch, 1978). For example, birds' *wings* are very different from airplanes' *wings,* and the *legs* of a zebra are very different from those of a desk. The importance of *feature instantiation* has been empirically demonstrated in various domains. Murphy (1988) demonstrated that an adjective can assume multiple different meanings when paired with different common nouns (e.g., *open hand* means *dealt face up* while *open eye* means *alert*). Brooks and Hannah (2006) found that people interpret ambiguous perceptual features (i.e., diagnostic of two categories they learned) by using another, unambiguously diagnostic feature also present in the exemplar. Yet despite such findings, there is a recent proposal in personality psychopathology that neglects the issue of feature instantiation. The main goal of the current study is to (i) provide basic research on whether the cognitive advantages of feature instantiation are less important for people with a high degree of domain expertise, and (ii) offer translational research making practical suggestions for mental disorder nosology.

Classification Systems of Personality Pathology

The current categorization system of personality pathology, the DSM-IV-TR classifies maladaptive personality into 10 discrete personality disorders (categories), each defined by unique criteria (features). For example, to be diagnosed with Antisocial Personality Disorder (PD), one must have, pervasively and across contexts, at least 3 of the 7 symptoms shown in Figure 1a.

There is currently an ongoing debate about whether the DSM should be dimensionalized rather than using categories. In particular, a proposal for personality pathology that has received much recent attention is based upon the Five Factor Model of personality (FFM; e.g., Costa & McCrae, 1992; see Clark, 2007 for a review). The FFM proposal advocates abandoning discrete personality disorders, instead describing people using ratings along 30 traits or facets grouped into 5 factors. Figure 1b shows an FFM profile of a prototypic patient with Antisocial PD.

Ambiguity and Utility of the FFM

The FFM is a promising candidate for the DSM-V because it has been shown to be biologically-based, universal, temporally stable, and can avoid problems with the DSM-IV Axis II categories including high comorbidity and arbitrary diagnostic thresholds (Widiger & Trull, 2007). Before adopting any new system, however, it is important to assess whether it can fulfill various clinical functions, including making treatment plans and prognoses, communicating with patients or other clinicians, and describing a patient's global personality or important personality problems (First et al., 2004; First, 2005). Such *clinical utilities* discussed in the psychopathology literature are strikingly similar to the functions of categories discussed in the cognitive literature (e.g., inductive inferences, communication; Murphy, 2002). Accordingly, we contend that abandoning categories, as

1a

1. Failure to conform to social norms with respect to lawful behaviors as indicated by repeatedly performing acts that are grounds for arrest.
2. Deceitfulness, as indicated by repeated lying, use of aliases, or conning others for personal profit or pleasure.
3. Impulsivity or failure to plan ahead.
4. Irritability and aggressiveness, as indicated by repeated physical fights or assaults.
5. Reckless disregard for safety of self or others.
6. Consistent irresponsibility, as indicated by repeated failure to sustain consistent work behavior or honor financial obligations.
7. Lack of remorse, as indicated by being indifferent to or rationalizing having hurt, mistreated, or stolen from another.

1b

Neuroticism		Extraversion		Openness		Agreeableness		Conscientiousness	
Anxiousness	2.00	Warmth	2.00	Fantasy	3.48	Trust	1.70	Competence	2.52
Angry Hostility	3.93	Gregariousness	3.48	Aesthetics	2.78	Straightforwardness	1.41	Order	2.74
Depressiveness	2.70	Assertiveness	4.07	Feelings	2.41	Altruism	1.41	Dutifulness	1.52
Self-consciousness	1.63	Activity	4.00	Actions	4.07	Compliance	1.81	Achievement Striving	2.33
Impulsivity	4.22	Excitement-Seeking	4.30	Ideas	3.26	Modesty	1.70	Self-Discipline	1.85
Vulnerability	2.07	Positive Emotions	3.52	Values	3.48	Tendermindedness	1.52	Deliberation	1.96

Figure 1: A sample description of a patient with Antisocial Personality Disorder using (a) the DSM-IV-TR symptoms and (b) the Five-Factor Model of Personality facet scores from Samuel and Widiger (2004).

suggested by the FFM proposal, will lead to cognitive difficulties. In this section, we outline some key problems.

Because it lacks categories, the FFM may be unable to instantiate the meanings of features; that is, each of the 30 traits may assume different meanings in different contexts (Mischel & Peake, 1982; Mischel & Shoda, 1995; Westen, 2006). For instance, a clinician might interpret a low score on the 'Gregariousness' facet to mean paranoid fears (as in Paranoid Personality Disorder), fear of not being liked by others (Avoidant PD), or indifference to others (Schizoid PD). A high score on 'Anger' can mean temper tantrums (Histrionic PD) or a lack of control over anger (Borderline PD; Benjamin, 1993). Thus, an FFM description alone, without the context of categories, may be ambiguous.

This ambiguity can pose problems not only for making diagnoses, but also for determining prognoses, developing treatment plans, and carrying out other clinical functions. For instance, the FFM's lack of categories may impair its utility for communicating about and remembering patients; saying that a patient has "paranoid personality disorder" is considerably easier than discussing 30 facet scores. This problem would also be exacerbated by the fact that users would have to integrate information across the 30 facets to form a coherent image of the patient. Taken together, the FFM's vague descriptors, lack of categories, and many dimensions could impede reasoning about patients.

Previous studies comparing the clinical utility of the FFM with that of the DSM have yielded mixed results, with some supporting the FFM (e.g., Samuel & Widiger, 2006) and others not (e.g., Spitzer et al., 2007; Sprock, 2003). However, in all of those studies, the utility of the FFM was assessed within the context of a patient vignette or personal knowledge about a patient, potentially disambiguating the descriptors. For example, if a clinician knows that her patient is avoidant, she can rate the patient low on gregariousness for a specific reason (e.g., fear of not being liked by others) and, as a result, may judge the FFM as useful within the context of the additional information. Since the case vignette could disambiguate the FFM's facets and also offer diagnostic information (e.g., Avoidant PD), which could further disambiguate the facets, the tasks in these previous studies were not capable of revealing any potential ambiguity in FFM patient descriptions.

In a recent study (Rottman, Ahn, Sanislow, & Kim, 2009) we assessed the FFM's ambiguity, using a method analogous to the back-translation paradigm commonly used in cross-linguistic studies. For example, if a stimulus in English is translated into Japanese, to ensure that the Japanese version contains the intended content, a separate translator back-translates the Japanese version into English and compares the back-translation with the original. In previous work, other researchers translated prototypic DSM patient descriptions into FFM profiles (Samuel & Widiger, 2004; 2006). Rottman et al. (2009) therefore tested whether clinicians could accurately back-translate Samuel and Widiger's (2004; 2006) FFM descriptions into DSM diagnoses. If FFM patient profiles are inherently ambiguous, then clinicians should have difficulty providing the correct corresponding DSM diagnosis, as one FFM profile could be mapped onto multiple DSM diagnoses.

Indeed, we found that without categorical information or vignettes (e.g., Figure 1b), practicing clinical psychologists, psychiatrists, and social workers had difficulty recognizing DSM diagnoses, although they showed little difficulty when the same disorders were presented in the DSM format (e.g., Figure 1a). This suggests that when DSM diagnoses are translated into FFM profiles, critical information is lost. The clinicians also rated the FFM as having less clinical utility than the DSM. In sum, the dimensional FFM appears to be difficult to use alone, without the instantiating effects of accompanying categories or vignettes.

FFM and Expertise

Our previous study (Rottman et al., 2009) examined practicing clinical psychologists, psychiatrists, and social workers, who were not necessarily personality disorder experts. It is possible that domain experts may overcome the above-mentioned challenges of working with the FFM.

For example, experts likely have more knowledge and theories about the causal workings of personality disorders, which could make it easier for them to identify important correlations between facets (see Ahn, Marsh, & Luhmann, 2002; Wattenmaker et al., 1986 for such demonstrations in the cognitive literature). As a result, experts may integrate information across the 30 facets to form a more coherent concept of a patient, resulting in better recognition of DSM diagnoses. For instance, although a low score on the 'Gregariousness' facet may be ambiguous on its own, a combination of low 'Gregariousness' and low 'Trust' scores may indicate that a patient has Paranoid PD, whereas a combination of low 'Gregariousness' and high 'Self-Consciousness' scores may indicate that a patient has Avoidant PD. (See Chase & Simon, 1973 for a related chunking phenomenon with chess experts.) The current study tests whether personality disorder experts can better recognize the DSM diagnoses and find the FFM to be more clinically useful, presumably because they can use configurations of facet scores to disambiguate the meaning of patient profiles.

Study

Methods

Participants As in Lynam and Widiger (2001), we operationalized expertise in personality pathology as having published research on personality disorders. Specifically, we searched the PsycInfo database for authors who had published at least three papers with the keyword "Personality Disorder" in peer-reviewed journals, and who had published at least one article from January 2006 through mid-November 2008 (the time of the search). We then excluded those for whom we could not find contact information and those who were highly likely to be familiar with Rottman et al. (2009). Recruitment emails were sent to 476 researchers in December 2008. At the beginning of the study, we requested that participants verify that they consider personality disorders to be among their primary research interests and that they have been conducting research on personality disorders for at least four years. This allowed us to exclude those who collaborated on personality disorder papers only because of expertise in other fields (e.g., statisticians). Seventy-three participants completed the experiment. The experiment took 29 minutes on average, and participants were compensated with either a $60 gift certificate to an online retailer or a $60 check.

Materials and Design Twelve different cases were described in both the FFM and DSM styles. Ten of these cases described prototypic patients for the 10 DSM-IV personality disorders. The remaining two were comorbid cases with two personality disorders each, as comorbid cases have been argued to be more representative of real-world patients (e.g., Bornstein, 1998). The FFM scores were taken from previous studies (Samuel & Widiger, 2004 for the prototypic cases; Samuel & Widiger, 2006 for the comorbid cases), in which practicing clinicians thought about these cases and rated each on the 30 FFM facets. When presenting the cases to participants, we provided both the average rating for each facet and a plot of the facet scores, anchored by high (e.g., "fearful, apprehensive" for anxiousness) and low (e.g., "relaxed, unconcerned, cool" for anxiousness) adjectives (Rottman et al., 2009).

For the DSM condition, each prototypic case comprised all the DSM-IV-TR diagnostic criteria for that personality disorder (e.g., Figure 1a). The comorbid DSM case descriptions were taken from Rottman et al. (2009), in which clinicians in a pretest identified all the DSM-IV-TR personality disorder symptoms they found to be present in the comorbid vignettes (Samuel & Widiger, 2006).

The 12 cases were divided into two groups, each containing five prototypic cases and one comorbid case. For diversity, each group included at least one disorder from the three clusters of personality disorders in the DSM-IV, and the diagnoses of the comorbid case did not match the diagnoses of any of the prototypic cases in the group. To the extent that it was possible, we also matched the two groups of prototypic cases for difficulty of diagnosis, as previously determined in Rottman et al. (2009).

Each participant saw one group of six cases presented in the FFM style and the other group in the DSM style. Thus, descriptive style (DSM vs. FFM) was a within-subject variable. The pairing of cases with descriptive style, presentation order of the two groups, and order of the styles were counterbalanced across participants. The order of the six cases within each group was randomized.

Procedure The study was performed online using Qualtrics software. Participants were told that they would be presented with descriptions of adult patients and were asked to imagine that these patients were referred to them along with a patient description from a previous consultation. Participants were told that the patients "do not have schizophrenia or any other psychotic disorder, and their symptoms do not occur due to the direct effect of any general medical condition." This instruction was included to prevent participants from avoiding giving personality disorder diagnoses for reasons not of experimental interest (e.g., a schizoid personality disorder diagnosis is not allowed if it occurs exclusively during the course of schizophrenia). Finally, participants were instructed not to consult the DSM or other references during the experiment.

Next, participants were presented with the first group of six cases in the DSM or FFM style. After each individual case, participants were asked to "provide any DSM-IV diagnoses you believe this patient to have." Participants rated their confidence in each diagnosis on a seven-point scale (where 1 = "not confident at all," 4 = "somewhat confident," and 7 = "very confident").

After the first group of cases was presented, participants rated the utility of the descriptive system that they just saw by answering the following six questions on a five-point

scale (1 = "not at all," 2 = "slightly," 3 = "moderately," 4 = "very," 5 = "extremely"):

1) "How informative is this description in making a prognosis for this person?"
2) "How informative is this description in devising treatment plans for this person?"
3) "How useful do you feel the system used to describe this person would be for communicating information about this individual with other mental health professionals?"
4) "How useful do you feel the system used to describe this person would be for communicating information about the individual to him or herself?"
5) "How useful is the system used to describe this person for comprehensively describing all the important personality problems this individual has?"
6) "How useful was the system used to describe this person for describing the individual's global personality?"

Participants then performed the same series of tasks for the second group of cases. Finally, participants provided demographic information and rated their own familiarity with the diagnostic systems (1 = "not at all familiar," 4 = "moderately familiar," 7 = "extremely familiar").

Results[1]

Demographics Fifty-one Ph.D.'s, 16 M.D.'s, 3 M.A.'s, 2 M.D./Ph.D.'s, and 1 M.S.W. participated. Participants had published a median of 15 papers on personality disorders (Mean=24, Range=[3,160]). Of these, 66% were also practicing clinicians. Because the DSM-IV is the current diagnostic system, participants were more familiar with the DSM than the FFM, $t(72)=7.70$, $p<.01$. The current participants were more familiar with the DSM ($M=6.40$, $SD=.95$) and, more importantly, with the FFM ($M=4.97$, $SD=1.66$) than the clinicians in Rottman et al. (2009; $M=5.68$, $SD=1.26$, for the DSM; $M=2.17$, $SD=1.65$, for the FFM, $t(174.26)=4.89$, $p<.01$,[2] for the DSM; $t(252)=12.24$, $p<.01$, for the FFM).

Diagnoses For each prototypic case, participants almost always gave the correct diagnosis in the DSM condition ($M=.99$ of cases, $SD=.06$) and were much more accurate than in the FFM ($M=.62$ of cases, $SD=.25$), $t(72)=12.36$, $p<.01$ (Figure 2a, Experts). These results closely replicate Rottman et al. (2009). Although the personality disorder experts in the current study provided more accurate diagnoses than the practicing clinicians in Rottman et al. (2009), they did so in both the DSM and FFM (Figure 2a).

For the comorbid cases, participants were also more likely to give correct diagnoses in the DSM ($M=.77$, $SD=.26$) than in the FFM condition ($M=.48$, $SD=.33$), $t(72)=6.32$, $p<.01$. Again, the direction and amount of the difference roughly replicated the results from the clinicians in Rottman et al. (2009; .60 in the DSM condition and .21 in the FFM).

We also examined incorrect diagnoses, defined as any DSM diagnosis mismatching the correct diagnosis, and any non-DSM-IV diagnosis. Participants gave significantly more incorrect diagnoses per case in the FFM ($M=.79$, $SD=.48$) than DSM condition ($M=.16$, $SD=.35$), $t(71)=9.82$, $p<.01$ (Figure 2b). For the comorbid cases, they also gave more incorrect diagnoses in the FFM ($M=.81$, $SD=.84$) than in the DSM condition ($M=.30$, $SD=.64$), $t(72)=4.52$, $p<.01$. Again, these results differed little from those of the practicing clinicians in Rottman et al. (2009).

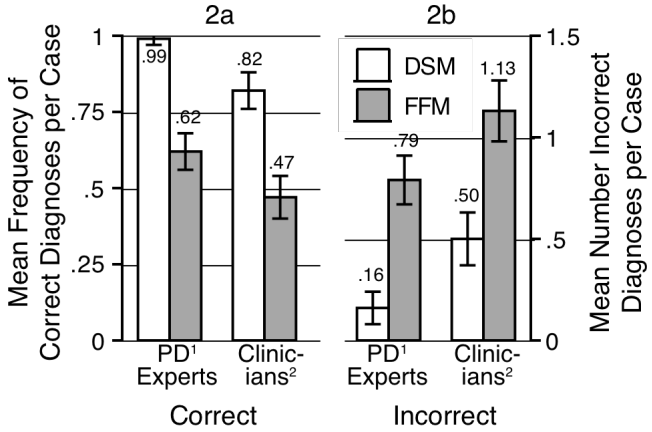

Figure 2: Correct and Incorrect Diagnoses for Prototypic Cases (95% Confidence Intervals)

Note. [1] PD = Personality Disorder. [2] Clinicians' data were taken from Rottman et al. (2009).

Confidence in Diagnoses The confidence ratings reflect participants' difficulty in providing accurate diagnoses (Figure 3; collapses prototypic and comorbid cases). A 2 (correct vs. incorrect diagnosis) × 2 (DSM vs. FFM) repeated-measures ANOVA[3] found that participants were more confident for correct than incorrect diagnoses, $F(1,25)=75.24$, $p<.01$, $\eta_p^2=.75$, and were more confident in the DSM than in the FFM condition, $F(1,25)=45.15$, $p<.01$, $\eta_p^2=.64$. In addition, there was a significant interaction, $F(1,25)=53.15$, $p<.01$, $\eta_p^2=.68$, because the experts were much more confident in correct than incorrect diagnoses for the DSM; participants were more aware of the accuracy of their diagnoses in the DSM condition.

Utility Ratings Paired t-tests revealed that participants found the DSM to be more useful than the FFM on three measures (prognosis, treatment plans, and communicating with professionals; (t's(69)>2.19, p's<.05; see Figure 4). Participants rated the FFM as more useful than the DSM for communicating with patients, $t(69)=3.03$, $p<.01$, possibly

[1] Where appropriate, comparisons are made between the current data from personality disorder experts and previous data (Rottman et al., 2009, Experiment 1) from clinicians not necessarily specializing in personality disorders. Due to differences in design, inferential statistics are not always possible.

[2] Equal variances not assumed.

[3] This analysis included only participants who gave at least one correct and incorrect diagnosis in both the DSM and FFM.

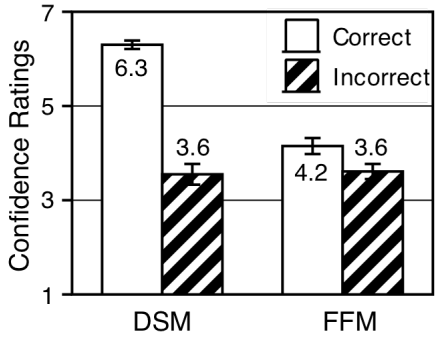

Figure 3: Confidence Ratings (Std. Error)

because the DSM disorder names are considered to be stigmatizing and because the FFM facets are common terms rather than technical disorder names. There was no difference for global personality description and comprehensively describing all important personality problems, *p's >.10*. All of these patterns of results also hold when only including data from the condition presented first.

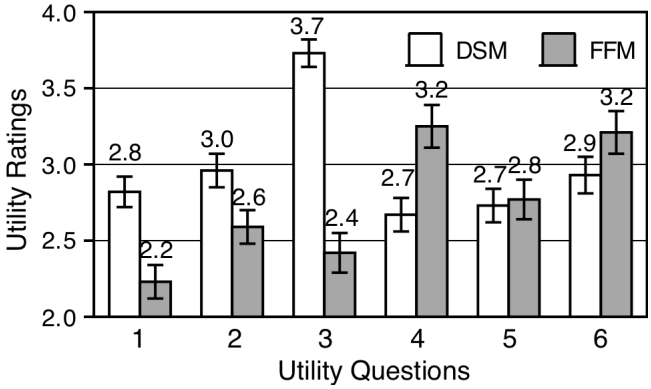

Figure 4: Utility Ratings (Std. Error)

Notes. Questions: (1) Making a prognosis; (2) Devising treatment plans; (3) Communicating with mental health professionals; (4) Communicating with patients; (5) Comprehensively describing all important personality problems; (6) Describing global personality.

General Discussion

Implications for the next version of the DSM

The FFM has been considered a highly promising candidate for describing personality pathology for the next version of the DSM. The five factors appear to provide reliable and valid summaries of personalities, and they can capture almost infinite varieties of personalities as opposed to just 10 personality disorders (Widiger & Trull, 2007). Despite these advantages, the current results, in conjunction with those reported in Rottman et al. (2009), suggest that it may be premature to adopt the FFM as a diagnostic tool for practicing clinicians. In the current study, experts in personality disorders (validated through publication records) had difficulty back-translating the FFM profiles of even highly familiar prototypic personality disorder cases.

These results are consistent with previous cognitive theories on two levels of specificity in representing features, termed *informational* vs. *instantiated* (Brooks & Hannah, 2006), or *global* vs. *local* (Solomon & Barsalou, 2001). For instance, "flying" is informational/global, whereas "bird-like flying" and "bee-like flying" are instantiated/local. Categories (e.g., birds) allow people to instantiate informational/global features, and develop more accurately nuanced and refined local/instantiated concepts (Brooks & Hannah, 2006). However, in the absence of categories, as in the current task and in a version of the FFM proposal, such instantiation is not feasible. We found that the loss of disambiguating information affects even domain experts.

Our experts also judged the clinical utility of the FFM to be low in a number of aspects, further suggesting that they found the FFM descriptors to be ambiguous. For instance, an abstract patient description (e.g., a neurotic, anxious, and introverted person) neither reveals the causes of the patient's pathology, nor does it help to make treatment plans and predictions about the course and outcome of the patient.

One obvious solution for the feature ambiguity problems demonstrated in the current study would be to provide disambiguating information. Thus, the FFM could be used as a supplement to diagnosis, which would provide categorical information, rather than as a stand-alone diagnostic tool for practicing clinicians. For example, a clinician could first identify that a patient has a particular personality disorder (e.g., Avoidant), and then use the FFM to further describe the patient, thereby instantiating the meanings of the FFM descriptors. For example, rather than thinking about a patient as 'withdrawn', a clinician could think of the patient as 'withdrawn due to paranoid fears' (as in Paranoid PD) or 'withdrawn due to indifference to others' (Schizoid PD). The instantiated descriptors will likely be more clinically meaningful and useful for clinicians than the ambiguous ones.

Implications for Research on Expertise

Rottman et al. (2009) speculated that the feature instantiation problems found in their study with practicing clinicians might be ameliorated with extensive training in the domain. Yet the current study demonstrated that the considerable training required in becoming a personality disorder researcher did not successfully reduce the effect of these problems. At first glance, these results may appear at odds with previous demonstrations of experts' impressive categorization and memorization abilities in their domain of expertise (e.g., Chase & Simon, 1973). For instance, Tanaka and Taylor (1991) showed that dog experts use subordinate category labels as quickly and frequently as they use basic-level category labels. Yet such demonstrations concern experts' rapid pattern recognition of objects presented with instantiated, unambiguous features (e.g., pictures of dog features). What the current study showed instead was a lack

of pattern recognition at an abstract level, an issue in itself deserving empirical investigation.

Nonetheless, it is also possible that a different form of training, unlike that of our participants, who after all were initially trained using the DSM system, could teach people to more productively use informational/global features. Furthermore, it is possible that different domains require experts to learn different types of information. For expertise in mental disorders, learning more accurate and nuanced instantiations of features may be necessary to differentiate patients, but for mathematics, expertise may require focusing on abstract informational or definitional concepts. Consequently, different domains may require classification systems that focus on different types of information.

Acknowledgments

This research was supported by an NSF Graduate Research Fellowship (Rottman), and NIMH Grants MH57737 (Ahn) and MH73708 (Sanislow). We thank Rachel Litwin for coding and recruiting participants.

References

Ahn, W., Marsh, J. K., Luhmann, C. C., & Lee, K. (2002). Effect of theory-based feature correlations on typicality judgments. *Memory & Cognition, 30*, 107-118.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: Text revision* (4th Ed.). Washington, DC: Author.

Benjamin, L. S. (1993). *Interpersonal diagnosis and treatment of personality disorders.* NY: Guilford Press.

Bornstein, R. F. (1998). Reconceptualizing personality disorder diagnosis in the *DSM-V*: The discriminant validity challenge. *Clinical Psychology: Science and Practice, 5,* 333-343.

Brooks, L. R., & Hannah, S. D. (2006). Instantiated features and the use of "rules." *Journal of Experimental Psychology: General*, *135*, 133-151.

Chase, W. G. & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4,* 55-81.

Clark, L. A. (2007). Assessment and diagnosis of personality disorder: perennial issues and emerging reconceptualization. *Annual Review of Psychology, 58,* 227-57.

Costa, P.T.J., & McCrae, R.R. (1992). *Revised NEO personality inventory: professional manual (NEO-PI-R)*. Odessa, Fl: Psychological Assessment Resources.

First, M. B. (2005). Clinical Utility: A Prerequisite for the Adoption of a Dimensional Approach in DSM. *Journal of Abnormal Psychology, 114,* 560–564.

First, M. B., Pincus, H. A., Levine, J. B., Williams, J. B. W., Ustun, B., & Peele, R. (2004). Clinical utility as a criterion for revising psychiatric diagnoses. *American Journal of Psychiatry*, *161*, 946-954.

Kamp, H. (1975). Two theories about adjectives. In E. Keenan (Ed.), *Formal Semantics of Natural Language*. Cambridge: Cambridge U.P. 123-55.

Lynam, D. R., & Widiger, T. A. (2001). Using the Five-Factor Model to represent the *DSM-IV* personality disorders: An expert consensus approach. *Journal of Abnormal Psychology, 110,* 401-412.

Mischel, W., & Peake, P. K. (1982). Beyond déjà vu in the search for cross-situational consistency. *Psychological Review, 89,* 730-755.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102,* 246-68.

Murphy, G. L. (1988). Comprehending complex concepts. *Cognitive Science, 12,* 529-562.

Murphy, G. L. (2002). *The big book of concepts.* Cambridge, MA: MIT Press.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, 92, 289-316.

Rosch, E. (1978). Principles of categorization. In E. Rosch and B. B. Lloyd (eds.), *Cognition and Categorization*. Hillsdale, NJ: Erlbaum.

Rottman, B. M., Ahn, W., Sanislow, C. A., & Kim, N. S. (2009). Can clinicians recognize DSM-IV personality disorders from Five-Factor Model descriptions of patient cases? *American Journal of Psychiatry, 166,* 427-433.

Samuel, D. B., & Widiger, T. A. (2004). Clinicians' personality descriptions of prototypic personality disorders. *Journal of Personality Disorders. 18,* 286-308.

Samuel, D. B., & Widiger, T. A. (2006). Clinicians' judgments of clinical utility: a comparison of DSM-IV and Five Factor Models. *Journal of Abnormal Psychology, 115,* 298-308.

Solomon, K. O., & Barsalou, L. W. (2001). Representing properties locally. *Cognitive psychology, 43,*129-169.

Spitzer, R. L., First, M. B., Shedler, J., Westen, D., & Skodol, A. E. (2008). Clinical utility of five dimensional systems for personality diagnosis: A "consumer preference" study. *Journal of Nervous and Mental Disease, 196*, 356-374.

Sprock, J. (2003). Dimensional versus categorical classification of prototypic and nonprototypic cases of personality disorder. *Journal of Clinical Psychology, 59,* 991-1014.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology, 23,* 457-482.

Wattenmaker, W. D., Dewey, G. I., Murphy, T. D., & Medin, D. L. (1986). Linear separability and concept learning: Context, relational properties, and concept naturalness. *Cognitive Psychology, 18,* 158-194.

Westen, D. (2006). Drizzling on the 5+3 factor parade: A response to Trull. In T. Widiger & E. Simonson (Eds.), *Research Agenda for DSM-V: Dimensional Classification of Personality Disorder.* DC: American Psychiatric P.

Widiger, T. A, & Trull, T. J. (2007). Plate tectonics in the classification of personality disorder: shifting to a dimensional model. *American Psychologist, 62,* 2 71-83.