Interests, Gender, and Science

Virginia Valian

Perspectives on Psychological Science 2014 9: 225
DOI: 10.1177/1745691613519109

aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Perspectives on Psychological Science
2014, Vol. 9(2) 225–230

DOI: 10.1177/1745691613519109
pps.sagepub.com
SAGE

Interests, Gender, and Science

Virginia Valian
Hunter College and CUNY Graduate Center

Abstract

In this commentary on Nye, Su, Rounds, and Drasgow (2012) and Schmidt (2011), I address the value of occupational interest inventories for understanding sex differences in occupational choice and the extent to which occupational interests are malleable. In particular, I argue (a) that some subscales in interest inventories are too heterogeneous to be given a single label and that the labels that are applied to some subscales are inaccurate and misleading; (b) that "things versus people" is an inaccurate and misleading characterization of a dimension that is frequently associated with interest inventories and linked to sex differences; (c) that vocational interests will be valid predictors of job performance primarily in cases in which the job has been held for some time by a diverse group of people and not in cases in which jobholders have been homogeneous; (d) that sex differences in interests are malleable and sensitive to small and subtle environmental cues; and (e) that women's interest in math and science will increase if they have a feeling of belonging and an expectation of success.

Keywords

vocational interests, gender differences, personnel selection, job performance ratings

A shift has taken place in reasoning about sex differences in the distribution of males and females in occupations. Although there is still reference to sex differences in ability in science and math in favor of men (e.g., Wai, Cacchio, Putallaz, & Makel, 2010), there is an increasing emphasis on sex differences in interest in science and math. Women can do science and math, but they prefer to do something else. That shift in thinking is good news for individuals who think that better science will result if the demographics of the scientific workforce resemble the demographics of the population. Interests, after all, can change. We know that because interests have changed. The percentage of women in the military, for example, increased from 2% in 1973 to 14% in 2011. If we understand what underlies male and female interests, and how the physical and social environment influences those interests, we can potentially enlarge the scope of both sexes.

Or can we? There is an alternate story that goes like this: (a) Vocational interests are correlated with performance—people perform better at and spend more time at work they are interested in (Nye, Su, Rounds, & Drasgow, 2012), (b) interests are sex-linked—men prefer things and women prefer people (Schmidt, 2011, drawing on Su, Rounds, & Armstrong, 2009), and (c) interests are probably hard to change (Schmidt, 2011). In this comment on Nye et al. (2012) and Schmidt (2011), I want to cast an eye that is simultaneously appreciative of their scrupulous work and skeptical of the story they tell about interests and sex differences.

Nye et al. (2012) have provided a much-needed study of the relation between vocational interests and job performance in which they go beyond prior work to examine whether the fit between someone's interest profile and an occupation's profile—the congruence index—predicts performance. If interests do not predict performance, there seems to be little practical point in measuring interests or using them for vocational counseling. Although interests might predict choice of college major or occupation, interests alone would not inform individuals about what jobs they would succeed in and enjoy and thus would be of limited value. Nye et al. convincingly showed that congruence indices moderately predict ratings of job performance and employee persistence. To create a congruence index, an individual's three major interest categories and an occupation's three major interest categories are compared. A congruence index measures the extent to which the resulting profiles match.

Corresponding Author:
Virginia Valian, Department of Psychology, Hunter College, 695 Park Avenue, New York, NY 10065
E-mail: virginia.valian@hunter.cuny.edu

When there is high overlap, ratings of job performance are higher. Thus, interest ratings can be helpful for vocational counseling. Nye et al. were not directly concerned with sex differences but with the predictive power of vocational interests. But as vocational interests do show sex differences, it is necessary to examine in detail how interests are characterized and measured. Despite my admiration for the advance Nye et al. (2012) have made, I am skeptical about the underlying model of vocational interests that they have used (Holland, 1997).

Heterogeneous Subscales With Inaccurate and Misleading Labels

Holland's (1997) model of interests has six scales: Investigative (Thinkers), Artistic (Creators), Conventional (Organizers), Enterprising (Persuaders), Social (Helpers), and Realistic (Doers). The terms in parentheses reflect commonly-associated labels for the personality types associated with each scale. Two of the subscales are relatively homogeneous, coherent, and well named. Investigative interests are those that natural and social scientists, mathematicians, and engineers, among others (detectives and muckrakers, perhaps), know and love. Artistic concerns artistic expression of all sorts.

Three more subscales are relatively homogeneous but have nondescriptive labels. Conventional interests correspond to whatever happens in a structured setting, like an office or business environment (even though business environments need not be highly structured). It might better be termed *Managing*. *Enterprising* refers to being interested in leadership, usually with an economic goal, and might better be termed *Leadership* or, as it is commonly referred to, *Persuaders*. *Social* is conceptually uniform but is not "social" in a broad sense. Instead, it indicates interest in helping people, as the commonly associated label *Helpers* suggests. *Nurturant* or *Communal* would be equally accurate descriptors.

The Realistic inventory stands out from the others in being neither coherent nor well-named. It includes interests in outdoor activities and mechanical activities. These two activities have little in common; they do not make up a natural kind. Nor is *Realistic* an accurate descriptor; *Instrumental* or *Task-Oriented* might come closer, as the label *Doers* suggests. When a scale has no obvious internal conceptual coherence, but the items nevertheless cluster together, the most likely reason is that there is an underlying concept that indirectly ties those items together. Radically different items that exemplify the underlying concept can cluster together if they are different aspects of an underlying theme. In this case, the underlying theme is "activities that men have tended to spend more time at than women have."

Not surprisingly, there are large sex differences in scores on two of the six scales: Social and Realistic. I suggest that the reason for the sex differences is that implicit gender schemas on the part of scale constructors unintentionally gave rise to those two scales. Gender schemas are nonconscious hypotheses about male and female personality characteristics (Fiske & Taylor, 1991; Martin & Halverson, 1987; Spence & Helmreich, 1978; Spence & Sawin, 1985; Valian, 1998). According to the male gender schema, men are agentic, instrumental, and task-oriented; according to the female gender schema, women are communal, nurturant, and expressive. Those schemas map very well onto the Realistic and Social components, respectively. If antecedent implicit hypotheses about gender differences—hypotheses that have a kernel of truth—played an implicit and unintentional role in creating the Realistic and Social scales, it would be expected that those scales would distinguish males and females. If you take activities and interests that are characteristic of men, they will cluster together; this also occurs for activities and interests that are characteristic of women. The informal characterizations of the two scales—Doers and Helpers—align exceptionally well with gender schemas.

Things Versus People

The bipolar dimension "things versus people" (Prediger, 1982), in which men excel at things and women excel with people, has been the label used to divide the six components. The three components that contribute to the *things* dimension are Realistic, Conventional, and Investigative. Note that there are no conceptual grounds on which one could call Realistic, Conventional, and Investigative *thing*-like. *Thing* would have to be interpreted so broadly that it could refer to anything at all: the global economy, string theory, mental representations, or tennis. *Thing* is vacuous in this use. Other than clustering empirically, the components have little or nothing in common. One struggles for a nonexistent cover term.

In one sense, the lack of an adequate cover term does not matter: the point is that the dimensions cluster and the name is irrelevant; one could call the dimension *X*. In another sense, however, the label does matter. Once a cluster name has been reified and associated with males, as is the case with *things*, people who do not know what the dimension actually consists of—which is most people—think that men are oriented to things, using the conventional meaning of the term. If the label were *X*, it would give no impression to the uninformed, and it would not indicate a unifying principle; that would be preferable because it is closer to the truth.

How *people*-like is the people dimension? In addition to Social, the other two components are Enterprising and

Artistic. Enterprising concerns people in that leadership is involved, but people are primarily a means to an end. Artistic involves the creation of things (novels, paintings, ballets) or the performance of things (music, dance, theater). People are involved, but no more so than they are involved in Conventional or Investigative or Realistic interests. (In addition to the hypothesized "things vs. people" dimension is a "data vs. ideas" dimension, Prediger, 1982—oddly, Realistic, Investigative, Social, and Artistic, are considered to be at the ideas end, whereas Conventional and Enterprising are considered to be at the data end.) Again, in one sense, the labeling of this empirical cluster as *people* instead of *Y* is harmless, but to those who do not know what underlies the impression, it gives a misleading impression. As even scientists who do know the underlying data seem to think that men are primarily interested in things and women are primarily interested in people, the power of the terms to mislead seems clear.

A similar example of mislabeling is from a short measure that uses everyday activities to reveal an orientation to things or people (Graziano, Habashi, & Woodcock, 2011). The *thing* items include examples like "Take apart and try to reassemble a desktop computer." They do not include any examples like "Take apart and try to reassemble a dress" or "Try to recreate a dish tasted in a restaurant." If *thing* items are limited to activities that primarily men do, women will appear uninterested in things. Again, an implicit idea of gender differences seems to have motivated the choice of items that then show gender differences. And, again, an inaccurate label leads the unwary to conclude that men like things and women like people.

In addition to the conceptual problems is an empirical problem. Interest categories cluster, but they are not bipolar (Tay, Su, & Rounds, 2011). All the interest categories are positively correlated, even those that are expected to be negatively correlated. Given the internal structure of the categories, and the fact that they are not conceptually opposite, it makes sense that they are also not empirically opposite. Perhaps most important, about 55% of people are interested in both things and people, with one subgroup somewhat more interested in things and another somewhat more interested in people. Interests, as one might think pretheoretically, are capacious and varied: Individuals can be interested in opera and football, in understanding the mechanism of language acquisition and in helping people, in being a leader and in writing a novel. Speaking of the dimensions as bipolar is factually wrong and mischaracterizes a fundamental truth about people, a truth that Holland recognized: People are not just one kind of thing and they are not interested in just one kind of thing.

In casting a cold eye on interests, I have emphasized two points. First, I have emphasized the lack of conceptual coherence within interests (e.g., Realistic) and within dimensions (e.g., *things*). I have suggested that features other than the intended ones might be responsible for the empirical clusters—namely, gender schemas. The fact that two factors accounted for a majority of the variance (Prediger, 1982) does not provide information about exactly what the categories have in common, only that they have something in common.

Second, I have emphasized how unsuitable or misleading the labeling is. I place that emphasis because people are likely to rely on what the names suggest in ordinary language, independent of their reference within a technical domain. *Realistic* is the most egregious example of divergence from ordinary language, but "things versus people" and "data versus ideas" suggest dichotomies that a close examination of the categories does not support. Using "things versus people" as a shorthand both exaggerates and mischaracterizes sex differences.

When Are Vocational Ratings Valid Predictors of Job Performance?

Now consider the correlation between congruence indices and performance (Nye et al., 2012). Ratings of job performance depend on the raters' conceptions of what the job requires. Those ratings in turn depend in part on the characteristics of the jobholders. Our conceptions of what a job entails derive in part from our observations of people who occupy the job. A rater can thus use standards that are irrelevant but that are in place because they were at one time typical of most jobholders. If the demography of an occupation becomes less homogeneous, raters run the risk of taking extraneous properties that the original jobholders shared as relevant to job performance. Having seen a limited number of ways of solving a problem, raters may see those ways as the only ways. Raters may be caught in a *local maximum* (Hong & Page, 2004; Page, 2008): People who think the same, no matter how intelligent they are, tend to think the same and approach problems the same way. Homogeneous groups are thus less likely than a more diverse group to come up with atypical solutions to problems.

Consider a hypothetical scenario of everyday problem solving in which police officers detain suspects. Say that using force on a suspect is typical of male police officers, the initial occupants of law enforcement. Say that female police officers slowly enter the field and, in contrast, are less likely to use force (Rabe-Hemp, 2008). Raters of performance might perceive a quick use of force as normative, desirable, and necessary for the job even if it leads to more injury of citizens and police (Rabe-Hemp, 2008).

All correlations between interests and job ratings are subject to this caution. Supervisors' standards for good performance are conditioned by the norms set by previous jobholders, whether those norms are objectively desirable or not. New, nontraditional occupants in previously homogeneous fields are rated by people with a relatively traditional idea of what the job requires. Congruence indices will be most useful—and more likely to be valid—for fields that have had a diverse group of jobholders for some time; that diversity will help reduce accidental correlations. (In the case of women in business, current findings suggest that supervisors rate women somewhat more highly than men on their performance in their current middle-manager position but simultaneously rate women as less promotable; Roth, Purvis, & Bobko, 2012—in this interesting twist, being good at your job means you should stay there if you are a woman.)

Another way that sex differences in interests can be misleading concerns measures of aptitude. In his elegantly argued contribution on sex differences in technical aptitude, Schmidt (2011) started from two premises: general mental ability—the ability to learn—predicts job performance better than any set of specific abilities; no sex differences in general mental ability exist. Thus, any specific aptitude test that is biased in favor of either sex will not correctly measure the underlying general mental ability of the sex with less experience in the specific area being tested. Females have less experience with technical activities than do males; as a result, tests of technical aptitude favor males and underestimate females' general mental ability. Thus, employers in technical areas would be ill-advised to make decisions about prospective employees based on technical aptitude tests. Schmidt's argument here is similar to that of Wainer and Steinberg (1992), who showed that the SAT math test underpredicts women's performance in college math courses. Only use of a composite that combines high-school grade point average with the math SAT reduces the underprediction of women's college math performance (Bridgeman & Lewis, 1996). Schmidt drew several predictions from his model, all of which receive support and range from weak to very strong. It was a pleasure to follow the argument.

Malleability of Interests

Schmidt's (2011) article suggests that sex differences in interests are not only responsible for the sex differences in familiarity with technical areas, but that they are also hard to change. The first half of the proposal has substantial support, the second much less. Although Schmidt also argued that it is not necessary to change interests in order to reduce sex differences in science fields, it seems clear that the differences would be reduced even more if one could change interests. Schmidt gave three main reasons for thinking interests are hard to change: (a) "men prefer working with things and women prefer working with people" (p. 570, citing findings by Su et al., 2009), (b) interests are stable within individuals over time, and (c) interests are heritable. I would argue that none of those reasons provide the necessary support.

The appeal to the "things versus people" sex difference founders, as I suggested earlier, because the things-people dichotomy is false and is not bipolar. The appeal to stability of interests over time ignores the stable social contexts that support interests differentially as a function of social class, ethnicity, region, sex, and other variables, and the difficulty of starting over once one has invested time and energy in an interest. It also ignores the fact that the representation of women in fields formerly restricted to men has greatly increased. Women's and men's interests as measured by vocational tests may be stable, but their interests as measured by vocational choices have changed: There are more women engineers and more men nurses than there were. Women may simultaneously want to help people and want to go to the moon. Men may simultaneously want to go to the moon and want to help people. Vocational interests and abilities are among the determinants, but hardly the only determinants, of occupational choice.

The appeal to heritability ignores the fact that highly heritable traits change, sometimes rapidly. Height—a highly heritable trait (depending on the population and sex, estimates range from .60 to about .80)—is also highly sensitive to the environment. For example, 6- to 12-year-old Mayan children growing up in the United States are, on average, 10.24 cm taller than those growing up in Guatemala (Bogin & Rios, 2003)—a huge difference. Most of that difference (7.02 cm) comes from an increase in leg length. Alternate possible causes that can be ruled out include: (a) differences between parents who migrated and those who remained in Guatemala; (b) climate differences between the United States and Guatemala—the climates of Florida and Los Angeles are similar to that of Guatemala; (c) genetic changes—the parents had typically migrated no more than 20 years earlier; and (d) earlier achievement of the adult height—9-year-old Mayan-Americans were taller than adult Guatemalan Mayans, and Mayan-American growth continued linearly through age 12 (at which point measurements stopped). A high heritability coefficient does not imply a fixed trait. In the case of height, we have a good idea of what contributes to an increase: better nutrition (though the Dutch continue to get taller despite the absence of recent nutritional changes; Fredriks et al., 2000); better sanitation; better health care; higher parental education; and in some cases, better climate.

Malleability and Feeling of Belonging

We should expect interests to be at least as malleable as height, and they are. In fact, interests are responsive to small environmental differences. Several recent studies reveal environmental characteristics that will increase or decrease female interest in math and science. For example, young adults who had identified themselves as good at math tasks and as believing that it is important to do well in math tasks watched a video about a summer leadership conference for college students in math, science, and engineering (Murphy, Steele, & Gross, 2007). One form of the video showed equal numbers of male and female students participating—a total of about 150; the other form of the video had three male students for every female student.

Women who saw the gender-unbalanced video showed greater physiological vigilance (changes in heart rate and skin conductance) than did women who saw the gender-balanced video or men who saw either video. Those women reported a lower sense of belonging than did the other three groups. Women also preferred to attend the gender-balanced conference (as did men, presumably for different reasons). Consider the effect of the accumulation of such small cues signaling lack of belonging for girls with a nascent—rather than already well-developed—interest in math and science. What is today a preference for or against going to a science conference is tomorrow a preference for or against a career in science.

Simply changing the objects in the classroom environment influences women's interest in computer science (Cheryan, Plaut, Davies, & Steele, 2009). Individual male and female college students were queried about their interest in computer science in a classroom that either had items that were associated with male computer science students, such as a Star Wars poster and soda cans, or items that were not associated, such as a nature poster and a water bottle. Men showed the same interest, regardless of environment; women showed more interest in the neutral environment. Only in the stereotypical environment did men show more interest in computer science than did women. Again, what is today the presence or absence of interest in taking a computer science course is tomorrow the presence or absence of interest in choosing computer science as a career.

Other studies similarly demonstrate the malleability and complexity of choice and interest (Cheryan, Siy, Vichayapai, Drury, & Kim, 2011; Stout, Dasgupta, Hunsinger, & McManus, 2011). Female college students who intend to major in science, technology, engineering, or mathematics disciplines are positively affected by the presence of same-sex experts, whereas men are unaffected; the effect on women appears to be mediated by their identifying with the female expert (Stout et al., 2011). Female college students express more interest in computer science if the fellow student with whom they interact is dressed nonstereotypically, whether that person is male or female (Cheryan et al., 2011).

A range of subtle and sometimes not-so-subtle cues influence women's and men's feelings of belonging and expectations of success, which in turn lead to interests in one field or another. When women and men think they belong and have a good chance of being successful, they are, not surprisingly, more interested in a field. Environments (and people) need not be overtly hostile to women or men in order to exclude them. To ensure the success of nontraditional jobholders, raters will have to develop appropriate standards of good performance that are uncontaminated by irrelevant properties of jobholders. If we change the environment of math and science, we will change women's interest in math and science.

Acknowledgments

I thank Martin Chodorow and Mary C. Potter for helpful comments.

Declaration of Conflicting Interests

The author declared that she had no conflict of interest with respect to their authorship or the publication of this article.

Funding

This work was supported in part by National Institutes of Health Grant R01GM88530 and benefited from workshops sponsored by the Leverhulme Trust (Grant IN-058). Any opinions, findings, and conclusions or recommendations expressed in this article are those of the author and do not necessarily reflect the views of the National Institutes of Health or the Leverhulme Trust.

References

Bogin, B., & Rios, L. (2003). Rapid morphological change in living humans: Implications for modern human origins. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, *136*, 71–84.

Bridgeman, B., & Lewis, C. (1996). Gender differences in college mathematics grades and SAT-M scores: A reanalysis of Wainer and Steinberg. *Journal of Educational Measurement*, *33*, 257–270.

Cheryan, S., Plaut, V. C., Davies, P. G., & Steele, C. M. (2009). Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology*, *97*, 1045–1060.

Cheryan, S., Siy, J. O., Vichayapai, M., Drury, B. J., & Kim, S. (2011). Do female and male role models who embody STEM stereotypes hinder women's anticipated success in STEM? *Social Psychological and Personality Science*, *2*, 656–664.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York, NY: McGraw-Hill.

Fredriks, A. M., van Buuren, S., Burgmeijer, R. J. F., Meulmeester, J. F., Roelien, J., Beuker, R. J., . . . Wit, J.-M.

(2000). Continuing positive secular growth change in the Netherlands 1955–1997. *Pediatric Research*, *47*, 316–323.

Graziano, W. G., Habashi, M. M., & Woodcock, A. (2011). Exploring and measuring differences in person–thing orientations. *Personality and Individual Differences*, *51*, 28–33.

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa, FL: Psychological Assessment Resources.

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences, USA*, *101*, 16385–16389.

Martin, C. L., & Halverson, C. (1987). The roles of cognition in sex role acquisition. In D. B. Carter (Ed.), *Current conceptions of sex roles and sex typing: Theory and research* (pp. 123–137). New York, NY: Praeger.

Murphy, M. L., Steele, C. M., & Gross, J. J. (2007). Signaling threat: How situational cues affect women in math, science, and engineering settings. *Psychological Science*, *18*, 879–885.

Nye, C., Su, R., Rounds, J., & Drasgow, F. (2012). Vocational interests and performance: A quantitative summary of over 60 years of research. *Perspectives on Psychological Science*, *7*, 384–403.

Page, S. (2008). *The difference: How the power of diversity creates better groups, firms, schools, and societies*. Princeton, NJ: Princeton University Press.

Prediger, D. J. (1982). Dimensions underlying Holland's hexagon: Missing link between interests and occupations? *Journal of Vocational Behavior*, *21*, 259–287.

Rabe-Hemp, C. E. (2008). Female officers and the ethic of care: Does officer gender impact police behaviors? *Journal of Criminal Justice*, *36*, 426–434.

Roth, P. L., Purvis, K. L., & Bobko, P. (2012). A meta-analysis of gender group differences for measures of job performance in field studies. *Journal of Management*, *38*, 719–739.

Schmidt, F. L. (2011). A theory of sex differences in technical aptitude and some supporting evidence. *Perspectives on Psychological Science*, *6*, 560–573.

Spence, J. T., & Helmreich, R. L. (1978). *Masculinity and femininity: Their psychological dimensions, correlates, and antecedents*. Austin: University of Texas Press.

Spence, J. T., & Sawin, L. L. (1985). Images of masculinity and femininity: A reconceptualization. In V. E. O'Leary, R. K. Unger, & B. S. Wallston (Eds.), *Women, gender, and social psychology* (pp. 35–66). Hillsdale, NJ: Erlbaum.

Stout, J. G., Dasgupta, N., Hunsinger, M., & McManus, M. A. (2011). STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology*, *100*, 255–270.

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interest. *Psychological Bulletin*, *135*, 859–884.

Tay, L., Su, R., & Rounds, J. (2011). People-things and data-ideas: Bipolar dimensions? *Journal of Counseling Psychology*, *58*, 424–440.

Valian, V. (1998). *Why so slow? The advancement of women*. Cambridge, MA: MIT Press.

Wai, J., Cacchio, M., Putallaz, M., & Makel, M. C. (2010). Sex differences in the right tail of cognitive abilities: A 30-year examination. *Intelligence*, *38*, 412–423.

Wainer, H., & Steinberg, L. S. (1992). Sex differences in performance on the mathematics section of the Scholastic Aptitude Test: A bidirectional validity study. *Harvard Educational Review*, *62*, 323–337.