A Room With a Cue:
Personality Judgments Based on Offices and Bedrooms

Samuel D. Gosling and Sei Jin Ko
University of Texas at Austin

Thomas Mannarelli
INSEAD

Margaret E. Morris
Sapient

The authors articulate a model specifying links between (a) individuals and the physical environments they occupy and (b) the environments and observers' impressions of the occupants. Two studies examined the basic phenomena underlying this model: Interobserver consensus, observer accuracy, cue utilization, and cue validity. Observer ratings based purely on offices or bedrooms were compared with self- and peer ratings of occupants and with physical features of the environments. Findings, which varied slightly across contexts and traits, suggest that (a) personal environments elicit similar impressions from independent observers, (b) observer impressions show some accuracy, (c) observers rely on valid cues in the rooms to form impressions of occupants, and (d) sex and race stereotypes partially mediate observer consensus and accuracy. Consensus and accuracy correlations were generally stronger than those found in zero-acquaintance research.

An animal resting or passing by leaves crushed grass, footprints, and perhaps droppings, but a human occupying a room for one night prints his character, his biography, his recent history, and sometimes his future plans and hopes. I further believe that personality seeps into walls and is slowly released. . . . As I sat in this unmade room, Lonesome Harry began to take shape and dimension. I could feel that recently departed guest in the bits and pieces of himself he had left behind.

—John Steinbeck, *Travels With Charlie*

Steinbeck's observation draws on an intuition that much can be learned about individuals from the spaces they inhabit. Purely on the basis of some laundry receipts, an unfinished letter in the wastebasket, an empty bottle of bourbon, and other assorted clues, Steinbeck pieced together a portrait of Lonesome Harry. It would seem that the environments that people craft around themselves are rich with information about their personalities, values, and lifestyles.

Interactionist theories (Buss, 1987; Snyder & Ickes, 1985; Swann, 1987) suggest that individuals select and create their *social* environments (e.g., friendships, social activities) to match and reinforce their dispositions, preferences, attitudes, and self-views; extraverts choose friends, colleagues, and relationship partners who enable them to express their extraverted nature. Consistent with Steinbeck's intuition, we hypothesize that individuals also select and craft *physical* environments that reflect and reinforce who they are. Furthermore, we propose that observers use the information available in everyday environments to form impressions of what the occupants of those environments are like.

The links between occupants and their personal environments and between personal environments and observers' perceptions of the occupants can be conceptualized in terms of Brunswik's (1956)
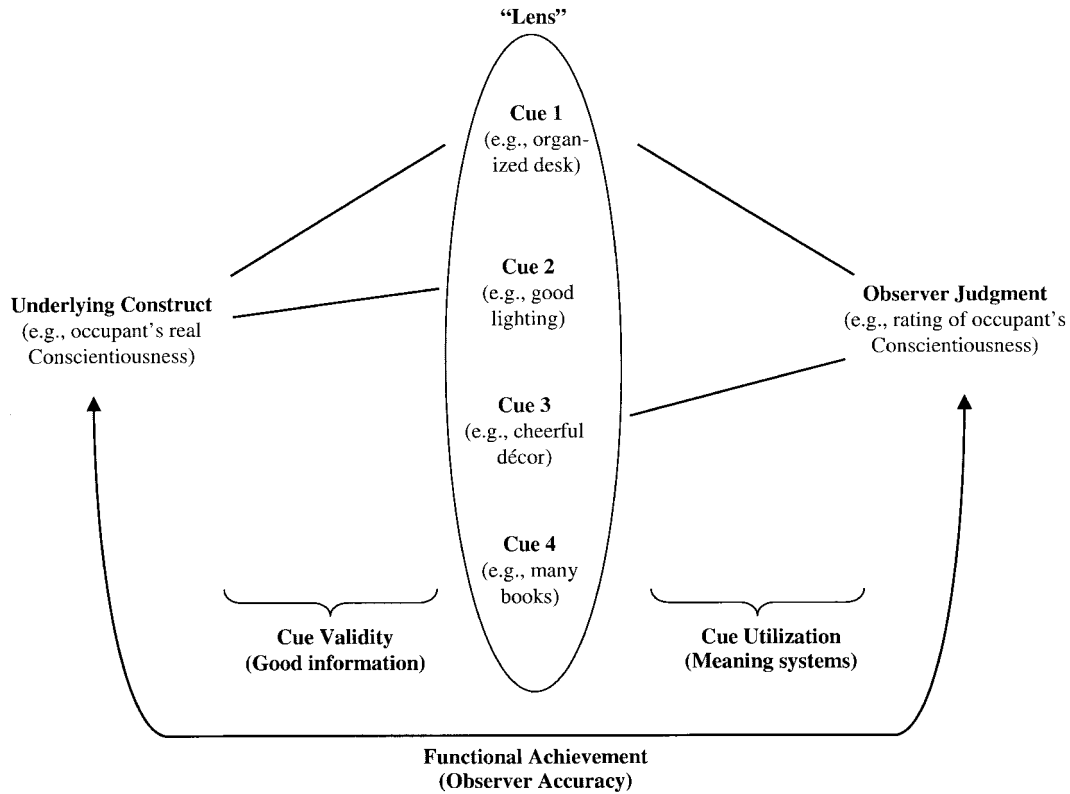
Figure 1. Brunswik's (1956) lens model.

lens model (see Figure 1). According to Brunswik, elements in the environment can serve as a kind of lens through which observers indirectly perceive underlying constructs. For example, an organized desk could serve as the lens through which an observer perceives an occupant's high level of Conscientiousness. In Brunswik's model, *cue utilization* refers to the link between the observable cue (e.g., the organized desk) and an observer's judgment (e.g., of conscientiousness). The link between the observable cue and the occupant's actual level of the underlying construct is referred to as *cue validity*. If both of these links are intact, then observer judgments should converge with the underlying construct being observed and will result in observer accuracy.[1] Thus, in terms of the example provided in Figure 1, an organized desk (Cue 1) will promote accuracy because it is both utilized and valid.

We draw on the logic of Brunswik's (1956) lens model first to explain the specific mechanisms by which individuals impact their physical surroundings and then to explain how personal environments can serve as repositories of individual expression from which observers can draw inferences about the occupants. As shown in Figure 2, our model proposes four parallel mechanisms linking individuals to the environments they inhabit and a two-step inference process linking environments to observer perceptions of the occupants.

Mechanisms Linking Individuals to the Environments They Inhabit

According to our model, the mechanisms linking individuals to the environments they inhabit fall into two categories: identity claims (self-directed and other directed) and behavioral residue (interior and exterior).

Self-Directed Identity Claims

People spend many of their waking hours in their personal living and work environments, and they often decorate these places. People choose colors, patterns, motifs, and décor that fit their own personal taste and aesthetic. To make these spaces their own, individuals may adorn them with self-directed *identity claims*— symbolic statements made by occupants for their own benefit, intended to reinforce their self-views. Many of these statements can make use of widely understood cultural symbols (e.g., a poster of Martin Luther King, university memorabilia), whereas other artifacts may have a more personal meaning (e.g., a pebble collected from a favorite beach). These latter objects can, nonetheless, convey a message to an observer even if the exact meaning of such private artifacts is obscure; for example, the pebble could signify that the occupant is sentimental or values nature.

Other-Directed Identity Claims

In addition to reinforcing their own self-views, occupants can display symbols that have shared meanings to make statements to others about how they would like to be regarded (Baumeister,

---

[1] Accuracy is referred to as functional achievement in Brunswik's (1956) model.

Mechanism	Underlying Disposition	Manifest Behavior	Residue or Evidence of Behavior	
1. Self-directed identity claims	Sentimental →	Collect memorabilia →	Pebble from a favorite beach	
2. Other-directed identity claims	Socially responsible →	Vote →	"I voted" sticker	CUE VALIDITY
3. Interior behavioral residue	Conscientious →	Organize belongings →	Organized workspace	
4. Exterior behavioral residue	Sensation seeking →	Go snowboarding →	Snowboard in room	

Step 2a: Observers infer disposition from behavior

OR

Step 2b: Behavior activates stereotype and associated traits

Stereotype (e.g., Asian)

↓

Suite of dispositions (e.g., socially responsible, conscientious, & introverted)

Step 1a: Observers infer behavior from residue

OR

Step 1b: Residue activates stereotype and associated traits
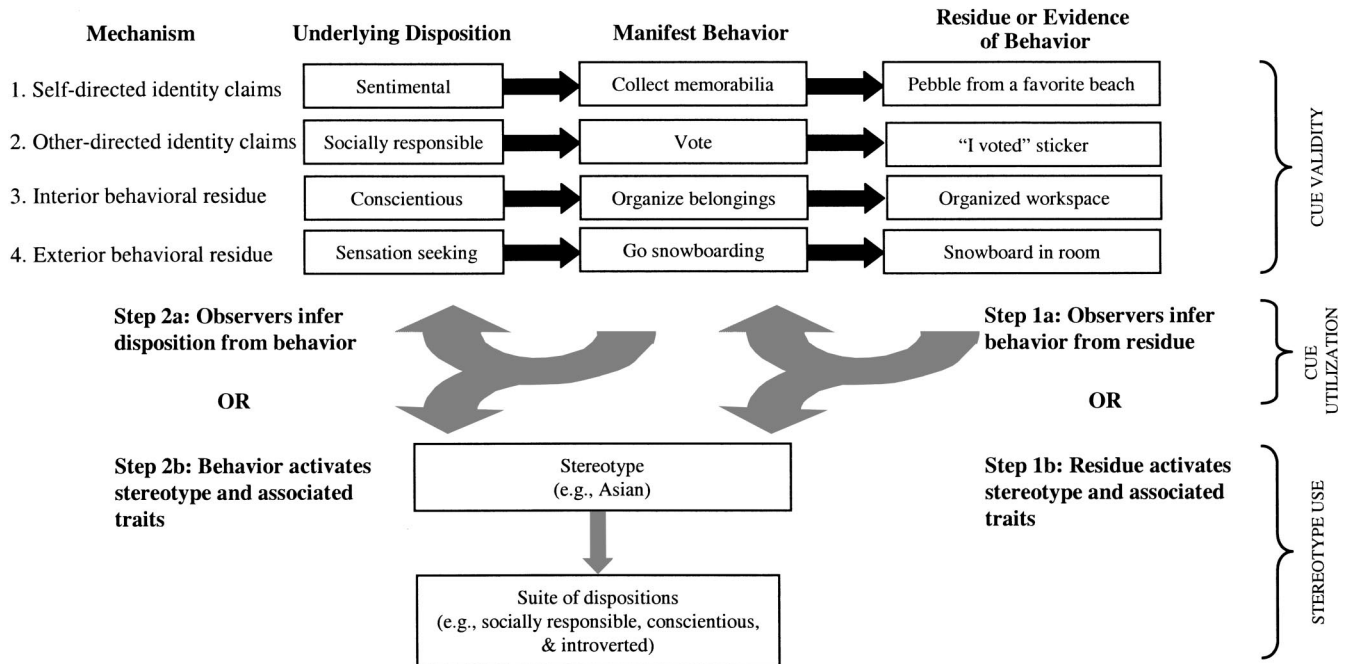
CUE UTILIZATION

STEREOTYPE USE

Figure 2. Mechanisms by which individuals impact their personal environments and processes by which observers infer personality.

1982; Goffman, 1959; Swann, 1987; Swann, Rentfrow, & Guinn, in press). By displaying such symbols (e.g., a poster of Martin Luther King, university memorabilia), occupants may be intentionally communicating their attitudes and values to others. These statements might be sincere and intended to convey truthful messages about what the individual is really like, but they may also be strategic, even deceptive statements intended to portray the individual in a certain light. For example, the occupant of an office who has frequent client visits may be motivated to decorate the space in ways that suggest to others a conscientious, hard-working, and honest disposition. Similarly, the occupant of a dorm room may be motivated to decorate his or her space with posters of rock stars and hip movies to be seen as cool by potential visitors.

We note that self-directed and other-directed identity claims may often result in similar environmental manifestations; for example, displaying a poster of Martin Luther King may serve both to reinforce one's self-view and also to communicate one's values to others. Our goal here is to note that although the outcomes of self-directed and other-directed identity claims may often overlap, they nonetheless reflect conceptually distinct motivations.

Interior Behavioral Residue

The act frequency approach to personality (Buss & Craik, 1983) defines personality traits in terms of behavioral conduct; hence, a person who is high on a particular trait would perform more acts that are prototypical of that trait than would a person who is low on the trait. Personal environments are places where individuals spend a great deal of time, and certain behaviors are performed repeatedly in these environments. Many of these behaviors leave behind discernible residue (Webb, Campbell, Schwartz, &

Sechrest, 1966). For example, the prototypically conscientious act of organizing one's space may leave an alphabetized CD collection as its residue, whereas the creative act of drawing may leave a set of the occupant's charcoal sketches lying on the floor as its residue. We use the term *behavioral residue* to refer to the physical traces of activities conducted in the environment. Although most cues will reflect past behaviors, there may also be some clues to anticipated behaviors; for example, an unopened bottle of wine and a set of beanbag chairs arranged in a circle on the floor may indicate a social occupant who is planning to entertain guests. We propose that personal environments such as offices and bedrooms are good receptacles of such interior behavioral residue.

Exterior Behavioral Residue

Just as personal environments may contain residue of behaviors conducted within that space (i.e., Mechanism 3 in Figure 2), they may also contain residue of behaviors performed by the individual entirely outside of those immediate surroundings. Remnants of past activities and material preparations for planned activities to be undertaken outside of the physical environment can also provide behavioral information about individuals. For example, a sensation seeker's space might contain a snowboard and a ski pass or perhaps even some parachuting equipment. A program from a recent opera and a plane ticket could suggest that the occupant enjoys the arts and likes to travel. The distinction between interior and exterior behavioral residue is important because it emphasizes the breadth of information that is potentially available in a personal space, extending to behavior occurring well beyond the limits of its four walls.

Again, we emphasize that these four mechanisms are not mutually exclusive, and it may not always be clear to observers which mechanisms are responsible for which cues. For example, the snowboard may indeed reflect exterior behaviors, but the occupant's decision to display the snowboard (rather than stow it in a closet) may also reflect a desire to make identity claims. Nonetheless, we propose that these are four distinct, albeit interlinked, mechanisms by which individuals impact their environments in ways that can guide observers to form impressions of personality.

Processes Linking Observer Judgments to Environmental Cues

Each of the four mechanisms above could provide information about an occupant's personality. As we show in the lower section of Figure 2, we propose that observers form impressions about occupants using a two-step inference process, with the possibility that stereotypes could intervene at either step. According to our model, if stereotypes do not intervene, observers first infer the behaviors that created the physical evidence (Step 1a), then infer the dispositions that underlie the behaviors (Step 2a). For example, they may infer from the organized workspace that the occupant organizes his or her belongings, and the observers may associate such behaviors with high levels of Conscientiousness. However, stereotypes might be activated by residue in the environment (Step 1b) or by inferred behaviors (Step 2b). In such cases, judgments about an occupant could be impacted by a stereotype that is associated with a whole set of traits, some of which may have no direct link to evidence in the environment. For example, an observer might notice some Asian books on the shelves, which might activate stereotypes about Asians. This may result in the observer inferring a suite of traits stereotypically associated with Asians, such as being socially responsible, conscientious, and introverted. When multiple observers hold similar stereotypes, as they might for common stereotypes such as sex and race, consensus among observers could be boosted. And when there is some truth to these stereotypes, accuracy could be boosted (Lee, Jussim, & McCauley, 1995). Note that stereotype-based inferences (Steps 1b and 2b) differ from Steps 1a and 2a in that observers using stereotypes may draw conclusions about traits for which they have no direct evidence.

Regardless of whether stereotypes intervene in the inference process, if observers make similar inferences, then consensus among observers' impressions should result. And if observers make correct inferences, then those impressions should also be accurate. However, observers may form inaccurate impressions by incorrectly matching evidence to the behaviors that caused them or by incorrectly matching the behaviors to the underlying dispositions. Observers' use of invalid stereotypes would also diminish accuracy.

Research Questions

Before we can test the specific processes hypothesized in our model, it is first necessary to document the broader basic phenomena that underlie it. The purpose of the present article is to achieve this initial goal. Thus, our primary aim was to document evidence in personal environments for observer accuracy, links between occupants and the physical features of their personal environments

(i.e., cue validity), and links between physical features of personal environments and observers' impressions of occupants (i.e., cue utilization). To examine the accuracy of observer judgments based on personal environments, we needed to test whether observers' judgments corresponded to what the individual was really like. Answering this question entailed first establishing whether observers' judgments corresponded with each other. Thus, we started by examining interobserver agreement. We could then move on to examine accuracy—testing whether observer impressions were correct. Third, we examined the links between observer impressions and features of the personal environments on which the impressions were based (i.e., cue utilization) and between the environments and what the occupants were really like (i.e., cue validity). Finally, as a preliminary step to understanding the inference processes in our model, we examined the role of stereotype use in accounting for interobserver agreement and observer accuracy.

We addressed these issues in two settings in which people spend a lot of their time: offices and bedrooms. As we describe below, these settings provide two diverse and potentially rich contexts in which to examine environment-based perceptions and the processes that link persons to the spaces they inhabit. We next describe the four questions that guided our research.

Question 1. Consensus: Do Observers Agree About Individuals' Personalities on the Basis of Their Personal Environments?

Past research on social perception has examined impressions of competence, ability, and personality formed on the basis of momentary impressions (Ambady & Rosenthal, 1992; Blackman & Funder, 1998; Borkenau & Liebler, 1992; Chaplin, Phillips, Brown, Clanton, & Stein, 2000; Paulhus & Bruce, 1992), photographs (Berry & Finch-Wero, 1993; Robins, Gosling, & Donahue, 1997), and attire (Burroughs, Drews, & Hallman, 1991; R. D. Gosling & Standen, 1998). One consistent finding to emerge from this research is that observers can agree on some aspects of what a target individual is like even when they base their judgments on minimal information. In a meta-analysis of nine of these so-called "zero-acquaintance" studies, the consensus correlations among observers averaged .12 (ranging from .03 to .27) across the Five-Factor Model (FFM) personality dimensions (Kenny, 1994; Kenny, Albright, Malloy, & Kashy, 1994).[2]

Where do observations based on personal environments fit with the previous zero-acquaintance research? In some sense, judgments based purely on personal environments are made under conditions of even less acquaintance than are judgments used in typical zero-acquaintance studies, because personal environments provide absolutely no direct exposure to the targets. However, we propose that personal environments may contain an abundance of potentially informative cues about an individual and therefore provide richer information to observers than is found in most

---

[2] Consistent with common usage in the literature (e.g., Kenny, 1994), we use the term *zero-acquaintance* to refer to studies examining impressions based on minimal exposure to the targets. However, it should be noted that the term is somewhat of a misnomer, because the judges in such studies have at least some information on the targets.

zero-acquaintance studies. Therefore, we expected to find significant consensus among observer judgments based on the physical characteristics of occupants' personal environments. To test this expectation, we examined the degree to which observers formed similar impressions on the basis of work and living spaces.

According to past zero-acquaintance studies, observer consensus is not equally strong for all traits judged (Paulhus & Bruce, 1992). Kenny's (1994) meta-analysis showed that, by far, the strongest consensus was obtained for Extraversion, with Conscientiousness a distant second, and the least consensus found for Agreeableness. These findings suggest that photos, short video clips, and brief interactions provide observers with more information about some traits than about others. Perhaps observers judge Extraversion on the basis of the target's facial expressions (e.g., a smile) and posture (e.g., an erect stance) and judge Conscientiousness from the target's level of grooming (e.g., well-combed hair). However, the informational base of previous zero-acquaintance research differs substantially from the informational base of the present research, so we did not expect to replicate the exact pattern of consensus correlations. Instead, we expected physical spaces to hold more cues to an occupant's level of organization (e.g., from alphabetized books and compact discs), tidiness (e.g., a neat vs. messy space), values (e.g., a poster supporting the legalization of marijuana), and recreational pursuits (e.g., tickets to the opera). The availability of such cues should promote relatively strong consensus for observers' judgments of Conscientiousness and Openness to Experience. More generally, we expected that physical spaces would provide observers with more information about some traits than about others, with the result that consensus would vary across the traits judged.

Question 2. Accuracy:
Are Observers' Impressions Correct?

Zero-acquaintance studies have shown that even judgments based on minimal information show some accuracy (e.g., Blackman & Funder, 1998; Paulhus & Bruce, 1992). Kenny's (1994) meta-analysis of 10 zero-acquaintance studies showed that across the FFM dimensions, observer judgments correlated .25 with criterion measures (self-reports).[3] Again, we propose that personal environments are richer in information than are zero-acquaintance contexts because environments probably hold many legitimate cues to what an occupant is like. Therefore, we expected significant accuracy correlations for observer judgments. To test this idea, we compared the observers' ratings with criterion ratings that were derived from self- and peer reports of the target occupants.

The zero-acquaintance research has shown that observer accuracy is not equally strong for all traits judged. In terms of the dimensions of the FFM, the greatest accuracy has generally been found for Extraversion and Conscientiousness and the least accuracy found for Openness and Agreeableness (Borkenau & Liebler, 1992; Funder & Dobroth, 1987; Kenny, 1994; Paulhus & Bruce, 1992). As noted above, the information available in personal environments differs from the information available in brief interactions, so we did not expect our findings to mirror the exact pattern found in zero-acquaintance research. However, we did expect variation in the levels of accuracy across the traits judged.

Question 3. Cue Utilization and Cue Validity: Which Cues in Personal Environments Do Observers Use to Form Their Impressions, and Which Cues Are Valid?

What judgment processes might explain why observers unacquainted with the occupants show consensus and accuracy in their impressions? Kenny's (1994) Weighted-Average Model (WAM) of consensus proposes nine parameters to explain the extent to which observers agree about a target individual. One of these WAM parameters, *similar meaning systems*, refers to the degree to which observers agree on the meaning of information (Kenny, 1994). For example, do observers agree that a poster of Martin Luther King and neatly stacked papers indicate that an occupant has liberal values and is conscientious? If observers notice the poster and the stacked papers, agree on the behavioral implications of this physical evidence, and agree on what the behaviors say about the occupant, then interobserver consensus should be strong (Hayes & Dunning, 1997). To determine which aspects of personal environments observers might have used to make inferences about the occupants, we correlated the observer judgments with coded features of the environments.

Funder's (1995, 1999) Realistic Accuracy Model (RAM) proposes that observer accuracy will be promoted when observers use *good information*. That is, accurate judgments should result when observers base their judgments on information that is actually related to the criterion. If underlying dispositions are actually related to the physical evidence in the personal environment (through manifest behavior; see Figure 2), then this evidence provides good information about the occupant. To determine which aspects of personal environments furnish good information, we correlated the criterion measures that tell us what the occupant is really like with coded features of the environments.

The WAM concept of meaning systems and the RAM concept of good information can be brought together by interpreting the concepts as the two halves of Brunswik's (1956) lens model (see Figure 1). Recall that cue utilization refers to the relation between judgments and observable information in the environment and that cue validity refers to the relation between the criterion and observable information in the environment. Thus, cue utilization is similar to the WAM parameter of meaning systems, and cue validity is similar to the RAM parameter of good information.

The lens model can represent all combinations of cue utilization and cue validity, revealing sources of good and bad judgments (Funder & Sneed, 1993; Gifford, 1994). Good judgment is promoted when observers use valid cues (represented by Cue 1 in Figure 1) and when they ignore invalid cues (Cue 4). Poor judgments occur when observers ignore valid cues (Cue 2) or use invalid cues (Cue 3). Thus, we hypothesized that accuracy would be promoted when observers use good cues and disregard bad ones. We tested this hypothesis by comparing the pattern of cue-utilization correlations with the pattern of cue-validity correlations for each trait.

---

[3] Note that Kenny (1994) made a distinction between self–other agreement and accuracy. Although Kenny's preferred criterion for accuracy was behavior, he acknowledged there are many cases in which behavioral observations are not practical and in which self-reports should be used.

Question 4. Stereotype Use: How Do Stereotypes Used by Observers Affect Consensus and Accuracy?

Kenny's (1994) WAM also places substantial emphasis on the role stereotypes play in observer judgments, with three of the WAM parameters focusing on stereotype use. According to WAM, consensus among observers will increase to the extent that observers hold similar stereotypes and use them in their judgments. Suppose that observers share the stereotype that women are more agreeable than men; if the observers believe Personal Environment A belongs to a female occupant and Personal Environment B belongs to a male occupant (using Steps 1b or 2b in Figure 2), then interobserver consensus may increase, because the observers will tend to rate the occupant they believe to be female as more agreeable than the occupant they believe to be male. To test the effects of sex and race stereotypes on observers' judgment, we examined the extent to which observers rated occupants they believed to belong to a given social category (e.g., female, White) differently from occupants they believed did not belong to that category.

Sex and race stereotypes might also account for some of the accuracy of judgments based on personal environments. The WAM parameter *validity of stereotypes* suggests that when observers use a stereotype with a kernel of truth, accuracy will be promoted (also see Lee et al., 1995). For example, if observers hold the stereotype that women are more agreeable than men and if there is some truth to this stereotype, then observers using this stereotype should show improved accuracy. To determine whether the sex and race stereotypes have a kernel of truth, we tested whether perceived sex and race differences for a given trait matched actual sex and race differences. To the extent that actual differences for each trait correspond to the perceived differences, accuracy should result. Thus, on the basis of the WAM, we predicted that observer accuracy would be mediated by the observers' use of accurate stereotypes. However, we anticipated that accurate stereotypes would only partially mediate observer accuracy because personal environments may contain valid, individuating information about what the occupant is like (through Steps 1a and 2a in Figure 2) beyond the accuracy associated with the use of valid sex and race stereotypes.

Design of the Studies

We examined the above four questions in offices and bedrooms, settings that may facilitate the accumulation of behavioral residue and permit other forms of self-expression. In Study 1, we examined perceptions of occupants based on their office spaces. Particular activities are consistently performed in offices, permitting the accumulation of residue for work-related traits. In Study 2, we examined perceptions based on personal living spaces such as apartment rooms and dorm rooms. These contexts permit a broader range of activities and self-expression than are permitted in offices. In both studies we collected four sources of data.

Observer Judgments

A team of observers made personality ratings of occupants based solely on an examination of occupants' personal environments. The observers were undergraduate students working on the project as research apprentices. The observers were unacquainted with the participants and did not discuss their ratings with one another. We obtained consensus estimates by computing the mean correlation among the observers' ratings.

Accuracy Criteria

To derive a criterion measure against which the accuracy of the observer reports could be gauged, we obtained self-ratings from occupants and peer ratings from the occupants' close acquaintances. We obtained accuracy estimates by correlating the observers' ratings with the combined self- and peer ratings.

Environmental Cues

A separate team of coders examined and recorded the features of each room. We obtained cue-utilization estimates by correlating these codings with the observer judgments. We obtained cue-validity estimates by correlating the codings with the accuracy criteria. Finally, we computed vector correlations between the cue-utilization correlations and the cue-validity correlations. These vector correlations provide evidence as to whether the cues observers used to make their judgments correspond to the cues that are actually related to the occupants' traits.

Sex and Race of Occupants

Observer estimates of the sex and race of occupants were used to assess the mediational role of sex and race stereotypes on interobserver consensus. Self-reported sex and race were used to examine the extent to which sex and race stereotypes mediated observer accuracy.

Study 1: Offices

Method

Research Setting

Five office locations in a large U.S. city were chosen. These locations were a commercial real estate agency, an advertising agency, a business school, an architectural firm, and a retail bank.

Occupants

With the consent of management, employees at each company were given an opportunity to participate in this research. In exchange for their participation, occupants received feedback based on observers' impressions of their offices. Ninety-four office occupants participated in this study. Of the occupants who provided the relevant information, 41 (59%) were women and 28 (41%) were men, and the average age was 37.0 years ($SD = 10.1$). The occupants' ethnicities were not sufficiently diverse to analyze race effects in this sample (of those who indicated their race, 3 [5%] were Asian, 57 [85%] were White, and 7 [10%] were of other ethnicity).

Observer Ratings

Eight observers examined the participants' personal workspaces and completed a set of ratings about each occupant.[4] We wanted to learn about everyday impressions made by ordinary people (rather than impressions formed by trained experts), so observers were given no instructions regarding what information they should use to make their ratings. The observers had no contact with the participants and made their judgments independently after entering the offices, using whatever information they thought was relevant. The number of observers in an office at any time ranged between 1 and 5, depending on the size of the office. The observers did not communicate with one another. All photos of occupants and references to occupants' names were covered before the observers entered the rooms.

Accuracy Criteria

The best method for assessing a psychological construct is to search for converging evidence across assessment methods (Wiggins, 1973). For example, an approach that uses a combination of self- and peer reports can overcome some of the biases of either method used alone (John & Benet-Martinez, 2000; McCrae, Stone, Fagan, & Costa, 1998). Thus, we obtained self- and peer reports on all of the dimensions examined in this study. Specifically, after providing self-ratings, occupants were asked to nominate two people who knew them well and could complete the peer ratings. The peers were sent the rating scales, which they completed confidentially and mailed back to us in return addressed, postage-paid envelopes. We were able to obtain 69 self-reports and ratings by one or two peers for 60 of the occupants (average acquaintance with occupant = 8.3 years, SD = 9.7). Averaged across the five dimensions examined in this study, the self-ratings correlated .40 with the peer ratings; this value is comparable to that reported in previous research (e.g., Funder, 1980; John & Robins, 1993; McCrae et al., 1998). Although these agreement correlations were far from perfect, the self- and peer reports each provide valuable information about the occupants and were combined to form an accuracy criterion.[5] Reliability of the criterion measures was computed for the composite of the self-report and two peer reports; alphas averaged .61 across the five dimensions examined.

Sex of Occupants

Observers estimated the sex of the occupants. The actual sex of the occupants was obtained from the occupants' self-reports.

Environmental Features

Access to the office spaces was granted for quite limited periods of time. To minimize the time taken to code the environmental features, we used two teams of two coders each. Once a team finished coding an office, it moved on to the next office. One team of coders coded 49 offices, and the other team coded 45. For each office, both coders independently coded the space in terms of 43 environmental features (e.g., ratings of neatness and organization) that had been consensually selected to represent a broad range of environmental attributes. The 43 items were selected using extensive item-generation and selection procedures, the details of which can be obtained from the first author. To control for differences in scale use by the two teams of coders, we standardized their codings within each team before combining and aggregating them into composites. The codings were reasonably reliable, with a mean coefficient alpha of .63 across the 43 features.

Instruments

Observer, self-, and peer reports of personality were based on the FFM (McCrae & Costa, 1999). The FFM is a hierarchical model with five broad factors that represent personality at the broadest level of abstraction. Each bipolar factor (e.g., Extraversion vs. Introversion) summarizes several more specific facets (e.g., Sociability), which, in turn, subsume a large number of even more specific traits (e.g., talkative, outgoing).

The FFM dimensions were assessed using the 44-item Big Five Inventory (BFI; John & Srivastava, 1999). The self-report version of the BFI shows high convergent validity with other self-report scales and with peer ratings of the FFM. The BFI items were rated on a 7-point Likert scale ranging from 1 (*disagree strongly*) to 7 (*agree strongly*). In the present sample, alpha reliabilities for Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experience averaged .88, .86, .85, .78, and .92, respectively, across the eight observers; averaged .85, .84, .86, .83, and .83, respectively, across the two peers; and were .84, .76, .75, .79, and .81, respectively, for the self-reports. These values are typical of those reported for the BFI (John & Srivastava, 1999).

Results and Discussion

Question 1: Consensus

Based on past zero-acquaintance research, we predicted that observers would agree about the occupants' personalities solely on the basis of the occupants' workspaces. To test this prediction, we computed an index of interobserver consensus by taking the mean of all 28 possible pairwise correlations among the eight observers. Across the traits, the mean of the 28 pairwise correlations was positive and significant (r = .34), supporting our prediction.

We further predicted that interobserver consensus would vary across traits. To test this hypothesis, we computed interobserver consensus separately for the five dimensions. The mean pairwise consensus correlations for each trait are shown in the first data column of Table 1. As can be seen, interobserver consensus varied substantially across the traits. Of the FFM dimensions, Openness showed the strongest consensus, followed by Conscientiousness and Extraversion. Agreeableness also showed some consensus, with the least consensus found for Emotional Stability.

The pattern of consensus correlations found here is different from that found in previous zero-acquaintance research, suggesting that the cues available from photographs or short interactions differ from those available from workspaces. Whereas zero-acquaintance research using the FFM has found the strongest consensus for Extraversion and Conscientiousness (e.g., Kenny, 1994; Kenny et al., 1994; Kenny, Horner, Kashy, & Chu, 1992; Park & Judd, 1989), we found the strongest consensus for Openness, although we still found strong consensus for Conscientiousness and Extraversion.

Question 2: Accuracy

We obtained an index of accuracy by correlating the aggregated observer ratings with the composite criterion ratings. The aggregate of observer ratings is independent of the idiosyncrasies of any

---

[4] In some cases only seven observers were present because of scheduling difficulties.

[5] An alternative way of gauging accuracy is to use the self- and peer reports as separate criteria rather than combining them. However, the biases of each of these sources of information could skew the accuracy findings. For example, the self-reports are probably saturated with other-directed identity claims, whereas peer reports partly reflect observations of behaviors that leave residues in the rooms.

Table 1
Judgments Based on Offices: Consensus, Accuracy, and Column–Vector Correlations

FFM personality dimensions	Question 1: Interobserver consensus (Mean $n = 76$)[a]	Question 2: Observer accuracy ($n = 70$)	Question 3: Vector correlations ($n = 43$)
Extraversion	.39**	.24*	.36*
Agreeableness	.23*	−.04	−.08
Conscientiousness	.42**	.24*	.80**
Emotional Stability	.14	.19	.09
Openness to Experience	.51**	.46**	.60**

Note. Interobserver consensus is the mean of the 28 correlations derived from all possible pairwise combinations of eight observers. Observer accuracy is the correlation between the aggregated observer ratings and the composite criterion ratings. The vector correlations reflect the convergence between the cue-utilization correlations and the cue-validity correlations. FFM = Five-Factor Model.
[a] Significance of consensus correlations was based on sample size of 76, the average number of cases across which the correlations were computed. When consensus was determined from intraclass correlations using the subset of rooms for which there was no missing data, the pattern of findings was almost identical, with all five consensus correlations reaching significance at the .01 level.
* $p < .05$, one-tailed.   ** $p < .01$, one-tailed.

single observer and thus provides the most reliable measure of how much strangers learn about an occupant by observing the occupant's personal workspace (Block, 1961). The accuracy findings are presented in the second data column of Table 1. Consistent with our hypothesis, the accuracy correlations were positive and significant—indeed, considering that the observers had absolutely no direct contact with the occupants, they were quite substantial—averaging .22 across the five dimensions examined.

Also consistent with our predictions, the level of accuracy varied across traits. Accuracy was highest for Openness, followed by Extraversion, Conscientiousness, and Emotional Stability. Judgments of Agreeableness were not accurate. Again, the pattern of findings differed from the general pattern of findings in zero-acquaintance research, in which Extraversion usually yields the highest accuracy, followed by Conscientiousness, with little or no accuracy for Emotional Stability and Openness (Kenny, 1994; Watson, 1989).

Question 3: Cues

Drawing on the logic of Brunswik's (1956) lens model, we conducted a series of analyses to test (a) the extent to which observers used physical characteristics of the rooms to make inferences about occupants' personalities, and (b) the extent to which the physical characteristics of the rooms were related to what the occupants were really like.

Cue utilization. The cue-utilization correlations in the right-hand section of Table 2 reflect the relationships between the aggregated observers' ratings and the coded physical features of the offices. These cue-utilization correlations show which cues may have acted as Brunswikian lenses through which the observers perceived characteristics of the occupants. It is not surprising that the personality dimensions that achieved the strongest consensus also yielded the most associations with specific cues in the work environment. Thus, Openness and Conscientiousness had the largest number of significant cue-utilization correlations.

Which particular cues might we expect observers to use? Within the FFM framework, Conscientiousness is associated with order,

efficiency, and self-discipline. Therefore, it would be reasonable for observers to expect the office of a conscientious person to be organized, efficiently arranged, clean, and uncluttered. The cue-utilization correlations suggest that the observers indeed used these types of cues, with their judgments of Conscientiousness correlating .56, .44, .73, and −.55 with good use of space, clean, organized, and cluttered, respectively. As shown in Table 2, the observers drew on a wide range of cues to infer the occupants' levels of Conscientiousness, most of which are consistent with the FFM definition of what it means to be conscientious (e.g., organized and tidy).

Individuals who are high in Openness tend to be curious, imaginative, and unconventional and to have wide interests. Thus, it makes sense that observers appeared to base their judgments of Openness on cues that reflect these traits, such as the distinctiveness of the spaces (.60), the level of decoration (.49), the quantity of magazines (.34), and the quantity and variety of books (.28 and .44) and compact discs (.32 and .61).

Extraversion is associated with sociability, enthusiasm, talkativeness, and assertiveness. It is difficult to imagine which concrete elements of physical environments might reflect these traits. As Table 2 shows, offices that were, among other things, decorated (.48), cheerful (.47), colorful (.46), cluttered (.24), inviting (.35), and unconventional (.41) were believed to belong to extraverted occupants. Few of the coded cues were related to observers' ratings of Agreeableness or Emotional Stability.

As has been noted by Funder and Sneed (1993), such correlational analyses must be interpreted cautiously until future experimental research can address two limitations. First, although the correlations showed that observers' judgments were associated with the presence of certain cues, the correlations did not show that the observers actually used these cues to make their judgments. Second, the present analyses did not have the power to assess (e.g., by multiple regression) the degree to which the environmental cues overlapped or were used independently by the observers.

Cue validity. The correlations in the left-hand section of Table 2 reflect the relationship between the accuracy criterion and the

Table 2
A Brunswik (1956) Lens Model Analysis of Judgments Based on Offices: Cue-Validity and Cue-Utilization Correlations

Cue-validity correlations						Cue-utilization correlations				
Extra.	Agree.	Cons.	Em. St.	Open.	Environmental cues ("lens")	Extra.	Agree.	Cons.	Em. St.	Open.
.11	−.03	−.02	.03	.12	Strong (vs. weak) odor	.12	−.21*	−.04	−.11	.01
.21	−.05	−.09	−.02	.15	Noisy (vs. quiet)	.14	.14	−.16	−.03	.21*
−.04	.00	.07	−.04	−.10	Well lit (vs. dark)	−.03	−.12	.02	.00	−.08
.10	.00	.09	.14	.14	Drafty (vs. stuffy)	−.03	−.06	−.02	.02	−.03
.10	−.04	.00	.11	.24*	Fresh (vs. stale)	.03	.00	−.08	−.04	.06
.26*	−.09	−.04	−.16	.01	Hot (vs. cold)	.12	.06	−.01	−.05	.21*
.02	.11	.25*	−.01	.05	Good (vs. poor) condition	.01	−.01	.57**	.20	.01
.27*	−.05	−.06	−.26*	.04	Decorated (vs. undecorated)	.48**	.09	−.12	.13	.49**
.22	.03	.05	−.22	.17	Cheerful (vs. gloomy)	.47**	.19	−.03	.09	.44**
.15	.06	.03	−.15	.16	Colorful (vs. drab)	.46**	.16	.02	.07	.45**
.01	.22	.24*	.16	−.08	Clean (vs. dirty)	.00	.09	.44**	.11	−.13
.14	.14	.35**	.07	−.02	Organized (vs. disorganized)	−.09	.08	.73**	.15	−.14
.17	.12	.30*	.09	−.07	Neat (vs. messy)	−.13	.10	.73**	.15	−.19
−.19	.00	−.29*	−.07	−.04	Cluttered (vs. uncluttered)	.24*	−.03	−.55**	−.07	.23*
−.14	.14	−.18	−.09	−.03	Full (vs. empty)	.33**	.03	−.47**	−.10	.30**
.13	−.08	.14	−.04	.08	Roomy (vs. cramped)	−.06	−.05	.42**	.06	.07
−.01	.09	.20	.01	.14	Expensive (vs. cheap)	.07	−.03	.28**	.02	.25*
.21	.04	.15	−.07	.15	Comfortable (vs. uncomfortable)	.15	.07	.44**	.16	.28**
.29*	−.08	.07	−.18	.15	Inviting (vs. repelling)	.35**	.27**	.33**	.29**	.38**
−.11	−.03	−.08	−.11	.12	Large (vs. small)	.01	−.10	.33**	.06	.28**
.11	−.04	−.03	−.24*	.30*	Distinctive (vs. ordinary)	.42**	−.01	−.24*	−.02	.60**
.14	.06	.12	−.12	.26*	Stylish (vs. unstylish)	.34**	.05	.14	.10	.51**
.03	.21	.07	.21	.02	Modern (vs. old fashioned)	.23*	−.11	.01	.00	.22*
.03	.14	.09	−.02	.00	New (vs. old)	−.03	−.14	.15	−.02	.05
.03	.03	−.05	−.06	.05	Multiple (vs. single) purpose	.21*	−.05	−.08	−.09	.24*
−.11	.16	.01	.17	−.12	Public (vs. private)[a]	.02	.05	−.18	−.05	−.36**
.09	.03	.18	.27*	−.14	Formal (vs. informal)[a]	−.10	−.06	.47**	.11	−.14
−.08	−.11	.01	.08	−.24*	Conventional (vs. unconventional)[a]	−.41**	.00	.29**	.00	−.53**
.04	.13	−.03	−.03	.00	Centrally (vs. peripherally) located[a]	−.04	.12	.12	.11	−.24*
.03	.25*	.02	.14	−.06	High (vs. low) traffic area[a]	.05	.15	−.03	.03	−.24*
.06	.10	.13	.03	−.05	Good (vs. poor) use of space[a]	.12	−.01	.56**	.18	−.01
−.15	.22	.09	.01	−.10	Matched (vs. mismatched) contents[a]	.08	.02	.33**	.10	.03
−.07	−.06	−.12	−.19	.06	Many (vs. few) books	−.05	−.01	.02	.03	.28**
−.15	−.08	.10	−.14	−.14	Organized (vs. disorganized) books	−.11	.02	.32*	.13	−.09
.03	−.01	−.32*	−.23	.30	Varied (vs. homogenous) books	.26*	−.05	−.27*	−.12	.44**
.01	.11	−.01	.07	.16	Many (vs. few) magazines	.23*	−.02	−.17	.05	.34**
−.43*	−.14	.12	.14	−.04	Organized (vs. disorganized) magazines	−.18	.09	.32	.03	.13
.00	.33	−.14	−.07	.18	Varied (vs. homogenous) magazines	.26	−.01	−.42**	−.15	.17
.12	−.04	−.03	−.18	.22	Many (vs. few) CDs	.23*	.15	−.19	.03	.32**
.07	.08	.04	−.14	.28	Organized (vs. disorganized) CDs	−.25	.00	.11	−.30	−.01
−.06	.22	−.62*	−.26	−.14	Varied (vs. homogenous) CDs	.50*	−.17	−.47	−.08	.61**
.04	−.02	.02	.09	.04	Many (vs. few) items of stationery	.07	.06	−.04	−.15	−.04
−.28	.27	.12	.11	−.28	Organized (vs. disorganized) stationery	−.04	.05	.53**	.12	−.13

Note. Extra. = Extraversion; Agree. = Agreeableness; Cons. = Conscientiousness; Em. St. = Emotional Stability; Open. = Openness.
[a] Cue was assessed in Study 1 but not in Study 2.
* $p < .05$, two-tailed.   ** $p < .01$, two-tailed.

coded physical features of the offices. These cue-validity correlations show which office cues were actually related to characteristics of the occupants.

As discussed in the previous section, the cue-utilization correlations shown in Table 2 suggest that observers relied on intuitively sensible cues to judge occupants' Conscientiousness, such as the degree to which the offices were well-organized, neat, and uncluttered. The cue-validity correlations shown in Table 2 suggest that these intuitions had some merit; the rooms of conscientious individuals were well organized (.35), neat (.30), and uncluttered (.29), but conscientious persons' offices did not make particularly good use of the space.

The Openness cue-validity correlations also support the observers' intuitions that people who were high in Openness occupied distinctive (.30) and unconventional (.24) offices. Compared with introverts, extraverts' offices appeared to be crafted to encourage interaction; they were relatively warm (.26), decorated (.27), and inviting (.29). However, very few of the coded cues related to the occupants' levels of Agreeableness and Emotional Stability in an intuitively compelling way.

Intercorrelations of column vectors. To test more formally the extent to which observers' cue-utilization patterns matched the patterns of cue-validity correlations, we computed intercorrelations among the column vectors in Table 2 (see Funder & Sneed,

1993). Specifically, we first transformed the cue-utilization and cue-validity correlations using Fisher's r-to-z formula and then, for each trait, correlated the transformed correlations across the cues. The vector correlations thus obtained (shown in the last data column of Table 1) are useful because they permit one to gauge the overall congruence between the cue-utilization and cue-validity patterns.

If the physical features of the rooms accounted for observers' accuracy, then the vector correlations should be highest for the accurately judged traits. The two most accurately judged traits—Openness and Conscientiousness—were indeed characterized by substantial vector correlations. Similarly, the least accurately judged traits—Agreeableness and Emotional Stability—were characterized by weak vector correlations. However, the magnitude of accuracy correlations was not universally matched by the magnitude of vector correlations. If direct inferences based on cues do not account for the accuracy (i.e., Steps 1a and 2a in Figure 2), perhaps accuracy can be explained by the use of stereotypes (i.e., Steps 1b and 2b in Figure 2). We next examine the role of sex stereotypes in promoting consensus and accuracy.

Question 4: Stereotype Use

Consensus. To test whether the observers rated the occupants differently on the basis of perceived sex, we examined the effects of sex using a hierarchical multiple regression, entering the stereotype variable (i.e., sex) as a first step. If sex stereotypes did mediate consensus for a trait, the mean ratings of occupants perceived to be female should be significantly different from the mean ratings of occupants perceived to be male. The first two data columns under the *Perceived sex differences* heading in Table 3 show the mean values of the observers' ratings separately for occupants perceived by the observers to be female and male. Women were perceived to be significantly more agreeable and less emotionally stable than were men. This finding is consistent with commonly held sex stereotypes (Eagly & Mladinic, 1989; Spence, 1993). Although perceived sex differences in personality do not guarantee that observers' use of sex stereotypes contributed to consensus, they are consistent with the interpretation that observers guessed the sex of occupants and modified their ratings of

Agreeableness and Emotional Stability in the stereotypically appropriate direction. However, further research that examines the effects of stereotypes in the laboratory, manipulating the perceived sex of occupants, would qualify our real-world findings by providing an experimental test of our interpretations.

Accuracy. Using hierarchical multiple regression, we tested for sex differences in the criterion ratings, which are shown in Table 3. The only significant real sex difference was in Emotional Stability, with women lower than men.

By comparing the perceived mean ratings with the real mean ratings, we were able to assess the accuracy of the sex stereotypes used. If perceived sex differences match real sex differences, then this is consistent with the hypothesis that observers' use of valid stereotypes could at least partially mediate the accuracy for that trait. Our analyses suggest that the observers' use of sex stereotypes was appropriate in the case of Emotional Stability but not in the case of Agreeableness. However, it is important to consider the occupant effects presented in Table 3. These effects reflect the variance in observer ratings that is not accounted for by differences associated with the perceived sex of the occupants. As shown in Table 3, the occupant effect sizes were far greater than the sex effect sizes, suggesting that the sex stereotypes accounted for only a small proportion of the observers' consensus and accuracy.

Summary of Study 1 Findings

In Study 1, we set out to answer four central questions. Question 1 asked whether observers agreed in their ratings of target occupants based purely on the occupants' workspaces. We found that observers generally agreed but that agreement varied across the traits, with Openness, Conscientiousness, and Extraversion showing the strongest agreement and Emotional Stability and Agreeableness showing the least agreement. Question 2 asked whether the observer ratings were accurate. Again, we found that the ratings were generally accurate but varied across traits, with Openness, Conscientiousness, and Extraversion showing the strongest accuracy and Agreeableness and Emotional Stability showing little accuracy. Question 3 asked which cues in the workspaces observers used to form their impressions and which ones were valid. We identified a wide range of cues relating to observer

Table 3
Impact of Stereotypes: Perceived and Real Sex Differences Based on Office Spaces

	Sex differences												
	Perceived (i.e., observers' ratings)									"Real" (i.e., criterion ratings)			
	Mean rating[a]			Effect of perceived sex			Occupant effect			Mean rating			Effect of sex	
FFM personality dimensions	Women ($n = 322$)		Men ($n = 350$)	η	$F(1, 119)$	p	η	$F(131, 532)$	p	Women ($n = 105$)		Men ($n = 71$)	η	$F(1, 67)$	p
Extraversion	4.6	=	4.5	.03	.09	.76	.71	4.10	<.001[b]	5.2	=	5.1	.09	.53	.47
Agreeableness	4.6	>	4.4	.34	14.22	<.001[b]	.57	2.01	<.001[b]	5.3	=	5.7	.21	3.19	.08
Conscientiousness	5.2	=	5.1	.12	1.82	.18	.72	4.49	<.001[b]	5.6	=	5.4	.18	2.30	.13
Emotional Stability	4.3	<	4.4	.23	5.79	.02	.54	1.68	<.001[b]	4.1	<	4.7	.32	7.58	.01[b]
Openness to Experience	4.1	=	4.2	.04	.26	.61	.76	5.62	<.001[b]	5.6	=	5.4	.14	1.34	.25

Note. FFM = Five-Factor Model.
[a] Ratings were made on a 7-point scale.   [b] Probability remains ≤ .05 after Bonferroni correction.

ratings of Conscientiousness and Openness. We also identified a number of cues relating to what the occupants were really like, especially for Conscientiousness and Openness. Of the cues we assessed, observers seemed to use valid cues to judge Conscientiousness, Openness, and, to a lesser extent, Extraversion but not to judge Agreeableness and Emotional Stability. Question 4 asked whether observers' use of sex stereotypes could account for interobserver consensus and observer accuracy. Consensus and accuracy were only partially mediated by sex stereotypes, with sex stereotypes potentially accounting for some of the consensus about occupants' Agreeableness and Emotional Stability and for some of the accuracy of observer ratings of occupants' Emotional Stability.

It is interesting that observers seemed to rely on stereotypes for those traits for which the office environment provided the fewest clues (see the cue-utilization correlations in Table 2). This is consistent with research that suggests that the more individuating information observers have about a target, the less likely they are to use sex stereotypes in their judgments (Locksley, Borgida, Brekke, & Hepburn, 1980).

Study 2: Bedrooms

The results of Study 1 were promising but raised the question of how well the findings generalize to other personal environments. Although office spaces clearly contain behavioral residue and expressions of identity, they may provide a window into only a subset of traits because of the limited scope of activities and forms of self-expression that occur in office environments. It is possible, for example, that the corporate setting promotes the expression of work-related traits (e.g., Conscientiousness) but stifles the expression of non-work-related traits (e.g., Agreeableness).

Thus, in Study 2 our goal was to focus on an ecologically richer environment that may facilitate a broader range of individual expression and serve as the repository for a wider range of behavioral residue. For this purpose, we needed to examine spaces in which occupants spend much of their time and that are used for a variety of purposes, not just for working. Given these requirements, we examined student accommodations; these spaces are multipurpose, often serving as places to entertain friends, relax, sleep, eat, and study. In Study 2, our sample was from a large public West Coast university with an ethnically diverse student body, permitting us to broaden our analyses to examine both sex and race stereotypes.

Method

The methods and procedure for this study essentially duplicate those used in Study 1, with seven observers independently rating 83 occupants solely on the basis of their personal living spaces.

Research Setting

The personal living spaces were rooms in private houses, apartments, dormitories, co-ops, and Greek-system housing situated in an urban setting close to a large West Coast public university. Most of the rooms we studied had single occupants, but we did examine shared rooms in which the occupants' section of the room could be clearly demarcated.

Occupants

Eighty-three college students or recent graduates with an average age of 21.9 (SD = 2.8) years participated in this study. Of those who indicated their gender and ethnicity, 54 (69%) were women, 24 (31%) were men, 32 (42%) were Asian, 27 (36%) were White, and 17 (22%) were of other ethnicities. In exchange for their participation, occupants received feedback based on the observers' impressions of their rooms.

Observer Ratings

Seven observers examined the participants' personal living spaces and completed a set of ratings about each occupant.[6] As in Study 1, observers were given no instructions regarding what information they should use to make their ratings. The observers had no contact with the participants and made their judgments independently after entering the rooms, using whatever information they thought was relevant. The number of observers in a room at any time ranged between 1 and 6, depending on the size of the room. The observers did not communicate with one another. All photos and references to occupants' names were covered before the observers entered the rooms.

Accuracy Criteria

We obtained self-ratings from 78 of the occupants as well as peer ratings by one or two friends for 77 of the occupants (average acquaintance with occupant = 3.4 years, SD = 3.5). Agreement between self- and peer ratings of personality was strong; across the five dimensions examined in this research, the self- and peer ratings correlated .53. We again combined the self- and peer reports to form a composite accuracy criterion index. Reliability of the criterion measures was computed for the composite of the self- and two peer reports; alphas averaged .78 across the five dimensions examined.

Sex and Race of Occupants

Observers estimated the sex and race of the occupants. The actual sex and race of the occupants were obtained from the occupants' self-reports.

Environmental Features

A team of 3 coders independently coded each room in terms of 42 environmental features (e.g., ratings of neatness and organization). The cue-rating instrument (S. D. Gosling, Martin, Craik, & Pryor, 2001) was tailored for use in personal living spaces and therefore differed slightly from the cue instrument used in Study 1. For example, two items referring to the amount of clothing in the room were added, and seven variables thought to be primarily relevant to office spaces (e.g., private vs. public) were omitted.

Coder ratings were aggregated into composites. These composite codings showed reasonable levels of reliability, with a mean coefficient alpha of .72 across the 42 features. The composite codings were more reliable than the codings of Study 1, reflecting the fact that the study of living spaces used 1 more coder than did the study of workspaces.

Instruments

Observer, self-, and peer reports of personality were made in terms of the FFM using the BFI. The BFI items were rated on a 5-point Likert scale ranging from 1 (*disagree strongly*) to 5 (*agree strongly*). In the present sample, alpha reliabilities for Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experience averaged .91, .87, .92, .75, and .94, respectively, across the 7 observers; averaged .92, .86, .88, .87, and .87, respectively, across the two peers; and were .90, .72, .79,

---

[6] In some cases only six observers were present because of scheduling difficulties.

.86, and .84, respectively, for the self-reports. These values are typical of those reported for the BFI (John & Srivastava, 1999).

Results and Discussion

Question 1: Consensus

We computed an index of interobserver consensus by taking the mean of all 21 possible pairwise correlations among the seven observers. The mean pairwise consensus correlations for each trait are shown in the first data column of Table 4. Across the traits, the mean of the 21 pairwise correlations was positive and significant ($r = .34$).

As in Study 1, there was substantial variation across traits in interobserver consensus. Furthermore, the variation in consensus correlations for the FFM dimensions had a remarkable resemblance to the pattern of correlations in Study 1; the strongest consensus was found for Openness and Conscientiousness, followed by Extraversion and Agreeableness, and little consensus was found for Emotional Stability.

Question 2: Accuracy

We again derived an index of accuracy by correlating the aggregated composite observers' scores with the composite criterion scores. The accuracy findings are presented in the second data column of Table 4. The correlations averaged .37 across the five dimensions examined, suggesting that greater overall accuracy was obtained from bedrooms than from offices (for which the mean correlation was .22). To put the magnitude of these accuracy correlations in perspective, one should recall that observers had absolutely no direct contact with the occupants.

As in the study of office spaces, the variation in accuracy across the FFM dimensions differed from that reported in past zero-acquaintance research. As in Study 1, accuracy was strongest for Openness, followed by Conscientiousness and Emotional Stability. Accuracy was significant but weaker for Extraversion and Agreeableness.

Question 3: Cues

We again examined the role of physical cues in the judgment process.

Cue utilization. The cue-utilization correlations in Table 5 reflect the relationship between the aggregated observers' ratings and the coded physical features of the bedrooms. As with the offices, the cue-utilization correlations suggest that the bedroom observers used intuitively sensible cues to judge Conscientiousness, with their judgments of Conscientiousness correlating .61, .70, and −.56 with clean, organized, and cluttered, respectively. As in Study 1, the results shown in Table 5 suggest that the observers drew on a wide range of intuitively compelling cues related to cleanliness and organization to infer the occupants' levels of Conscientiousness. In addition, they rated the occupants of rooms that were comfortable and inviting as conscientious, perhaps reflecting a belief that a consistent and concerted effort is required to make a place pleasant.

The patterns of cue utilization for judgments of Openness also made sense, with observers apparently basing their judgments of Openness on the distinctiveness of the rooms (.35), the level of decoration (.35), and the quantity or variety of books, magazines, and compact discs.

Agreeableness is associated with interpersonal warmth, sympathy, modesty, and trust. Unlike offices, bedrooms contained many cues that were associated with observer judgments of Agreeableness. The correlations in Table 5 show that rooms that were cheerful (.66), colorful (.51), clean (.37), organized (.26), neat (.33), comfortable (.43), and inviting (.52) and did not have clothes strewn about (−.39) were believed to be occupied by agreeable individuals. It is difficult to imagine how all of these elements might relate to Agreeableness, but one possible explanation is that the observers could have based their ratings on a folk belief that like goes with like (Gilovich & Savitsky, 1996), perhaps concluding that pleasant people occupy pleasant rooms. Another possibility is that observers associate Agreeableness with an occupant's concern for the aesthetic comfort of visitors.

Table 4

Judgments Based on Bedrooms: Consensus, Accuracy, and Column–Vector Correlations

FFM personality dimensions	Question 1: Interobserver consensus (Mean $n = 68$)[a]	Question 2: Observer accuracy ($n = 79$)	Question 3: Vector correlations ($n = 42$)
Extraversion	.31*	.22*	.24
Agreeableness	.20	.20*	−.23
Conscientiousness	.47**	.33**	.79**
Emotional Stability	.08	.36**	.16
Openness to Experience	.58**	.65**	.80**

Note. Interobserver consensus is the mean of the 21 correlations derived from all possible pairwise combinations of seven observers. Observer accuracy is the correlation between the aggregated observer ratings and the composite criterion ratings. The vector correlations reflect the convergence between the cue-utilization correlations and the cue-validity correlations. FFM = Five-Factor Model.

[a] Significance of consensus correlations were based on sample size of 68, the average number of cases across which the correlations were computed. When consensus was determined from intraclass correlations using the subset of rooms for which there were no missing data, the pattern of findings was almost identical, with all dimensions except Emotional Stability reaching significance at the .01 level.

* $p < .05$, one-tailed.   ** $p < .01$, one-tailed.

Table 5
A Brunswik (1956) Lens Model Analysis of Judgments Based on Bedrooms: Cue-Validity and Cue-Utilization Correlations

Cue-validity correlations						Cue-utilization correlations				
Extra.	Agree.	Cons.	Em. St.	Open.	Environmental cues ("lens")	Extra.	Agree.	Cons.	Em. St.	Open.
−.05	−.04	−.15	−.08	−.09	Strong (vs. weak) odor	−.15	−.15	−.05	−.19	.01
.13	−.11	.00	−.05	.05	Noisy (vs. quiet) in room	.05	−.07	−.03	−.09	.02
.25*	.00	.06	.10	.06	Noise (vs. quiet) in house[a]	.21	.06	−.04	.01	.04
.03	−.13	.16	.19	−.16	Noise (vs. quiet) outside[a]	.12	.10	.05	.17	−.19
−.14	−.05	.26*	.18	−.05	Well lit (vs. dark) [overall]	−.02	.04	.07	.23*	−.01
−.13	−.01	.24*	.20	−.18	Well lit (vs. dark) [natural light][a]	.04	.06	.22*	.22*	.00
−.17	−.15	.04	−.01	−.17	Well lit (vs. dark) [artificial light][a]	.07	.20	.08	.17	−.10
.08	−.16	.09	.07	.01	Drafty (vs. stuffy)	.06	.20	.15	.17	.02
−.04	−.11	.17	−.16	−.02	Fresh (vs. stale)	.20	.20	.13	.24*	.06
.00	.17	.13	.01	−.03	Hot (vs. cold)	−.06	−.14	.05	−.06	−.18
.03	−.09	.15	−.09	−.02	Good (vs. poor) condition	.03	.37**	.57**	.02	−.03
.06	−.11	−.10	−.15	.21	Decorated (vs. undecorated)	.41**	.20	.04	.11	.35**
.02	−.05	.07	−.03	.00	Cheerful (vs. gloomy)	.16	.66*	.46*	.12	.00
.07	−.16	.05	−.08	.12	Colorful (vs. drab)	.21	.51**	.42**	.15	.11
.08	−.06	.17	−.08	.02	Clean (vs. dirty)	−.02	.37**	.61**	−.06	−.11
.13	−.12	.29**	.08	−.01	Organized (vs. disorganized)	.01	.26*	.70**	−.02	−.02
.13	−.09	.27*	.04	.04	Neat (vs. messy)	−.05	.33**	.75**	−.06	−.08
−.06	−.01	−.32**	−.14	.14	Cluttered (vs. uncluttered)	.24*	−.15	−.56**	.05	.26*
−.01	−.04	−.11	.07	.03	Clothing everywhere (vs. none visible)[a]	.13	−.39**	−.57**	.12	.17
−.01	.01	−.24	−.18	−.22	Clothing strewn around (vs. organized)[a]	−.04	−.23	−.28*	.02	−.06
−.04	−.01	−.26*	−.16	.15	Full (vs. empty)	.19	−.05	−.35**	.02	.22*
−.03	−.03	.17	.12	−.02	Roomy (vs. cramped)	−.01	.12	.34**	−.07	−.05
−.02	−.08	.04	−.07	−.09	Expensive (vs. cheap)	.11	.21	.31**	.13	.04
.01	.03	.24*	−.05	.03	Comfortable (vs. uncomfortable)	−.07	.43**	.62**	−.15	.03
.06	.00	.19	−.07	.05	Inviting (vs. repelling)	−.01	.52**	.64**	−.03	−.01
.08	.01	.03	.13	.16	Large (vs. small)	−.02	.04	.20	−.07	.10
.19	−.03	−.06	.04	.35**	Distinctive (vs. ordinary)	.20	.01	.12	−.04	.35**
.01	.01	.14	.04	.07	Stylish (vs. unstylish)	.15	.33**	.34**	.20	.11
−.04	−.10	.24*	.06	−.09	Modern (vs. old fashioned)	.05	.27*	.23*	.18	−.03
−.08	−.20	.14	−.09	−.03	New (vs. old)	.11	.38**	.29**	.18	−.01
.02	−.12	.04	−.02	.13	Multiple (vs. single) purpose	.14	.03	.05	.15	.23*
−.09	−.08	−.01	.03	.16	Many (vs. few) books	−.13	−.17	.00	−.20	.37**
−.03	−.13	.24*	.07	−.02	Organized (vs. disorganized) books	−.16	.10	.50**	−.05	.08
.14	−.13	.06	−.02	.44**	Varied (vs. homogenous) books	−.07	−.01	.01	−.07	.50**
.01	−.01	.11	−.02	.18	Many (vs. few) magazines	.05	−.21	−.07	.07	.16
.29	−.38**	.22	.00	.14	Organized (vs. disorganized) magazines	.36*	.23	.27	.12	.01
.15	−.03	−.14	−.11	.51**	Varied (vs. homogenous) magazines	−.05	−.23	−.17	−.19	.33*
−.03	−.14	−.01	−.02	.17	Many (vs. few) CDs	.11	−.10	−.03	.10	.32**
.08	−.15	.27*	.06	−.06	Organized (vs. disorganized) CDs	.04	.26*	.47**	−.01	.02
−.02	−.26*	.01	−.19	.22	Varied (vs. homogenous) CDs	.01	.03	.15	−.08	.09
−.18	.17	−.17	.02	.13	Many (vs. few) items of stationery	−.01	−.07	−.29**	−.06	.19
.26*	−.03	.21	.04	.06	Organized (vs. disorganized) stationery	.00	.41**	.59**	−.12	−.13

Note. Extra. = Extraversion; Agree. = Agreeableness; Cons. = Conscientiousness; Em. St. = Emotional Stability; Open. = Openness.
[a] Cue was assessed in Study 2 but not in Study 1.
* *p* < .05, two-tailed.   ** *p* < .01, two-tailed.

In contrast to judgments made in offices, bedroom-based judgments of Extraversion did not reliably correlate with features of the personal living spaces, with the few exceptions including decoration (.41) and clutter (.24). Few of the coded cues were related to observers' ratings of Emotional Stability, although it is interesting that cues related to certain ambient factors (e.g., lighting, air quality) showed modest correlations.

Cue validity. The cue-validity correlations in Table 5 reflect the relationship between the accuracy criterion and the coded physical features of the bedrooms. These correlations show which cues in the rooms were actually related to characteristics of the occupants.

As in the case of offices, the correlations in Table 5 suggest that observers' intuitions about Conscientiousness cues were largely

valid; the rooms of conscientious individuals were indeed well-organized (.29), neat (.27), and uncluttered (−.32).

Again, the Openness cue-validity correlations provide some validation of the observers' intuitions, such that the distinctiveness of the rooms (.35) and the variety of books (.44) and magazines (.51) were valid cues to an occupant's Openness. Note that it is the variety, not the quantity, of books and magazines that served as the crucial cue to an individual's Openness.

Very few of the coded cues were related to the occupants' levels of Extraversion, Agreeableness, and Emotional Stability, with the few significant correlations quite possibly having arisen by chance.

Intercorrelations of column vectors. The patterns of cue-utilization and cue-validity correlations for bedrooms reveal a range of success in judgments. For traits such as Conscientious-

Table 6
Impact of Stereotypes: Perceived and Real Sex Differences and Race Differences Based on Bedrooms

FFM personality dimensions	Sex differences														
	Perceived (i.e., observers' ratings)									"Real" (i.e., criterion ratings)					
	Mean rating[a]			Effect of perceived sex			Occupant effect			Mean rating			Effect of sex		
	Women (n = 350)		Men (n = 179)	η	F(1, 86)	p	η	F(89, 432)	p	Women (n = 162)		Men (n = 70)	η	F(1, 76)	p
Extraversion	3.3	=	3.1	.15	2.07	.15	.64	3.47	<.001[b]	3.4	=	3.3	.04	.15	.70
Agreeableness	3.7	>	3.5	.26	5.92	.02	.55	2.13	<.001[b]	3.8	=	3.8	.03	.04	.84
Conscientiousness	3.5	=	3.3	.17	2.67	.11	.74	5.88	<.001[b]	3.6	=	3.7	.09	.70	.40
Emotional Stability	3.2	<	3.4	.39	14.16	<.001[b]	.45	1.26	.07	3.0	<	3.6	.40	14.12	<.001[b]
Openness to Experience	3.3	=	3.1	.13	1.39	.24	.80	8.79	<.001[b]	3.8	=	3.6	.07	.38	.54

Note. W = Whites; A = Asians; FFM = Five-Factor Model.
[a] Ratings were made on a 5-point scale.　[b] Probability remains ≤ .05 after Bonferroni correction.

ness, the observers appear to have done a good job, making use of the valid cues and ignoring the invalid cues. Again, we formally tested the extent to which observers' cue-utilization patterns matched the patterns of cue-validity correlations by computing a series of correlations between the column vectors for each trait. The vector correlations thus obtained are shown in the last data column of Table 4.

As in the study of offices, we generally found that in bedrooms the vector correlations were highest for the accurately judged traits. Two of the most accurately judged traits—Openness and Conscientiousness—were characterized by substantial vector correlations. Similarly, the least accurately judged traits—Extraversion and Agreeableness—were characterized by weak vector correlations. However, the magnitude of accuracy correlations was not uniformly matched by the magnitude of vector correlations; for example, Emotional Stability had a substantial accuracy correlation but a weak vector correlation. Thus, it again seems possible that observers' use of cues (i.e., Steps 1a and 2a in Figure 2) is not the whole story and that observers' use of stereotypes (i.e., Steps 1b and 2b in Figure 2) may account for at least some of their accuracy.

Question 4: Stereotype Use

Consensus. As in Study 1, we used a hierarchical multiple regression to examine the effects of sex and race, entering the stereotype variables (i.e., sex or race) as the first step. These analyses tested whether observers rated the occupants differently on the basis of perceived sex and race. We limited our race analyses to comparisons between Asians and Whites, the two groups sufficiently represented in this sample. Any significant differences between women and men and between Asians and Whites for those traits would be consistent with the hypothesis that sex and race stereotypes mediate consensus for judgments of those traits. The first two data columns under the *Perceived* heading under *Sex differences* of Table 6 and the first two data columns under the *Race differences* of Table 6 show the observers' mean ratings.

There were only a few perceived sex differences overall. We found significant sex differences for perceptions of Agreeableness

and Emotional Stability, the same traits that showed perceived sex differences in the study of office spaces. Thus, the directions of the significant sex differences were again consistent with commonly held stereotypes (Eagly & Mladinic, 1989; Spence, 1993). However, our earlier analyses of consensus suggest that there was no consensus for these sex stereotypes to mediate—interobserver agreement for Agreeableness and Emotional Stability was not significant (see Table 4). Although the consensus correlations for Agreeableness failed to reach conventional levels of statistical significance, the direction and magnitude replicate that found in Study 1, suggesting there was some consensus for this trait. On the basis of the findings reported in Table 6, stereotype use could provide one partial explanation for the (weak) levels of consensus associated with Agreeableness; that is, observers adjusted their ratings on the basis of the assumed sex of the occupants.

Observers perceived many more race differences than sex differences. In keeping with widely held cultural stereotypes (e.g., Madon et al., 2001; Miller, 1999; Zhang, Lee, Liu, & McCauley, 1999), Whites were perceived to be significantly more extraverted and open and significantly less agreeable than were Asians. In addition, Whites were perceived to be more emotionally stable than Asians. Although Conscientiousness failed to reach significance, the race difference was in the stereotypically consistent direction, with Asians perceived as more conscientious than Whites. On the basis of the perceived race findings reported in Table 6, one partial explanation for the interobserver agreement associated with Extraversion, Openness, and, to a marginal extent, Conscientiousness could be that observers adjusted their ratings on the basis of the assumed race of the occupants.

Accuracy. Just as in Study 1, we tested for real sex and race differences using hierarchical multiple regression. As shown in Table 6, the only significant real sex differences were found in Emotional Stability; women were less emotionally stable than men. The only significant real race differences were found in Openness, with Whites more open than Asians.

Next we compared the perceived mean ratings with the real mean ratings to assess the accuracy of the sex and race stereotypes used. A comparison of the perceived sex differences with the real sex differences reported in Table 6 shows that only the observers'

Race differences												
Perceived (i.e., observers' ratings)								"Real" (i.e., criterion ratings)				
Mean rating[a]		Effect of perceived race			Occupant effect			Mean rating		Effect of race		
W ($n = 257$)	A ($n = 185$)	η	$F(1, 88)$	p	η	$F(101, 333)$	p	W ($n = 81$)	A ($n = 96$)	η	$F(1, 57)$	p
3.4 > 3.1		.26	6.46	.01[b]	.67	2.75	<.001[b]	3.4 = 3.2		.13	.95	.33
3.5 < 3.7		.25	5.57	.02	.58	1.66	<.001[b]	3.8 = 3.7		.12	.85	.36
3.4 = 3.6		.16	2.54	.11	.76	4.57	<.001[b]	3.7 = 3.5		.10	.60	.44
3.3 > 3.2		.27	5.93	.02	.53	1.30	.04	3.2 = 3.1		.04	.12	.73
3.6 > 2.7		.50	32.30	<.001[b]	.79	5.48	<.001[b]	3.9 > 3.5		.33	7.03	.01[b]

use of sex stereotypes to rate the occupants on Emotional Stability was warranted. For race, only stereotypes used to rate occupants on Openness were warranted. These data are consistent with the idea that observers use sex and race stereotypes, some of which help judgments but most of which hinder them. One should recall that the occupant effects reflect the variance in observer ratings that is not accounted for by differences associated with the perceived sex and race of the occupants. As in Study 1, even the significant perceived sex and race effects were generally much smaller than the occupant effects. This suggests that the use of sex and race stereotypes only partially explains observer consensus and accuracy.

Summary of Study 2 Findings

In Study 2, we set out to determine whether the effects we obtained in workspaces would replicate in a different context (living spaces). Question 1 asked whether observers agreed in their ratings of target occupants based purely on the occupants' personal living spaces. We found that observers generally agreed but that agreement varied across the traits. As in Study 1, we found that Openness, Conscientiousness, and Extraversion showed the strongest agreement and Emotional Stability and Agreeableness showed the least agreement. Question 2 asked whether the observer ratings were accurate. We found that the accuracy correlations were significant for all five FFM domains, with Openness, Emotional Stability, and Conscientiousness showing the most accuracy and Agreeableness and Extraversion showing less accuracy. Question 3 asked which cues in the rooms were used by observers to form impressions and which cues were valid. We identified a wide range of cues relating to observer ratings of Conscientiousness, Agreeableness, and Openness. We also identified a number of cues relating to what the occupants were really like, especially for Conscientiousness. Of the cues we assessed, observers seemed to use valid cues to judge Conscientiousness and Openness. Question 4 asked whether observers' use of sex and race stereotypes could account for interobserver consensus and observer accuracy. Consensus and accuracy were only partially mediated by sex and race stereotypes. Sex stereotypes may have contributed to consensus for Agreeableness, and race stereotypes may have contributed

to consensus for Extraversion, Openness, and, to a lesser extent, Conscientiousness. Our results also suggest that observers' use of sex stereotypes was warranted for judgments of Emotional Stability and that observers' use of race stereotypes was warranted for judgments of Openness.

General Discussion

Overview of the Results

Overall, there are striking similarities between the findings of the office and bedroom studies. It is important to note that with the exception of a single person who served as an observer in both studies, the two studies were entirely independent, with no overlap in observers, occupants, or coders. Our findings suggest that an observer who has briefly examined an individual's living or working environment will form impressions that are remarkably consistent with other observers' impressions. Furthermore, these impressions are often accurate.

Moreover, the remarkable similarity in the patterns of FFM correlations across the two studies suggests that personal environments consistently yield more cues for certain traits than for others. Specifically, the highest accuracy was found for Openness and the least accuracy was found for Agreeableness in both the studies. However, all the accuracy correlations in the study of bedrooms (Study 2) were consistently stronger than those in the study of offices (Study 1). There are five possible explanations for why students' personal living spaces may have provided an unusually good view of their occupants. First, college is a time when individuals are negotiating identity issues, so students may be particularly prone to self-expression. Second, people generally have the freedom to decorate their personal living spaces as they please, but office décor is often restricted by company guidelines. Third, individuals in offices, both voluntarily and because of extrinsic pressure and norms, are typically concerned about the positive and professional image they project. As a result, they may be pressured to arrange and decorate their offices in ways that are contrary to their actual preferences and personalities. Fourth, the observers in both studies were students themselves and perhaps were relatively well versed in the cultural meaning of the possessions and icons

found in student living spaces but less so in the meanings of cues in office spaces. Fifth, the rooms in our study of bedrooms were often multipurpose spaces where the occupants spent a great deal of their time, using them for a variety of activities such as working, sleeping, relaxing, and entertaining. In college dwellings, individuals often invite guests into their living spaces, which become the context for a broad variety of social interactions ranging from TV-watching gatherings and Bible-study sessions to raucous parties and intimate romantic encounters. Thus, the personal living spaces in the study of bedrooms may have been ecologically richer environments than were the office spaces.

Where we were able to make cross-study comparisons in stereotype use, we found interesting similarities. For instance, recall the sex stereotypes for the FFM dimensions; in both studies observers held the stereotypes that men were less agreeable but more emotionally stable than women. The results suggest that observers did rely on stereotypes to form impressions. However, stereotype use probably accounted for only a small portion of the consensus and accuracy findings in both studies. In the two studies, the occupant effects were all significant, with the exception of the Emotional Stability dimension in the bedroom study, which just failed to reach significance. The implication of these findings is that even though the observers in both studies used stereotypes to form impressions, they did not base their judgments solely on stereotypes but may have drawn more heavily on the physical cues in the rooms. When we link these findings back to the model we propose in Figure 2, it seems that stereotype use (Steps 1b and 2b) played a relatively minor role in the judgment process.

The column–vector correlations shown in the last column of Tables 1 and 4 suggest that in both offices and bedrooms, observers were consistently better at judging some traits than others. By better, we mean that observers tended to use cues in their judgments that were actually related to the traits they were trying to judge (cf. Cue 1 in Figure 1) and ignored cues that were unrelated to the traits (cf. Cue 4 in Figure 1). Across studies, judges seemed to make particularly good use of the environmental information to judge Conscientiousness and Openness. In light of these findings, the question arises as to whether the same cues were diagnostic of traits across contexts. For example, are some cues diagnostic of Conscientiousness across contexts, or are the cues that are diagnostic of Conscientiousness in offices different from those that are diagnostic of Conscientiousness in bedrooms?

To answer this question, one must compare the correlations in Tables 2 and 5. Therefore, we computed a new set of column–vector correlations comparing the cue-utilization correlations from Study 1 with the cue-utilization correlations from Study 2 and comparing the cue-validity correlations from Study 1 with the cue-validity correlations from Study 2. These column–vector correlations were computed across the 36 environmental cues that were common to both studies. The resulting correlations are shown in Table 7. These correlations suggest that observers tended to use similar patterns of cues across office and bedroom contexts when judging occupants' Extraversion, Agreeableness, Conscientiousness, and Openness. However, the correlations suggest that the actual relationships between occupants and their environments were less consistent across the office and bedroom contexts; cue-validity correlations were similar across studies only for Conscientiousness. It seems, then, that personality attributes are manifested somewhat differently across these contexts.

Table 7
Column–Vector Correlations: How Similar Are the Cue-Utilization and Cue-Validity Correlations Across the Office and Bedroom Studies?

	Cross-study similarity in pattern	
FFM personality dimension	Cue-utilization correlations[a]	Cue-validity correlations[b]
Extraversion	.37*	−.18
Agreeableness	.44**	−.04
Conscientiousness	.86**	.58**
Emotional Stability	.05	.17
Openness to Experience	.54**	.27

Note. Column–vector correlations were computed across the environmental cues that were common to both studies. $N = 36$ cues. FFM = Five-Factor Model.
[a] Office cue-utilization correlations (from Table 2) are compared with the bedroom cue-utilization correlations (from Table 5).   [b] Office cue-validity correlations (from Table 2) are compared with the bedroom cue-validity correlations (from Table 5).
* $p < .05$, one-tailed.   ** $p < .01$ one-tailed.

Comparisons With Previous Research

Recall that we questioned whether personal environments provide less or more information than that provided by zero-acquaintance studies. In past zero-acquaintance studies, observers were exposed to limited information in the form of brief interactions, short videotapes of the targets, or photographs. We predicted that personal environments would provide more information than typical zero-acquaintance contexts and would therefore result in relatively strong consensus and accuracy correlations. To test this prediction, we next compare our findings with findings from previous zero-acquaintance research.

Consensus and Accuracy

Figure 3 summarizes consensus correlations from the studies of offices and bedrooms reported here along with findings from a meta-analysis of zero-acquaintance and long-term acquaintance (e.g., year-long friendships) studies (Kenny, 1994; Kenny et al., 1994). All the consensus correlations based on environments are larger, in many cases substantially so, than the zero-acquaintance correlations. For Conscientiousness and Openness, the environment-based consensus was substantially stronger than that found even in past studies of long-term interactions. Even for Extraversion and Agreeableness, environment-based consensus was stronger or comparable to consensus obtained after long-term acquaintance. Thus, it seems that personal environments contain richer sources of information from which to form impressions than are contained in zero-acquaintance contexts.

Figure 4 summarizes the accuracy correlations from studies of offices and bedrooms reported here along with findings from a meta-analyses of zero-acquaintance and long-term acquaintance studies (Kenny, 1994). For Emotional Stability and Openness, environment-based accuracy was substantially stronger than was accuracy found in zero-acquaintance research. However, Extraversion was slightly more accurately judged in zero-acquaintance
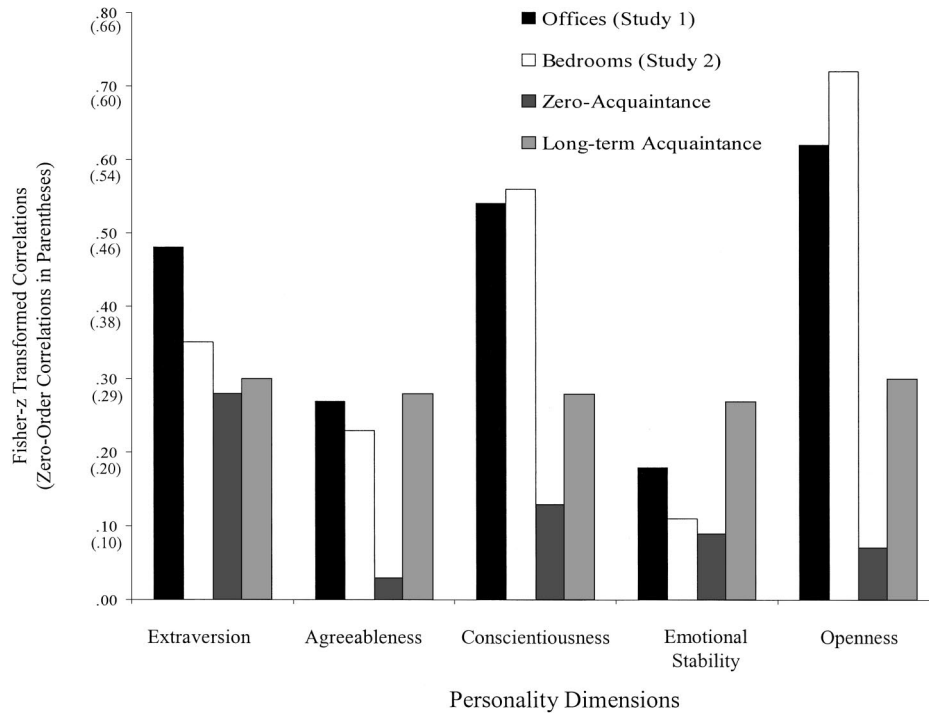
Figure 3. Observer consensus as a function of informational base: office environments (Study 1), living environments (Study 2), zero-acquaintance contexts, and long-term acquaintance contexts. Zero-acquaintance and long-term acquaintance data are from Kenny (1994). To provide a linear representation of the correlations on the *y*-axis, we report the correlation coefficients in terms of Fisher's *z* metric. To make the comparisons with Kenny's data parallel, the office and living environment correlations have been corrected for unreliability.

contexts than in the context of personal environments, with neither of the contexts consistently providing more information than the other for Agreeableness and Conscientiousness. Figure 4 again supports the idea that physical environments elicit accurate impressions of personality, even when compared with studies of long-term acquaintance. One clear exception is for Extraversion, where accuracy is strongest in long-term acquaintance studies. More generally, different contexts afford different information about targets; for example, if you want to learn about someone's Extraversion you should meet him or her but if you want to learn about a person's Openness, an examination of his or her bedroom would be more useful.

Although both contexts provide important and perhaps distinct information, it is not surprising that personal environments contain richer sources of behavioral information than do zero-acquaintance contexts. Information accumulated in personal environments is often the result of repeated behaviors. For example, to have an organized office it is not sufficient to organize the office just once; instead, the occupant must continually engage in organizing behaviors—returning the phone directory to the bookshelf after use, throwing away used paper cups, and placing documents in neat stacks. Multiple acts are more likely to have an impact on the environment than are single acts. Because environmental cues tend to reflect repeated acts, they may offer more reliable evidence than the few acts that observers witness in many zero-acquaintance contexts.

Cue Use

How do the vector correlations shown in Tables 1 and 4 compare with previous findings? Funder and Sneed (1993; Sneed, McCrae, & Funder, 1998) examined a similar issue, but instead of behavioral residue they focused on behavior itself, relating 62 actual behaviors to the FFM dimensions. Their index of cue utilization was the correlation between observers' ratings of target individuals portrayed in a 5-min videotaped interaction and behaviors coded from the videotapes. Their index of cue validity was the correlation between ratings of the targets made by well-acquainted informants and the coded behaviors. Thus, the structure of our study is remarkably similar to theirs.

A comparison of Funder and Sneed's (1993) findings with ours reveals a number of interesting differences between the value of information in behavioral and environmental contexts. Recall that we found that observers made best use of environmental cues for Openness and Conscientiousness. Funder and Sneed (1993) did find that observers made good use of behavioral cues for Conscientiousness. However, observers did not make good use of behavioral cues for Openness. Instead, in stark contrast to our findings for environmental cues, observers made relatively good use of behavioral cues in their judgments of Extraversion, Agreeableness, and, to a lesser extent, Emotional Stability. Thus, although observers might have a hard time identifying the behavioral manifestations of Openness (Sneed et al., 1998), they appear to be rather successful at identifying the environmental manifestations of this
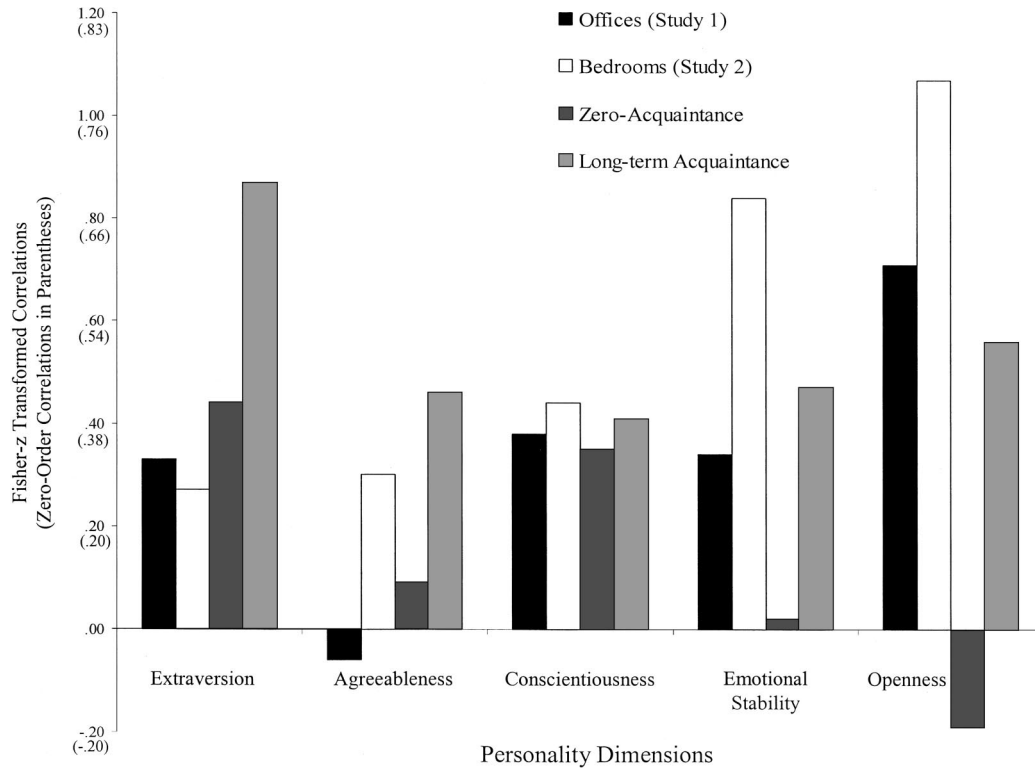
Figure 4. Observer accuracy as a function of informational base: office environments (Study 1), living environments (Study 2), zero-acquaintance contexts, and long-term acquaintance contexts. Zero-acquaintance and long-term acquaintance data are from Kenny (1994), who used the term *self–other agreement* to refer to these correlations. To provide a linear representation of the correlations on the *y*-axis, we report the correlation coefficients in terms of Fisher's *z* metric. To make the comparisons with Kenny's data parallel, the office and living environment correlations have been corrected for unreliability.

dimension. Together, these studies support our expectation that the manifestation and judgment of personality traits would vary across domains.

Potential Limitations

We next turn to a number of potential limitations of the research reported here. The first two limitations relate to alterations occupants may have made to their personal environments. The first of these concerns is that occupants may have tidied or altered their rooms before the assessment team arrived. The occupants knew that an assessment team would be going through their room to learn about their personality, and they knew when the team would be there. However, there are several reasons to think that the rooms were not tidied or altered. First, we believe that occupants complied with our specific request not to tidy or alter their rooms because the rooms were assessed under conditions of complete anonymity and confidentiality. Second, the main incentive for the occupants to take part in this research was to receive feedback on the impressions formed on the basis of their personal environments. We reminded participants that meaningful feedback depended on observers seeing the environments in their unaltered state. Third, the occupants' peers indicated in their confidential ratings how much they thought the rooms had been altered for the assessment; in Study 1 the mean rating was 1.5 on a scale ranging

from 1 (*very little*) to 7 (*very much*), and in Study 2 the mean rating was 1.4 on a scale ranging from 1 (*very little*) to 5 (*very much*), both indicating the rooms had not been substantially altered.

The second potential limitation regards more permanent manipulation of the environment, not just alterations for our assessments. Occupants may craft their environments to project specific impressions that they deem desirable. For example, they may display socially desirable symbols (e.g., an award for public service), they may fabricate behavioral residue (e.g., a made-up bed), or they may deceivingly display anticipated behavior cues they have no intention of using. The present design did not permit us to examine the extent to which self-presentational concerns drove the occupants to actively manipulate their personal environments to portray themselves in a positive light, and this is an area that would benefit from experimental and field research.

A third potential limitation of the present design is that our analyses of cues were limited to the set of cues we measured. Although we strove to assess a broad range of cues, we cannot be certain that our inventory was comprehensive. We may have oversampled cues that were relevant to some dimensions (e.g., Conscientiousness) and undersampled cues that were relevant to others (e.g., Emotional Stability). In addition, the analyses reported here focus on broad cues and do not include specific cues (e.g., a poster of Martin Luther King) that may have conveyed a lot of

information to observers. Moreover, we did not have the statistical power to search for configurations of cues (e.g., the combination of organized and dark). It seems quite likely that many judgments are made on the basis of such configural patterns. The findings reported here are limited to the extent that the codings neglected such information. However, there is some evidence to suggest that our codings did not miss too much information. If the codings had neglected information that judges used to make judgments, then the pattern of accuracy correlations would not mirror the pattern of vector correlations. But inspection of Tables 1 and 4 shows that the accuracy correlations generally do mirror the vector correlations, suggesting that we did capture most of the relevant cues.

Future Directions

The present research sought to document the basic links between occupants and their personal environments and between personal environments and observers' perceptions of the occupants. We conceptualized these links in terms of processes specified by Brunswik's (1956) lens model: accuracy, cue utilization, and cue validity. We found evidence for each of these processes, demonstrating clear links between individuals and the physical spaces in which they live or work. Thus, the studies presented here lay a strong foundation of basic findings on which future research can build. Research can now directly examine the specific mechanisms and processes hypothesized to account for these basic phenomena. Models such as the one depicted in Figure 2 should be used to guide researchers as they embark on this vital second phase in understanding the complex relations between individuals and the places they inhabit.

Conclusion

This research takes a first look at what one can and cannot learn about people by examining the environments in which they live and work. Although this line of empirical research is relatively new, the practice of assessing individuals by examining their environments is not. In fact, one formal application was put in place by the U.S. government shortly after entering the Second World War. In May 1942, the Office of Strategic Services began a program of assessments designed to identify candidates suitable for work behind enemy lines. One of the selection tests was the Belongings Test, in which candidates were required to describe individuals solely on the basis of what they had left in their bedrooms—items included clothing, a time table, and a ticket receipt (MacKinnon, 1977). Our research provides strong support for the assumption underlying this test—much can be learned about persons from the spaces in which they dwell.

References

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111,* 256–274.

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin, 91,* 3–26.

Berry, D. S., & Finch-Wero, J. L. (1993). Accuracy in face perception: A view from ecological psychology. *Journal of Personality, 61,* 497–520.

Blackman, M. C., & Funder, D. C. (1998). The effect of information on consensus and accuracy in personality judgment. *Journal of Experimental Social Psychology, 34,* 164–181.

Block, J. (1961). *The Q-sort method in personality assessment and psychiatric research.* Springfield, IL: Charles C Thomas.

Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology, 62,* 645–657.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments.* Berkeley: University of California Press.

Burroughs, J. W., Drews, D. R., & Hallman, W. K. (1991). Predicting personality from personal possessions: A self-presentational analysis. *Journal of Social Behavior and Personality, 6,* 147–163.

Buss, D. M. (1987). Selection, evocation, and manipulation. *Journal of Personality and Social Psychology, 53,* 1214–1221.

Buss, D. M., & Craik, K. H. (1983). The act frequency approach to personality. *Psychological Review, 90,* 105–126.

Chaplin, W. F., Phillips, J. B., Brown, J. D., Clanton, N. R., & Stein, J. L. (2000). Handshaking, gender, personality, and first impressions. *Journal of Personality and Social Psychology, 79,* 110–117.

Eagly, A. H., & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin, 15,* 543–558.

Funder, D. C. (1980). On seeing ourselves as others see us: Self–other agreement and discrepancy in personality ratings. *Journal of Personality, 48,* 473–493.

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102,* 652–670.

Funder, D. C. (1999). *Personality judgment: A realistic approach to person perception.* San Diego, CA: Academic Press.

Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology, 52,* 409–418.

Funder, D. C., & Sneed, C. D. (1993). Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology, 64,* 479–490.

Gifford, R. (1994). A lens-mapping framework for understanding the encoding and decoding of interpersonal dispositions in nonverbal behavior. *Journal of Personality and Social Psychology, 66,* 398–412.

Gilovich, T., & Savitsky, K. (1996, March/April). Like goes with like: The role of representativeness in erroneous and pseudoscientific beliefs. *Skeptical Inquirer, 20,* 34–40.

Goffman, E. (1959). *The presentation of self in everyday life.* Garden City, NY: Doubleday Anchor.

Gosling, R. D., & Standen, R. (1998). Doctor's dress. *British Journal of Psychiatry, 172,* 188–189.

Gosling, S. D., Martin, N. R., Craik, K. H., & Pryor, M. R. (2001). *Personal living spaces: Development and analysis of a cue inventory.* Manuscript submitted for publication.

Hayes, A. F., & Dunning, D. (1997). Construal processes and trait ambiguity: Implications for self–peer agreement in personality judgment. *Journal of Personality and Social Psychology, 72,* 664–677.

John, O. P., & Benet-Martinez, V. (2000). Measurement, scale construction, and reliability. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social psychology* (pp. 339–369). Cambridge, England: Cambridge University Press.

John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality, 61,* 521–551.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality theory and research* (pp. 102–138). New York: Guilford Press.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis.* New York: Guilford Press.

Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus in interpersonal perception: Acquaintance and the Big Five. *Psychological Bulletin, 116,* 245–258.

Kenny, D. A., Horner, C., Kashy, D. A., & Chu, L. (1992). Consensus at zero acquaintance: Replication, behavioral cues, and stability. *Journal of Personality and Social Psychology, 62,* 88–97.

Lee, Y., Jussim, L. J., & McCauley, C. R. (1995). *Stereotype accuracy: Toward appreciating group differences.* Washington, DC: American Psychological Association.

Locksley, A., Borgida, E., Brekke, N., & Hepburn, C. (1980). Sex stereotypes and social judgment. *Journal of Personality and Social Psychology, 39,* 821–831.

MacKinnon, D. W. (1977). From selecting spies to selecting managers: The OSS Assessment program. In J. L. Moses & W. C. Byham (Eds.), *Applying the assessment center method* (pp. 13–30). New York: Pergamon.

Madon, S., Guyll, M., Aboufadel, K., Montiel, E., Smith, A., Palumbo, P., & Jussim, L. (2001). Ethnic and national stereotypes: The Princeton trilogy revisited and revised. *Personality and Social Psychology Bulletin, 27,* 996–1010.

McCrae, R. R., & Costa, P. T. Jr. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality theory and research* (pp. 139–153). New York: Guilford Press.

McCrae, R. R., Stone, S. V., Fagan, P. J., & Costa, P. T., Jr. (1998). Identifying causes of disagreement between self-reports and spouse ratings of personality. *Journal of Personality, 66,* 285–313.

Miller, L. (1999). Stereotype legacy: Culture and person on Japanese/American interactions. In Y. Lee, C. R. McCauley, & J. G. Draguns (Eds.), *Personality and person perception across cultures.* Mahwah, NJ: Erlbaum.

Park, B., & Judd, C. M. (1989). Agreement on initial impressions: Differences due to perceivers, trait dimensions, and target behaviors. *Journal of Personality and Social Psychology, 56,* 493–505.

Paulhus, D. L., & Bruce, M. N. (1992). The effect of acquaintanceship on the validity of personality impressions: A longitudinal study. *Journal of Personality and Social Psychology, 63,* 816–824.

Robins, R. W., Gosling, S. D., & Donahue, E. M. (1997). Are personality judgments based on physical appearance consensual and accurate? In J. Bermúdez, B. de Raad, J. de Vries, A. M. Pérez-García, A. Sánchez-Elvira, & G. L. van Heck (Eds.), *Personality psychology in Europe* (Vol. 6, pp. 70–75). Tilburg, the Netherlands: Tilburg University Press.

Sneed, C. D., McCrae, R. R, & Funder, D. C. (1998). Lay conceptions of the five-factor model and its indicators. *Personality and Social Psychology Bulletin, 24,* 115–126.

Snyder, M., & Ickes, W. (1985). Personality and social behavior. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (Vol. 2, pp. 883–947). New York: Random House.

Spence, J. T. (1993). Gender-related traits and gender ideology: Evidence for a multifactorial theory. *Journal of Personality and Social Psychology, 64,* 624–635.

Swann, W. B., Jr. (1987). Identity negotiation: Where two roads meet. *Journal of Personality and Social Psychology, 53,* 1038–1051.

Swann, W. B., Jr., Rentfrow, P. J., & Guinn, J. S. (in press). Self-verification: The search for coherence. In M. Leary & J. Tangney (Eds.), *Handbook of self and identity.* New York: Guilford Press.

Watson, D. (1989). Strangers' ratings of the five robust personality factors: Evidence of a surprising convergence with self-report. *Journal of Personality and Social Psychology, 57,* 120–128.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences.* Chicago: Rand McNally.

Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment.* Reading, MA: Addison-Wesley.

Zhang, K., Lee, Y., Liu, Y., & McCauley, C. R. (1999). Chinese–American differences: A Chinese view. In Y. Lee, C. R. McCauley, & J. G. Draguns (Eds.), *Personality and person perception across cultures.* Mahwah, NJ: Erlbaum.