

UNDERSTANDING THE STRUCTURE OF STEREOTYPES OF WOMEN: VIRTUE AND AGENCY AS DIMENSIONS DISTINGUISHING FEMALE SUBGROUPS

C. Nathan DeWall
Florida State University

T. William Altermatt
Hanover College

Heather Thompson
University of Michigan–Flint

A two-part study investigated the dimensional structure of stereotypes of women. In one sample ($n = 258$), participants sorted traits according to the likelihood that they would co-occur in the same woman. In a separate sample ($n = 102$), participants were given the same traits and were asked to judge the traits' desirability and to judge the moral virtue, sexual liberalism/conservatism, warmth, competence, and power of a woman who possessed high levels of each trait. Results from hierarchical cluster analysis indicated that participants perceived women in terms of six subgroups: professional, feminist, homemaker, female athlete, beauty, and temptress. Large differences among these subgroups were identified based on ratings of their moral virtue and sexual conservatism (i.e., virtue) and competence and power (i.e., agency). The implications of a virtue–agency model of female subgroups for gender stereotyping research are discussed.

What is the structure of stereotypes of women? For many years, psychologists have sought to answer this question using various strategies intended to highlight perceived similarities and differences between different “types” of women. Researchers in this tradition have consistently shown that people possess a schema of the stereotype of women that is organized in terms of three major subgroups: homemaker, professional, and sex object (Deaux, Winton, Crowley, & Lewis, 1985; Eckes, 1994a, 1994b; Six & Eckes, 1991). As the field of gender stereotyping evolved, the range of female subgroups under investigation expanded to include subgroups identified by their political stance (e.g., feminist) and physical activity (e.g., female athlete), domains previously dominated by men¹.

C. Nathan DeWall, Department of Psychology, Florida State University; T. William Altermatt, Department of Psychology, Hanover College; Heather Thomson, Department of Psychology, University of Michigan–Flint. The first two authors contributed equally to this work.

The study reported in this article was part of C. Nathan DeWall's master's thesis, completed at the University of Chicago. Portions of this article were presented at the 2003 Annual Meeting of the American Psychological Society, Atlanta, GA. We wish to thank Penny Visser and Elizabeth Davies for their comments on an earlier version of this article.

Address correspondence and reprint requests to: C. Nathan DeWall, Department of Psychology, Florida State University, One University Way, Tallahassee, FL 32306-1270. E-mail: dewall@darwin.psy.fsu.edu

Subsequent research supported the subgroup approach, indicating that people use subgroups to organize feminine trait information in memory (Noseworthy & Lott, 1984), to define the boundaries of in-groups and out-groups (Vonk & Olde-Monnikhof, 1998), and to discriminate in favor of some subgroups and against others (Haddock & Zanna, 1994).

Recently, researchers have examined how well a small number of dimensions could describe the differences among the subgroups. In an earlier study (Altermatt, DeWall, & Leskinen, 2003), we found that the subgroups professional women, housewives, and sex objects could be distinguished from one another based on two dimensions: agency (a combination of power and competence) and virtue (both sexual and moral). We selected agency as a likely dimension because of the evidence for discrimination against women who express competence (for a review, see Lott, 1985) or power (Eagly, Makhijani, & Klonsky, 1992; Rudman & Glick, 1999; Rudman & Kilianski, 2000). Historically, moral and sexual virtue have also been important dimensions in evaluating women (Hunter, 1976; Tavis & Wade, 1984). In particular, the importance of sexuality to the structure of female subgroups has been acknowledged by several researchers (Deaux, 1995; Eckes, 1994a, 1994b; England, 1988; Green & Ashmore, 1998; Six & Eckes, 1991) but has been largely neglected in gender stereotype research (Ashmore, Del Boca, & Wohlers, 1986). In our previous study (Altermatt et al., 2003), we found that professional women were perceived to be high in agency, homemakers neutral, and sex objects low, while homemakers were

perceived to be high in virtue, professional women neutral, and sex objects low.

In addition to describing how the subgroups differ, the dimensional approach may also provide clues as to why they differ. Glick and Fiske's (2001) theory of paternalistic prejudice proposes that negative stereotypes about women (e.g., that women are less agentic than men) are less likely to meet resistance if they are accompanied by positive stereotypes (e.g., that women are more virtuous than men). The stereotype of low agency is likely to impede women's progress toward status equality, while virtue grants nominal respect but little actual power. Thus, agency and virtue may be important dimensions of female subgroups because those stereotypes are likely to maintain gender differences in status. Although it provided initial support for an agency–virtue structure for the stereotype of women, the Altermatt et al. (2003) study had several limitations. The present study was designed to address those limitations.

First, the Altermatt et al. (2003) study examined only the three subgroups most commonly found by researchers: professional, homemaker, and sex object. To test the generality of the agency–virtue model to other subgroups, the present study included the feminist and female athlete subgroups occasionally identified by subgroup researchers (Deaux et al., 1985; Eckes, 1994b; Noseworthy & Lott, 1984).

Second, virtue (especially moral virtue) is closely associated with social desirability. In large part, a morally virtuous person is someone who is “good” and not “bad.” As a result, we were concerned that our dimension of moral virtue was not distinct from a more general evaluation or “good–bad” dimension (Osgood, Suci, & Tannenbaum, 1957). To measure the relation between virtue and evaluation, we included measures of both moral virtue and general evaluation in the present study.

Third, the sex object subgroup is composed of traits that connote both physical attractiveness and seductiveness (Clifton, McGrath, & Wick, 1976; Eckes, 1994a; Noseworthy & Lott, 1984). One of the dimensions identified by Ashmore, del Boca, and Titus (1984) in their analysis of female subgroups was a sexual good–bad dimension, with “beautiful” at one end and “vixen” at the other. In addition, Glick, Diebold, Bailey-Werner, and Zhu (1997) theorized that sexually “bad” women may be perceived as threatening to men because they could use men's attraction to them to manipulate men. Because the distinction between the attractive and seductive aspects of the sex object subgroup may be theoretically meaningful, we included traits from both aspects in the present study.

Finally, Fiske, Cuddy, Glick, and Xu (2002) proposed a two-dimensional alternative to agency and virtue. Their stereotype-content model proposed that the content of out-group stereotypes is a function of the out-group's relative status and competition (or interdependence) with the in-group. The dimensions they proposed were competence (a function of status) and warmth (a function of interdependence; Fiske, 1998; Fiske, Xu, Cuddy, & Glick, 1999; Fiske

et al., 2002). From this perspective, there exist three main categories of traditionally stereotyped groups, each organized in terms of their levels of competence and warmth relative to the in-group: (a) those who elicit disrespect for perceived lack of competence (e.g., housewives, disabled people, elderly people), (b) those who elicit dislike for perceived lack of warmth (e.g., Asians, Jews, career women), and (c) those who elicit dislike and disrespect due to low levels of both competence and warmth (e.g., welfare recipients; Fiske et al., 2002). The dimensions used in the stereotype-content model are similar to those employed by the agency–virtue model of female subgroups. However, it remains unclear how closely the dimensions specified by each model relate to each other and to what degree certain dimensions may be used interchangeably. Competence plays a significant role in each model, serving as a subcomponent of agency in the agency–virtue model and as one of the two dimensions in the stereotype-content model. The relation between the second dimensions of the models (virtue and warmth) is less clear. In distinguishing among subgroups, both may be used as a proxy for social desirability. However, some of the attributes of the sex object subgroup (e.g., flirtatious, seductive) may be perceived as warm but not sexually virtuous. By including the dimensions of both models, in the present study we hope to reconcile any differences between these two models.

In addition to the different dimensions used, the two models use different methods to identify the content of female subgroups. Following previous person perception research (Ashmore & Tumia, 1980; Rosenberg, Nelson, & Vivekananthan, 1968; Rosenberg & Olshan, 1970), research on the agency–virtue model (Altermatt et al., 2003) employs the use of trait adjectives that participants sort into piles based on their perceived co-occurrence in the same woman. Cluster analysis is used to identify subgroups of traits that tend to be placed together. To explore the perceived differences among the subgroups, the trait ratings are averaged within subgroups and then compared across subgroups using repeated measures analysis of variance (ANOVA). In contrast, research in support of the stereotype-content model (e.g., Eckes, 2002; Fiske et al., 2002) has made use of methods in which participants are given several subgroup labels and asked to sort them in terms of their perceived similarity and dissimilarity with other subgroups. This approach allows for a consideration of many different subgroups (e.g., gay men vs. career women) and has also been used to include contextual factors such as physical appearance, social role, and typical environment (Eckes, 1996). Given these methodological differences, a comparison of the results from the two models is problematic. By incorporating both models into a single study with a uniform methodological design, a more accurate comparison is possible.

In summary, the present study was designed to make the following contributions to research on the structure of the stereotype of women: (a) to expand the number of subgroups from three (homemaker, career woman, sex object)

to five (adding feminist and female athlete), (b) to examine the relations between a general evaluative or good–bad dimension and the other dimensions, (c) to better represent both the “beautiful” and “seductive” aspects of the sex object subgroup, and (d) to directly compare the agency–virtue model (Altermatt et al., 2003) to the competence–warmth model (Fiske et al., 2002). It should be noted that the primary purpose of the study was to examine the dimensional structure of female subgroups and not to verify or expand upon the content of the subgroups.

In this study, one sample of participants ($n = 258$) sorted 45 traits into piles based on their likelihood of co-occurring in the same woman. These trait-sorting data were submitted to cluster analysis to identify subgroups. A separate sample ($n = 102$) rated each of the traits in terms of its desirability and the degree to which a woman scoring high on that trait would be perceived as warm, competent, powerful, morally virtuous, and sexually virtuous. The trait-rating data were used to test the distinctiveness of the subgroups on each dimension.

METHOD

Participants

Two hundred fifty-eight undergraduates from the University of Chicago ($n = 123$) and the University of Michigan–Flint ($n = 135$) participated in the first study (trait sorting) and 102 participants from the University of Chicago ($n = 42$) and the University of Michigan–Flint ($n = 60$) participated in the second study (trait rating). The median age of participants in the trait-sorting study was 20 years old; 51% were male; and 61% identified themselves as European American or Caucasian, 13% African American, 13%

Asian/Indian, 6% Latino/Latina, and 7% other. The median age of participants in the trait-rating study was 20 years old; 52% were male; and 76% identified themselves as European American or Caucasian, 5% African American, 9% Asian/Indian, 3% Latino/Latina, and 7% other. Of the 42 Chicago trait-rating participants, only 34 rated the traits along the competence dimension, 35 for the negative/positive dimension, and 37 for the sexually liberal/conservative dimension. In addition, two participants from the Chicago trait-sorting sample were eliminated from analyses due to incomplete study packets and an inability to understand the task. This left 256 trait-sorting participants for the cluster analysis.

Trait Sorting

Materials. Participants were given 45 3" × 5" index cards with one trait printed on the front of each. Nine traits were selected to represent each of the five subgroups. Thirty-seven of the traits were drawn from previous factor analytic and multidimensional scaling research (Clifton et al., 1976; Deaux et al., 1985; Eckes, 1994a, 1994b; Noseworthy & Lott, 1984; Six & Eckes, 1991). Eight additional traits (“Free-thinking,” “Liberal,” “Well-read,” “Competitive,” “Fit,” “Manipulative,” “Promiscuous,” and “Scheming”) were chosen to round out each subgroup. A complete list of traits and their sources is given in Table 1. Participants were also given a brief demographic questionnaire in which they reported their age, sex, and ethnicity.

Procedure. Participants were recruited to participate in a study on gender and personality traits. In the Chicago sample, participants were recruited at the main building for student activities at the University of Chicago and

Table 1
Traits Selected to Represent Each Subgroup

<i>Professional</i>	<i>Feminist</i>	<i>Female Athlete</i>	<i>Homemaker</i>	<i>Sex Object</i>
Ambitious ^{fg}	Believes in equality ^f	Athletic ^{ae}	Caring ^f	Beautiful ^{abfg}
College-educated ^{fg}	Critical of society ^d	Competitive ^h	Cleans things ^{bf}	Fashionably dressed ^{bcfg}
Goal-oriented ^f	Demanding ^g	Coordinated ^f	Dependent ^{ag}	Flirtatious ^f
Independent ^{cdfg}	Free-thinking ^h	Energetic ^f	Devoted to family ^f	Manipulative ^h
Intelligent ^{abcf}	Giving a speech ^e	Fit ^h	Faithful ^a	Pleasure-loving ^{ac}
Nice clothes ^{bc}	Liberal ^h	Healthy ^f	Gentle ^f	Promiscuous ^h
Self-confident ^{cf}	Politically committed ^d	Muscular ^b	Maternal ^f	Scheming ^h
Sophisticated ^a	Rebellious ^d	Sports-oriented ^f	Sensitive ^a	Seductive ^f
Strong-minded ^f	Well-read ^h	Strong ^b	Takes care of kids ^b	Sexually attractive ^g

Note. Traits obtained from previous research are identified by the following superscripts:

^aClifton, McGrath, & Wick (1976);

^bDeaux, Winton, Crowley, & Lewis (1985);

^cEckes (1994a);

^dEckes (1994b);

^eGreen & Ashmore (1998);

^fNoseworthy & Lott (1984);

^gSix & Eckes (1991);

^htraits added for the current study.

completed the task in return for a candy bar. Participants in the Flint sample were recruited from introductory psychology classes and received extra credit for their participation. Chicago participants completed the task individually, whereas Flint participants completed the task in groups of 4 to 12. After listening to an overview of the experiment and asking questions, participants gave written informed consent. Participants placed the traits into piles based on the likelihood that most people would believe that the same traits would go together or co-occur in the same woman. They were told that they could use as many piles as they liked, but that each trait could be placed into only one pile. When they finished sorting the cards, they recorded the traits on a pile-recording sheet and then completed the demographics questionnaire. Afterward, participants received both an oral and written debriefing, were given an opportunity to ask further questions, and were dismissed.

Trait Rating

Materials. Participants rated each of the 45 traits from the trait-sorting task (see Table 1) using a seven-page questionnaire. The first five pages asked participants to imagine a woman who possessed high levels of each trait and then to estimate how most people would rate that woman in terms of one of five properties: (a) moral virtue, which was defined as fair, trustworthy, and free from traditional vices such as swearing, smoking, and drinking; (b) sexually liberal/conservative, which was defined as likely to believe that sex is appropriate only in a committed relationship (sexually conservative) or that sex is okay in casual, short-term relationships; (c) power, which was defined as being able to influence others (e.g., give orders) and resist their influence when she wants; (d) competence, which was defined as being able to solve problems and effectively achieve her goals; and (e) cold/warm, which was defined as kind, generous, and compassionate (warm) as opposed to critical and selfish (cold). For example, the first item on the moral virtue page asked participants to imagine a woman who was very ambitious and then to indicate the degree to which most people would consider the woman morally virtuous. The rating scale ranged from 1 (*low*) to 7 (*high*) and was framed with regard to each property, with the midpoint labeled neutral. For example, the rating scale for the moral virtue property ranged from 1 (low moral virtue), 2–3 (below-average moral virtue), 4 (neutral), 5–6 (above-average moral virtue), and 7 (high moral virtue). The sixth page was designed to measure the evaluative or good–bad aspect of the traits and asked participants to evaluate each trait in terms of how positive or desirable the trait was for a woman. The evaluative measure allowed for a comparison of whether participants perceived women who differed in terms of their virtue (e.g., homemakers) and agency (e.g., professionals) as more, less, or equally desirable to one another. On the seventh page, participants reported their age, sex, and ethnicity.

Procedure. Recruitment and informed consent procedures were identical to those for the trait-sorting task. Participants completed the seven-page questionnaire packet, received both an oral and written debriefing, were given an opportunity to ask further questions, and were dismissed.

RESULTS AND DISCUSSION

Analysis proceeded in two steps. First, subgroups were identified using cluster analysis of the trait-sorting data. Second, distinctions among the subgroups were examined using repeated-measures ANOVA of the trait-rating data.

Identifying Subgroups

Although the traits sorted by participants were selected to represent five subgroups, it is possible that some traits were not representative of their subgroup or that some subgroups were similar enough that participants did not distinguish between them. To provide an empirical means of assessing subgroup membership, cluster analysis was used.

Data were generated by constructing a trait-by-trait matrix for each participant in which a 1 indicated that two traits had been placed into the same pile and a 0 indicated that the two traits had not been placed into the same pile. By summing over participants, a trait-by-trait matrix was obtained in which the entries corresponded to the number of participants who sorted the two traits into the same pile. Values in this matrix ranged from a high of 230 for muscular and sports-oriented to a low of 3 for, among others, caring and rebellious. Finally, the values were transformed from measures of similarity to measures of dissimilarity by subtracting each value from N .

In order to ascertain whether the Chicago and Flint samples were equivalent, we computed the correlation between the 990 corresponding cells of the two matrices. The sortings of the two samples were highly similar, $r(990) = .93$, $p < .0001$. Given their high degree of congruence, the Chicago and Flint matrices were combined for analysis.

The combined matrix was submitted to four types of cluster analysis: single-link, complete-link, average-link, and Ward's minimum variance procedure. Although there is no significance test for a clustering solution, Milligan and Cooper (1985) proposed a method for measuring how well a clustering solution approximates the clustering present in data. As the clustering procedure progresses, stimuli that are rated similarly are combined. At each combination, Milligan and Cooper's (1985) method correlates the original similarity data with a binary matrix of cluster membership. The correlation generally increases as the clustering solution more closely approximates the naturally occurring clustering in the data. At some point, clusters that are very different are combined, which causes the correlation to fall. The number and content of the clusters is determined by the clustering solution that produces the maximum correlation between the original similarity data and a binary

Table 2
Content of the Clusters

<i>Professional/Feminist</i>	<i>Female Athlete</i>	<i>Homemaker</i>	<i>Sex Object</i>
Professional	Athletic	Caring	Beauty
Ambitious	Competitive	Cleans things	Beautiful
College-educated	Coordinated	Dependent	Fashionably dressed
Goal-oriented	Energetic	Devoted to family	Nice clothes
Independent	Fit	Faithful	Sexually attractive
Intelligent	Healthy	Gentle	Tempress
Self-confident	Muscular	Maternal	Demanding
Sophisticated	Sports-oriented	Sensitive	Flirtatious
Strong-minded	Strong	Takes care of kids	Manipulative
Feminist			Pleasure-loving
Believes in equality			Promiscuous
Critical of society			Rebellious
Free-thinking			Scheming
Giving a speech			Seductive
Liberal			
Politically committed			
Well-read			

matrix of cluster membership. The largest correlation for the complete-link method ($r = .84$) occurred at four clusters, as it did for average-link ($r = .84$) and Ward's minimum variance procedure ($r = .82$). For single-link, the largest correlation ($r = .83$) was at five clusters, with the fifth cluster composed of a single item, dependent, which in all other methods was clustered into the homemaker cluster. Because three of the four methods placed the trait dependent with the homemaker cluster, we felt it would be appropriate to classify it there. The content of the clusters, as determined by the Milligan and Cooper (1985) maximum correlation method, is presented in Table 2.

In general, the content of the clusters conformed to our expectations and replicated results from previous subgrouping research. However, two findings were notable. First, participants did not distinguish between the professional and feminist subgroups; participants sorted the traits we selected to represent those groups together. Second, participants also tended to sort the beauty and temptress aspects of the sex object subgroup into a single cluster. However, these combinations occurred at the very end of the clustering process for both the professional-feminist and sex object subgroups. That is, cluster analysis initially created a professional group, a feminist group, a beauty group, and a temptress group. Only at the last stages of clustering were these combined into two groups. Although people may have believed that, for example, traits relating to beauty tend to be found with traits relating to seductiveness, they may still have seen many differences between beautiful and seductive traits. To explore the possible tensions between combinations of dissimilar traits, we split the professional-feminist and sex object subgroups into their lower-order groupings, as indicated in Table 2 by the labels professional, feminist, beauty, and temptress.

Distinguishing Among Subgroups

A subgroup's score on a particular property was calculated by averaging the property ratings for all of the traits within that subgroup cluster. For example, the moral virtue of the homemaker cluster was calculated by averaging the moral virtue rating of caring, cleans things, and each of the other seven traits listed under the homemaker subgroup in Table 2. The degree to which properties distinguished among subgroups was measured using a repeated measures ANOVA for each property. The results are presented in Table 3. As Table 3 indicates, between 79 and 89% of the variance in property ratings could be explained by subgroup membership.

Post hoc comparisons with Bonferroni correction were used to test whether subgroups were perceived significantly differently from one another on each of the six properties (see Table 4). The contrast between the beauty and temptress subgroups was remarkable, especially given the cluster analysis results indicating that they tended to be

Table 3

Repeated Measures Analysis of Variance for Property Ratings of Subgroups

<i>Property</i>	<i>df</i>	<i>F</i>	η^2
Moral virtue	5, 97	130.09***	.87
Sexually conservative/liberal	5, 92	68.42***	.79
Power	5, 97	115.42***	.86
Competence	5, 89	132.34***	.88
Warm/cold	5, 95	127.30***	.87
Negative/positive	5, 90	150.40***	.89

*** $p < .001$.

Table 4
Mean Ratings of Subgroups (Traits Within Subgroup Clusters)

	<i>Athlete</i>	<i>Feminist</i>	<i>Professional</i>	<i>Beauty</i>	<i>Homemaker</i>	<i>Temptress</i>
Evaluation	4.92 _a	4.54 _b	5.63 _c	5.78 _c	5.55 _c	2.91 _d
Moral virtue	4.41 _a	4.73 _b	4.99 _c	4.04 _d	5.52 _e	2.72 _f
Warmth	4.01 _a	4.09 _a	4.11 _a	3.98 _a	5.89 _b	2.99 _c
Sexually conservative	4.01 _a	3.99 _a	4.55 _b	3.62 _c	5.36 _d	2.60 _e
Power	5.06 _a	5.12 _a	6.00 _b	5.04 _a	3.67 _c	4.39 _d
Competence	4.93 _a	5.08 _a	5.98 _b	4.31 _c	4.16 _c	3.72 _d

Note. Means in the same row that share the same subscript do not differ at $p < .01$ according to post hoc comparisons with Bonferroni correction.

placed into a single sex object subgroup. On every property except for power, the temptress subgroup received the lowest rating. The homemaker subgroup was regarded as significantly less powerful than the temptress, who may be seen as more influential through her manipulation of men. On the evaluation dimension, the beauty subgroup received the highest rating and the temptress subgroup the lowest. Beauty may have been regarded as desirable because of the assumption that physically attractive people possess more positive traits (the “what is beautiful is good” effect; Dion, Berscheid, & Walster, 1972). Recent research has also shown that physical attractiveness is positively associated with status to a greater extent for females than for males (Maner, Gailliot, Kenrick, & Li, 2004). The current findings indicate that although participants in the trait-sorting sample placed traits associated with the beauty and temptress subgroups into the same cluster, participants in the trait-rating sample perceived the beauty and temptress subcomponents of the sex object subgroup as distinct from each other in terms of their relative levels of the properties under investigation. Thus, these results fill a previously identified gap in the gender stereotyping literature (Ashmore et al., 1986) by showing that these two aspects of the sexual image of women (i.e., beauty and temptress) are evaluated differently on a number of dimensions but are perceived as belonging to a single sex object subgroup.

Another distinctive pattern was observed for the professional, feminist, and athlete subgroups. All three subgroups were seen as relatively high in power and competence and relatively neutral on the other properties, with the professional subgroup typically receiving higher scores than the feminist and female athlete subgroups. The feminist and female athlete subgroups were not significantly different from one another on four of the six properties: sexual conservatism, power, competence, and warmth. As for moral virtue and evaluation (desirability), the feminist group was regarded as significantly higher in moral virtue but significantly lower in desirability than the athlete subgroup. These findings suggest that the feminist and athlete subgroups may have been clustered separately due to the distinctive physical traits that made up the female athlete subgroup that may be associated with higher desirability

and decreased moral virtue (e.g., muscular, strong), rather than differences in the ways the subgroups are generally perceived. Results from the cluster analysis indicated that participants perceived traits associated with the feminist subgroup as similar to traits associated with professional women. These two categories were subsumed by a high-power, virtue-neutral subgroup. By splitting this subgroup into two categories, we found that professional women were regarded significantly more positively than feminists in terms of power, competence, sexual conservatism, moral virtue, and overall evaluation.

Finally, the homemaker subgroup was seen as high in warmth, virtue, and sexual conservatism, but low in power and competence. In addition, the homemaker subgroup was tied for first with the professional and beauty subgroups in terms of overall evaluation. This lack of differentiation in evaluation ratings suggests that despite the trade-offs of the professional and homemaker subgroups with regard to virtue and agency, participants ultimately regarded them as equally desirable. Taken together, the obtained differences among the subgroups illustrate the utility of considering subgroups as differing along a number of dimensions.

Significant gender differences in the ratings of the subgroups were identified for three of the six subgroups. Compared to male participants, female participants considered the professional subgroup to be significantly more morally virtuous, $F(1, 97) = 9.1, p < .01$, sexually conservative, $F(1, 93) = 15.2, p < .001$, competent, $F(1, 91) = 6.1, p < .02$, and warm, $F(1, 96) = 8.0, p < .01$. Perhaps these differences reflect women’s greater appreciation for the multiple roles—both agentic and virtuous—often enacted by the professional subgroup. Female participants also gave significantly higher evaluation ratings to the homemaker subgroup than males did, $F(1, 92) = 10.3, p < .01$. This difference may reflect a tendency for women to value traditional female roles (e.g., nurturer, mother) more than men. Finally, male participants rated the temptress subgroup as significantly more desirable than female participants, $F(1, 92) = 4.3, p < .04$. Men may see the temptress as a potential candidate for a short-term sexual relationship. If so, women may see the temptress as driving down the value of monogamy (i.e., sex

can be obtained without commitment; Guttentag & Secord, 1983), and thus as a threat to male commitment.

Comparing the Agency–Virtue to the Stereotype–Content Model

The major difference between the agency–virtue and stereotype–content models is in their choice of properties distinguishing subgroups. The utility of a property to describe the structure of the stereotype of women is based in part on its ability to distinguish among subgroups. The results presented in Table 3 indicate that significant variation exists among subgroups on all six properties. However, the results of the post hoc analyses in Table 4 suggest that certain properties may better distinguish among the subgroups than others. On both the warmth and moral virtue properties, the homemaker subgroup received the highest rating and the temptress subgroup received the lowest. However, warmth did not distinguish among the four subgroups arranged between these endpoints—they were all rated around the midpoint of the 7-point scale. In contrast, all six subgroups were significantly distinguished from one another in terms of their perceived moral virtue. Ratings of sexual conservatism showed a pattern similar to those for moral virtue, with the exception that the feminist subgroup was not distinguished from the athlete subgroup in the case of sexual conservatism. These results suggest that, for the purpose of understanding the structure of subgroups of women, moral and sexual virtue may offer a better guide than warmth.

One of the concerns we had regarding moral virtue was that it might be synonymous with a global evaluation property. The results presented in Table 5 indicate that although the correlation between moral virtue and evaluation was high ($r = .83$), the two dimensions were not synonymous. On the moral virtue property, the order of subgroups from highest to lowest was homemaker, professional, feminist, athlete, beauty, and temptress. Comparing the moral virtue property with the evaluation property, beauty moved from fifth place to a three-way tie for first place with homemaker and professional. In addition, whereas the feminist subgroup was considered more morally virtuous than the

athlete, it was evaluated more negatively than the athlete. Thus, participants distinguished between the moral virtue and evaluation of particular traits in a meaningful manner.

The differences between power and competence are less clear than those between virtue and warmth. Despite the high correlation between trait ratings on these two properties ($r = .85$), the means reported in Table 4 indicate unique contributions from each. Whereas the beauty subgroup was regarded as more powerful than the homemaker subgroup, the two were considered equally competent. And although the feminist subgroup was considered more competent than the beauty subgroup, the two were equally powerful. Thus, both competence and power (which we collectively label agency) appear to be important properties in distinguishing among subgroups of women.

In summary, virtue and agency effectively describe much of the perceived differences among subgroups of women. The major objective of the agency–virtue model is to examine the structure of the stereotype of women. In contrast, the stereotype–content model is concerned primarily with explaining stereotypic conceptions of many different kinds of groups in society (Fiske, 1998). As a result, the stereotype–content model might generate predictions that better generalize to more diverse observations. Given the more narrow focus of the agency–virtue model, it is not surprising that this model may provide a more precise perspective on the structure of female subgroups. To be sure, the results of the current study demonstrated that both models captured important aspects of how the stereotype of women is organized. Competence, for example, is a useful part of both models and our findings support its use in understanding the degree to which members of certain subgroups are perceived as agentic. Warmth, on the other hand, does not appear to capture as much of the nuance of subgroups of women as moral virtue and sexual conservatism. Warmth may explain other, less common, subgroups of women as well as (or better than) moral virtue and sexual conservatism. Unfortunately, the design of the current study did not provide a direct statistical test comparing the power of the two models. A possible avenue for future research would be to manipulate descriptions of women in terms of their levels of certain properties (e.g., warmth, moral virtue) and examine their effect on the classification of women into subgroups. For the present, virtue and agency should be considered as properties that adequately describe much of the differences among commonly identified subgroups of women.

Table 5

Correlations Between Properties

Properties	Sexual				
	Conservatism	Power	Competence	Warmth	Evaluation
Moral virtue	.89**	.09	.51**	.83**	.83**
Sexual conserv	–	.00	.40**	.72**	.65**
Power		–	.85**	–.33*	.25
Competence			–	.09	.53**
Warmth				–	.75**

Note. Correlations based on average ratings of between 93 and 102 participants. For purposes of significant testing, the number of observations is equal to the number of traits (45).

* $p < .05$. ** $p < .001$.

Limitations

A general limitation of the current study lies in its methodological approach, which involves the reliance upon subgroup trait clusters rather than category labels representing groups of individual women. Two advantages to the use of category labels are (a) more categories can be represented than if trait labels are used and (b) category labels

may represent a more basic level of cognitive organization than traits. With regard to the second point, our results are limited by the degree to which subgroups are distinct from the traits we selected to represent them. That is, subgroup membership may include qualities that are distinct from the traits participants used in the current study, such as the social role members of the subgroup typically occupy (Eckes, 1996). However, by using traits rather than category labels, we gain more precise control over the meaning of a particular subgroup and the advantage of knowing exactly what traits make up each subgroup. Additionally, the method employed in this study may represent a more indirect measure than the group labels approach. Group labels may invoke a participant's social identity and their personal relation to those groups. For example, a conservative participant may feel more threatened by the label feminist than by the traits believes in equality or politically committed. Similarly, progressive participants may think it inappropriate to report negative characteristics about professional women, but may feel more comfortable responding to the trait strong-minded. Thus, a trait approach may offer a view of the organization of subgroups that is more cognitive and less polarized by motivational concerns than the category approach.

It remains uncertain whether the results from the current study are indicative of person perception processes used in everyday life, as we have addressed primarily the structure of classification in this study and not the process of classification. Future research should examine how perceptions of a particular woman are influenced by the traits she possesses, the subgroups of which she is a member, and the properties (e.g., moral virtue, power) associated with both. For the present, we can reflect on what the structure of subgroups might tell us about the stereotype of women.

A final limitation is that our participants were entirely from an undergraduate population. It is possible that a larger, more nationally representative sample may have yielded different results. Older participants, for example, may have rated the female athlete and professional/feminist subgroups more negatively than undergraduate participants because women occupied these roles less frequently in the past than in current times. Future research may benefit from examining perceptions of female subgroups among participants of many age groups.

Function of the Dimensions

What can be made of the finding that female subgroups can be reliably distinguished along dimensions of virtue and agency? Jost and Banaji (1994) proposed that one function of stereotypes is to justify existing status differences between groups. Building on this theory, Glick and Fiske (2001) suggested that paternalistic stereotypes, which combine negative stereotypes about a group's competence with positive stereotypes about its warmth, might be more re-

sistant to criticism than unmitigated negative stereotypes. We propose that power might be more immediately related to status than competence, even though the two are generally found together. If the function of paternalistic stereotypes is to preserve a status hierarchy, then a subordinate group that considered itself influential and capable of self-determination (powerful) would pose a greater threat than a group that considered itself proficient in some domain (competent). In the present study, the homemaker subgroup exemplifies the paternalistic stereotype—low in competence and power, but high in virtue. Members of a group stereotyped as low in competence and power are unlikely to be considered for high-status roles, restricting their access to political and economic resources. Nevertheless, the homemaker possesses traits considered to be very desirable, as indicated by its three-way tie for first place on the evaluation dimension. The traits of the homemaker subgroup are thus desirable but politically inert, a combination likely to keep the homemaker subgroup low in status.

Another theory of the function of stereotypes is more directly relevant to subgroups of women. Glick and colleagues (1997) proposed that the differences among female subgroups may be related to the different ways that women threaten existing status differences between men and women. According to their theory, professional women threaten men's status directly by competing for economic resources. Although not mentioned in their theory, feminists would presumably also pose a direct threat by seeking equal status for men and women. Female athletes could pose a less direct threat because they offer examples of women who are physically powerful. All three groups pose a threat to status differences because they are powerful—they are examples of women who are strong and independent. Glick and colleagues (1997) also identified an indirect threat posed by the beauty and temptress subgroups, which could exploit men's attraction to them. These hypotheses are difficult to test, and our data do not provide much confirming or disconfirming evidence for them. Although we might expect male participants to give lower evaluation ratings to traits that are threatening, their instructions were to rate the desirability of the trait for a woman, not their personal liking or disliking of the trait. The correlation between ratings of power and evaluation was positive for both male ($r = .24$) and female ($r = .25$) participants, suggesting that powerful traits were generally regarded as desirable traits. Similar ambiguity plagues assessments of the temptress subgroup. Although the temptress subgroup may pose a threat to men's status, the threat exists only because men are attracted to the subgroup, making a global negative evaluation unlikely. In fact, men gave the temptress subgroup significantly higher desirability ratings than women. The development of the subgroup-divisions-as-threats theory may benefit from future research that directly measures the perceived threat of each subgroup to existing status differences.

Concluding Remarks

The current research makes two contributions to the existing literature on subgroups of women. First, it elucidates some interesting aspects of the subgroups themselves. Most surprising were the contrasts between the temptress and beauty subgroups, which tended to be combined in participants' sorting of traits. Whereas the beauty subgroup was one of the most positively evaluated, the temptress subgroup was the lowest. A question for future research is how and why such polarized combinations of traits tend to be combined into a single subgroup. Whereas the temptress and beauty subgroups were rated differently but sorted together, two other subgroups were rated similarly but sorted separately: feminists and female athletes. Feminists and female athletes, as well as the professional subgroup, were generally seen as high in competence and power and relatively neutral with regard to virtue. Finally, the homemaker subgroup was seen as lowest in power but highest in virtue.

The second contribution made by the current study is additional support for the agency–virtue framework of female subgroups (Altermatt et al., 2003). Although the stereotype-content model (Fiske, 1998; Fiske et al., 1999, 2002) suggests that many stereotypes can be described in terms of the two dimensions of warmth and competence, the results of the current study indicated that female subgroups may be better described by moral and sexual virtue than by warmth. The current study also demonstrated the value of considering both competence and power in relation to the structure of female subgroups. Members of the homemaker and beauty subgroups, for example, were not judged differently with regard to their perceived competence, but were judged as distinct in terms of their perceived power. Whereas competence was strongly correlated with (and thus to some degree redundant with) virtue, power was relatively independent of virtue. Power—defined as the capacity to influence others and to resist their influence—may play a more important role than competence in maintaining status differences between groups because of its direct association with self-determination. In summary, an agency–virtue model not only describes the perceived differences among female subgroups but is also consistent with a system-justification model of stereotypes (Jost & Banaji, 1994) in which the traditional stereotype of women (homemaker) trades the potentially status-enhancing benefits of agency for the nominally desirable but status-neutral benefits of virtue.

Initial submission: May 15, 2004

Initial acceptance: December 5, 2004

Final acceptance: March 14, 2005

NOTE

1. It is important to note that the identification of additional subgroups was not the result of gender stereotyping researchers seeking to broaden the range of female subgroups. Rather, it

was the result of changes in people's (i.e., participants') beliefs about the variability of different groups of women.

REFERENCES

- Altermatt, T. W., DeWalla, C. N., & Leskinen, E. (2003). Agency and virtue: Dimensions of female stereotypes. *Sex Roles, 49*, 631–641.
- Ashmore, R. D., Del Boca, F., & Titus, D. (1984). *Types of women and men: Yours, mine, and ours*. Paper presented at the 1984 American Psychological Association Convention, Toronto, Canada.
- Ashmore, R. D., Del Boca, F., & Wohlers, A. (1986). Gender stereotypes. In R. D. Ashmore & F. Del Boca (Eds.), *The social psychology of female-male relations: A critical analysis of central concepts* (pp. 69–119). New York: Academic Press.
- Ashmore, R. D., & Tuma, M. L. (1980). Sex stereotypes and implicit personality theory: I. A personality description approach to the assessment of sex stereotypes. *Sex Roles, 6*, 501–518.
- Clifton, A. K., McGrath, D., & Wick, B. (1976). Stereotypes of woman: A single category? *Sex Roles, 2*, 135–148.
- Deaux, K. (1995). How basic can you be? The evolution of research on gender stereotypes. *Journal of Social Issues, 51*, 11–20.
- Deaux, K., Winton, W., Crowley, M., & Lewis, L. (1985). Level of categorization and content of gender stereotypes. *Social Cognition, 3*, 145–167.
- Dion, K. K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*, 207–213.
- Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin, 111*, 3–22.
- Eckes, T. (1994a). Features of men, features of women: Assessing stereotypic beliefs about gender subtypes. *British Journal of Social Psychology, 33*, 107–123.
- Eckes, T. (1994b). Explorations in gender cognition: Content and structure of female and male subtypes. *Social Cognition, 12*, 37–60.
- Eckes, T. (1996). Linking female and male subtypes to situations: A range-of-situation effect. *Sex Roles, 35*, 401–426.
- Eckes, T. (2002). Paternalistic and envious gender stereotypes: Testing predictions from the stereotype content model. *Sex Roles, 47*, 99–114.
- England, E. M. (1988). College student stereotypes of female behavior: Maternal professional women and assertive housewives. *Sex Roles, 19*, 365–385.
- Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 2, 4th ed., pp. 357–411). New York: McGraw-Hill.
- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*, 878–902.
- Fiske, S. T., Xu, J., Cuddy, A. C., & Glick, P. (1999). (Dis)respecting versus (dis)liking: Status and interdependence predict ambivalent stereotypes of competence and warmth. *Journal of Social Issues, 55*, 473–489.

- Glick, P., Diebold, J., Bailey-Werner, B., & Zhu, L. (1997). The two faces of Adam: Ambivalent sexism and polarized attitudes toward women. *Personality and Social Psychology Bulletin*, 23, 1323–1334.
- Glick, P., & Fiske, S. (2001). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychologist*, 56, 109–118.
- Green, R. J., & Ashmore, R. D. (1998). Taking and developing pictures in the head: Assessing the physical characteristics of eight gender types. *Journal of Applied Social Psychology*, 28, 1609–1636.
- Guttentag, M., & Secord, P. F. (1983). *Too many women? The sex ratio question*. Beverly Hills, CA: Sage.
- Haddock, G., & Zanna, M. (1994). Preferring “housewives” to “feminists”: Categorization and the favorability of attitudes toward women. *Psychology of Women Quarterly*, 18, 25–52.
- Hunter, J. (1976). Images of woman. *Journal of Social Issues*, 32, 7–17.
- Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology*, 33, 1–27.
- Lott, B. (1985). The devaluation of women’s competence. *Journal of Social Issues*, 41, 43–60.
- Maner, J. K., Gailliot, M. T., Kenrick, D. T., & Li, N. P. (2004). *Halos or sex-specific heuristics? Specifying the links between perceptions of attractiveness and social status*. Manuscript submitted for publication.
- Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, 50, 159–179.
- Noseworthy, C. M., & Lott, A. J. (1984). The cognitive organization of gender-stereotypic categories. *Personality & Social Psychology Bulletin*, 10, 474–481.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.
- Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality & Social Psychology*, 9, 283–294.
- Rosenberg, S., & Olshan, K. (1970). Evaluative and descriptive aspects in personality perception. *Journal of Personality & Social Psychology*, 16, 619–626.
- Rudman, L. A., & Glick, P. (1999). Feminized management and backlash toward agentic women: The hidden costs to women of a kinder, gentler image of middle managers. *Journal of Personality & Social Psychology*, 77, 1004–1010.
- Rudman, L. A., & Kilianski, S. E. (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin*, 26, 1315–1328.
- Six, B., & Eckes, T. (1991). A closer look at the complex structure of gender stereotypes. *Sex Roles*, 24, 57–71.
- Tavris, C., & Wade, C. (1984). *The longest war: Sex differences in perspective*. San Diego, CA: Harcourt Brace Jovanovich.
- Vonk, R., & Olde-Monnikhof, M. (1998). Gender subgroups: Intergroup bias within the sexes. *European Journal of Social Psychology*, 28, 37–47.