

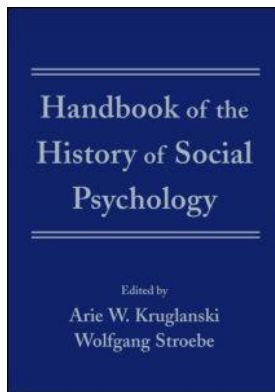
This article was downloaded by: 10.2.97.136

On: 10 Jun 2023

Access details: *subscription number*

Publisher: *Routledge*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London SW1P 1WG, UK



## **Handbook of the History of Social Psychology**

Arie W. Kruglanski, Wolfgang Stroebe

### **A History of Social Judgment research**

Publication details

<https://test.routledgehandbooks.com/doi/10.4324/9780203808498.ch11>

J. Richard Eiser

**Published online on: 01 Dec 2011**

**How to cite :-** J. Richard Eiser. 01 Dec 2011, *A History of Social Judgment research from: Handbook of the History of Social Psychology* Routledge

Accessed on: 10 Jun 2023

<https://test.routledgehandbooks.com/doi/10.4324/9780203808498.ch11>

**PLEASE SCROLL DOWN FOR DOCUMENT**

Full terms and conditions of use: <https://test.routledgehandbooks.com/legal-notices/terms>

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

# 11 A history of social judgment research

*J. Richard Eiser*

The field of social judgment, perhaps more than any other, has been one where the boundary between social and more general psychology is constantly shifting and hard to draw. This has also meant that the field is difficult to circumscribe. Either it tends to be defined so broadly, for instance in terms of how we “make sense of our social world,” that it potentially includes topics from all the other chapters of this volume; or it is defined far more narrowly, with reference to particular methodological paradigms, but here the problem is that research within a number of these paradigms has tended to develop rather independently, so that the resulting narrower definitions overlap only partially with one another. In view of this lack of a clear consensus on the scope of the field, the review that follows will necessarily be selective, and perhaps somewhat more narrative than systematic. The cast of characters in this narrative likewise will be less than exhaustive, and I apologize in advance for any omissions that disappoint. My main purpose, however, will be to highlight the influences that have shaped the field, as I see it, and offer prospects for its continued development. I shall be offering my personal view of the key ideas that are worth taking from this area of work, how these have evolved, and where they may lead.

## What kind of history?

Before attempting to define the concept of (social) judgment more precisely, let me try to characterize the kind of history this will be, and to do so by way of an analogy. Imagine a small state, strategically placed near the borders of more powerful neighbors, not especially populous but with enviable resources and its own, somewhat distinct, culture. Much of its early history was quite stable and prosperous, being well integrated into (and influential within) the dominant regional power of the time. Then, however, the old empire declined and fractured, and the region as a whole entered a time of flux, with new movements and alliances vying for supremacy. Our little state was of course left vulnerable to invasion, first from one side, then another. Some of these invaders attempted to assimilate local customs, others simply sought to impose their own interests and language. During this period the very identity of the state was threatened. This was also, to some extent, a time of diaspora, as some citizens left their homeland, at least temporarily, to seek new opportunities elsewhere.

To write a history of this period presents difficulties. Should it be a history of the occupying powers? Partly it has to be. (One can't, for instance, describe the history of Roman Britain without saying something about Rome itself—but not *everything* that could be said about Rome.) My own focus of interest, however, is not on the occupiers themselves (of which much will have been written elsewhere), but on what they found and what they left behind, and if any continuities can be traced from the state's earlier history to the present day, as it emerges, cautiously perhaps, toward a new independence. So the kind of history I shall offer here is one that attempts to focus on those continuities that, I believe, have survived despite, and perhaps because of, these upheavals. The continuities that, I believe, matter—that constitute the identity of the field (or the culture of the state, in my analogy)—are *ideas*; that is, theoretical principles and priorities. All such ideas are, of course, originally the work of individuals who have shaped and expressed them, but ideas, if they are any good, can take on a life of their own that outlasts the careers of their authors. So it is ideas that will be the main characters in this drama. We researchers and writers provide the chorus.

## What kind of judgment?

Although there is no single agreed definition of “social judgment,” some boundaries need to be placed around the field, for purposes of this chapter at least. What do we mean by “judgment”? Clearly, there is a sense in which *all* cognitive and emotional responses to events in the environment involve *some* form of information-processing, that is, judgment. But defining judgment as “thinking about things,” while correct, doesn't get us very far in terms of formulating more focused research questions. Likewise, all behavioral decisions can be said to involve some form of action selection or choice between alternatives; in other words, a kind of comparative judgment. But again, saying that all behavior involves judgment doesn't tell us very much about what judgment is (or isn't). So the approach I propose to adopt is to confine my attention to research that has studied these general processes in relation to more specific objects of judgment, methods, and outcome variables. Put simply, I take the term “judgment” to refer to the process(es) whereby participants in psychological experiments *rate objects and events* in terms of some attribute or attributes, usually specified by the

researcher. This may sound rather limited and artificial, but the assumption (not only mine) is that such ratings reflect more general processes that are involved whenever we encounter objects and (especially) compare them with one another.

But this chapter is not just entitled “judgment” but “social judgment.” So alongside all this there has been the question lurking in the background of what makes social judgment especially *social*. In many ways, this question is far less important than it might seem, and has more to do with aspirations by some for a kind of group distinctiveness than with anything that is more substantive intellectually. To anticipate my general argument on this point—the processes underlying social judgment phenomena are extremely general. For the most part, what distinguishes *social* judgment is not the kind of underlying process, but the objects or stimuli being judged. The term “social judgment” itself has an uncertain provenance, but owes most to Sherif and Hovland’s (1961) book of the same name. They argued for the study of judgment of “motivationally relevant stimuli” as distinct from the “motivationally neutral stimuli” typically used in laboratory experiments of the time. Their book was the fourth volume in the series *Yale Studies in Attitude and Communication*, so their emphasis was very much on links with attitude theory. A parallel tradition, explicitly derived from Brunswik’s (1943) theory of organization in perception, is represented by Hammond’s (e.g., 1955; Hammond, Stewart, Brehmer, & Steinmann, 1975) “social judgment theory.” This essentially offers a systematic framework for the prediction of choices among multiattribute objects. The main influence of this work has been on subsequent research in judgment and decision-making, rather than more mainstream social psychology, but Hammond shares with Sherif and Hovland a commitment to the importance of studying the representation of “ecologically valid” objects. (I shall say more about these two traditions shortly.)

The common thread in these and other bodies of work is to use the term “social judgment” when the objects we judge have features that carry meaning, and especially *value*, derived from social relationships and interactions. However, this does not imply that we have evolved or developed a completely different set of psychological capacities for processing social as distinct from nonsocial information. Indeed, this argument cuts both ways. Why should we have evolved separate capacities for dealing with nonsocial information when so much of our experience is social? In short, there is nothing “derivative” or “secondary” about social information processing. *All* our affective, cognitive and behavioral capacities facilitate our functioning as social animals. If any psychological research deserves to be called derivative, it is that conducted in attempted isolation from any social context.

### Judgment as perception

The origins of social judgment research, like those of so much of experimental psychology, can be traced back to 19th-century work on psychophysics. The most easily identified legacy, as

we shall see shortly, is mainly methodological—in the development of rating scales for reporting subjective experiences in comparative and absolute terms. But we can find an even deeper resonance with contemporary concerns if we look at some of the more philosophical assumptions that guided such early work. The classic text, of course, is Gustav Fechner’s (1860) *Elemente der Psychophysik*.

We take many things for granted now, including even the very term “psychophysics.” Perhaps by now most psychologists, if they consider it at all, take it as some kind of shorthand (like many other compound German nouns) for “the psychology of how we perceive or judge physical stimuli,” or such like. But for Fechner it was no mere shorthand. It amounted to a manifesto: that the psychological and the physical, the mental and the material, are essentially one. This unity is expressed within the single word, just as deliberately (and arguably more elegantly) than when contemporary writers introduce terms such as “neurophilosophy” (Churchland, 1986) or “mind/brain” (Dennett, 1991). Of course, this was well before any kind of brain science in the sense we know it today. Fechner was not trying to tell us *how* the brain enables action, thought, and feeling. Nor was he just trying to make the obvious point *that* the brain is necessary for behaviour and mental life. He was not waiting, like Damasio (1994), for neurological case studies—still less neuroimaging—to reject Descartes’ dualism. His argument was even more fundamental—that subjective experience is a physical process, governed by physical *laws*.

This assertion, that mental life is lawful, is something that needs to be taken quite strictly and seriously if we are to understand the real contribution of the psychophysical tradition. Even a century later, Stevens (1957, 1966) was writing about “the psychophysical *law*.” Although not always stated explicitly, the assumptions here run very deep: that (social) judgment can tell us something very basic about how, in a physical system, thought is possible, and that the underlying processes can be described, not just qualitatively, but in terms of quantitative laws possessing a high degree of generality. And not just generality. Within their proper range of convenience, the laws proposed by Fechner and updated by subsequent researchers have a remarkable degree of precision. Even though learning theorists in the early part of the last century were also looking for laws, few, with the notable exception of Hull (1952), aspired to anything like the same kind of mathematical sophistication. But, of course, it is not just psychology, but physics too, that has been radically transformed over the past one-and-a-half centuries. The laws of physics now are far more dynamic and less narrowly deterministic than in Fechner’s day. Far from undermining the psychophysical project, however, this opens up new areas where the unity of the psychological and the physical can be profitably explored.

But back to the specific context within which classical psychophysics laid the foundation for subsequent research in social judgment. Quantitatively relating the psychological to the physical meant relating psychological quantities to physical quantities. It was relatively unproblematic to measure most

physical quantities of interest, such as length, weight, and (later) other intensities such as light and sound. But what of “psychological quantities,” such as perceived size or felt heaviness? These could not be measured directly, but could only be inferred from observed responses, such as judgments on a rating scale. Thus, whereas the aim of a “psychophysical law” was to relate the intensity of perception to the physical intensity of a presented stimulus, the data available to researchers could only be in the form of *reported* perception or sensation. In other words, the relationship between a physical continuum (e.g., objective weight) and a psychological continuum (e.g., subjective heaviness) could only be established indirectly, through asking participants to express their subjective experiences on a judgmental continuum, such as a rating scale.

Experimental research in psychophysics then set about developing procedures for stimulus presentation and response elicitation to enable reasonably confident inferences to be made about the “psychological magnitude” of particular perceptual stimuli under specified conditions. By the early part of the last century, many of these techniques and conventions were well established. Then two related things happened that converged to transform judgment research and in the process make it much more accessible, and integral, to social psychology. One (and the earlier chronologically) was that, as psychology moved into new fields of application, it turned to these established techniques to see if they could be adapted to the measurement of a wider range of attributes and constructs, such as intelligence, personality, and attitudes. I shall return to this shortly. The other was that researchers took a renewed interest in the sensitivity of perceptual and social judgments to the *context* in which stimuli were presented.

This latter line of work brings out into the open a persistent tension within judgment research—whether we are looking primarily at what people do *right*, or at what they do *wrong*. For those (like Fechner) committed to a search for general “laws,” the starting point for any theory is a consideration of the functions that need to be fulfilled by any cognitive system that fits an organism to its environment. For us humans at least, these functions include the ability to identify and discriminate between relevant objects in our environment, to differentiate between stimulus magnitudes, and to focus our attentional or perceptual capacities on those stimulus differences and attributes that are informative for the discriminations we need to make (where “need” is defined in terms of motivational goals or survival considerations). For myself, any judgment theory worthy of the name should aim to improve our understanding of these functions. However, a fair amount of research appears to take a rather different line, through emphasizing our apparent vulnerability to a variety of errors and biases. Of course there is no necessary contradiction between these approaches, since one can learn much about the workings of any system by seeing when it fails. Even so, the discovery, naming, and prediction of “effects” suggestive of cognitive failures or limitations has (perhaps too) often seemed to have been taken by researchers as an end in itself.

## Context effects and adaptation

How judgments are affected by the context in which stimuli are presented is one of the most pervasive questions in judgment research. There are very many kinds of context effects, with some being much more reliable and extensively researched than others. However, it took some time before judgment researchers realized that context effects could be informative theoretically in their own right as opposed to just a source of nuisance variance. I shall therefore confine myself to a few of those that have been most significant historically for the development of theory.

By far the best documented context effect is that of *contrast*. This refers to what happens when judgments of a stimulus appear to shift away from the background context. Thus, the same object, presented against a background of larger stimuli, will be judged as smaller than if presented in the context of other stimuli smaller than itself. The same room can appear dim if one is coming in from bright sunlight, or well-lit if coming in from the dark. Judgments, in other words, are relative—but what exactly are they relative to?

The most influential early answer to this question was that judgmental contrast reflected a form of perceptual or sensory adaptation. Adaptation is a familiar process of adjustment to prevailing levels of stimulation along dimensions such as brightness (as above), loudness, heat, and such like. As such, it fulfills an essential function (in the terms I’ve just mentioned) for organisms that need to adjust to changes in their environment. Through such adjustment, what previously might have been experienced as bright, loud, hot, etc. is now experienced as average, whereas stimuli previously experienced as average now seem dim, quiet, cool, and so on. According to Helson’s (1947, 1964) adaptation-level theory, every stimulus is perceived relative to a subjective average or neutral point (along the relevant dimension of stimulus intensity) termed adaptation level (AL). Stimuli that are more intense (e.g., larger) than the AL are perceived as large; stimulus intensities below AL are perceived as small or weak. AL is assumed to represent the pooled effect of past and present stimulation on the relevant dimension, calculated as a weighted average. In other words, it reflects a form of cumulative *learning*. Since each new stimulus presentation adds into this weighted average, the assumed process is essentially dynamic, reflecting a continuous process of adjustment and readjustment. Several studies testify to the fact that such contrast effects are highly general. In the days before computer-controlled stimulus presentations, the paradigm of choice involved presenting series of small brass knobs or cylinders, precisely engineered to weigh a specified number of grams. Importantly, these cylinders were all the same size, and so indistinguishable visually. Participants then lifted the weights and reported how light or heavy each appeared to be.

Helson’s AL theory faced challenges on two main fronts. How valid were his theoretical assumptions about the underlying causal process? And how accurate was his model predictively? These two questions turn out to be interrelated. At the

level of theory, Helson places essentially the whole causal burden on the hypothesized process of *sensory* adaptation. This may be thought of literally in the context of dimensions such as weight or brightness, but when (as in his 1964 book) Helson speculates on the relevance of his theory to other domains such as aesthetic and moral judgments, it is no longer clear what kinds of stimulus intensities are being averaged together, or how, to form the relevant AL. Moreover, even in the field of perceptual judgment, the process of adaptation is more obvious for some dimensions than for others. Stevens (1966) distinguished between what he called “prothetic” continua, where an increase in physical intensity produces a quantitative (additive) increase in perceived intensity (e.g., brightness), and “metathetic” continua, where a physical increment produces a qualitative change in sensation (e.g., hue). Stevens introduced this distinction to account for the difficulty of providing a general psychophysical law for metathetic continua, but it is relevant also to the applicability of the concept of sensory adaptation, and how any neutral point could be computed. The point is not that contrast effects cannot occur, for instance, in perceptions of colour, where different hues are juxtaposed. It is rather that accounting for such local effects does not require any notion of a subjectively “average” colour.

There is an additional difficulty here, stemming from the same problem that confronted Fechner. Perceptions can only be inferred from expressed judgments. But judgments may change, may have their own dynamics, for reasons other than mere changes in sensation. This is especially so when (as is typical) participants are required to express such judgments in semantic terms (such as “heavy” or “light”) and/or on numerical scales anchored by such terms. Such terms are not absolutely tied to any physical quantity but can themselves be interpreted relative to context. A light aeroplane is a good deal heavier than a heavy book. We need to know what kinds of objects we are talking about in order to use such adjectives appropriately, in accordance with Gricean maxims (Grice, 1975). If we are given novel stimuli (such as brass weights) to judge, we may rely to some extent on our past experience (Tresselt, 1948) but essentially we need to work out, that is *learn*, what “light,” “average,” and “heavy” *mean* in the context of the actual stimuli presented. So, even when sensory adaptation effects are plausible as a factor contributing to judgmental contrast, it is also often the case that “semantic shifts” could have occurred in participants’ use of the judgmental terms, and in the same direction.

Distinguishing the relative contribution of sensory and semantic processes to contrast effects is therefore difficult, both practically and logically, and probably self-defeating as a problem to get fixated upon. However, one consequence of at least recognizing the problem was that researchers began to question the notion of AL as merely a weighted average of past and present stimulation, and considered other statistical properties of the distribution of stimuli presented for judgment. Most influential among such attempts was Parducci’s (1963) “range–frequency compromise.” Parducci was content to start from the notion that judgments are made relative to a psychological

neutral point or AL, but disputed Helson’s argument that AL was computed through a process of subjective averaging of pooled experience.

For Parducci, the main question was not how any single stimulus might be judged, but how judgments of the whole set of presented stimuli were distributed over the separate categories of the rating scale. Two principles govern this distribution. The range principle states that judges will match or “anchor” the end-points of the rating scale to the (actual or expected) extremes of the presented stimulus series (see Upshaw, 1962, for an application of this principle to social judgment). In other words, the range of judgments matches the range of stimuli. Thus, if the rating scale extends from “very light” to “very heavy,” judges will use or define “very light” to signify the lightest stimulus in the series, and so on. This implies that the stimulus value most likely to be rated as neutral or average will be that at the *midpoint* of the range between the two extremes. The frequency principle states simply that, other things being equal, judges will attempt to use the separate rating categories with equal frequency. Thus, if there are 50 stimulus presentations and the rating scale has five categories, the ideal would be to place 10 stimuli in each of the five categories. This implies that the stimulus most likely to be rated neutral would be that closest to the *median* of the distribution.

The critical issue is then how evenly spread the stimuli are. If the distribution is very even, the midpoint, median, and mean will all be close to each other. As soon as any skew appears, the mean, that is, the AL, will lie between the midpoint and the median. Thus, when Helson predicts AL from an averaging process, without reference to features of the rating scale, Parducci also predicts that AL will be close to an average because of how judges deliberately (by implication) use the rating scale: Their judgments are a *compromise* between an attempt, on one hand, to use the separate scale categories to cover equal proportions of the stimulus range, and an attempt, on the other hand, to use the categories with equal frequency. Both principles assume that judgment is a form of communication, rather than merely a passive form of adaptation. However, Parducci by no means excludes the possibility of changes in subjective representation or sensation as a consequence of distributional features of the experienced stimuli. That this might occur even without the explicit requirements of a judgment task becomes clear from his extension of the range–frequency compromise to provide a relativistic theory of happiness (Parducci, 1984). One of the paradoxes in the literature on happiness is that most people think of themselves as (on average) relatively happy, i.e., they rate their happiness as being above average most of the time. Parducci reasons that, if Helson’s view of adaptation were the whole story, everyone should habituate to their typical level of happiness, and so regard themselves as neither especially happy nor unhappy. However, suppose someone’s distribution of life events was positively skewed, i.e., characterized by a majority of at least mildly positive experiences, offset by a few disasters. In that case, the median of their distribution would lie above the

midpoint and, since AL (here, the kind of event that would be felt to be merely neutral) would lie between the median and the midpoint, most events would be experienced as better than average, i.e., on the happy side of the neutral point.

### Attitudinal judgment

Leon Thurstone's (1928) classic paper entitled "Attitudes can be measured" not only provided a launch pad for social judgment research but was a turning point for social psychology generally. Quite simply (along with Thurstone and Chave's, 1929, book, *The Measurement of Attitude*), it introduced the principle of quantitative measurement into the social arena. The basic argument was straightforward:

It will be conceded at the outset that an attitude is a complex affair which cannot be wholly described by any single numerical index. For the problem of measurement this statement is analogous to the observation that an ordinary table is a complex affair which cannot be wholly described by any single numerical index. So is a man such a complexity which cannot be wholly represented by a single index. Nevertheless, we do not hesitate to say that we measure the table. The context usually implies what it is about the table that we propose to measure . . . Just in the same sense we shall say here that we are measuring attitudes. We shall state or imply by the context the aspect of people's attitudes that we are measuring. The point is that it is just as legitimate to say that we are measuring attitudes as it is to say that we are measuring tables or men.

(Thurstone, 1928, p. 530)

The "aspect of people's attitudes" on which Thurstone then focused was the degree of favorability or unfavorability (approval or disapproval) toward an issue or attitude object. In other words, he conceptualized attitude measurement as the task of quantifying each individual's viewpoint on an attitudinal continuum, analogous to the (subjective) psychological continuum of psychophysics, from unfavorable to favorable, or anti to pro. As in psychophysics, subjective or psychological magnitudes cannot be observed directly, but need to be inferred from expressed judgments. Thurstone developed a number of techniques for doing this, but the one of most relevance here (the "method of equal-appearing intervals") relied on respondents expressing their agreement or disagreement with a series of attitude statements. (The by now ubiquitous use of Likert scales derives directly from this early work.) The basic logic of this technique was very simple: If person A, on average, tends to agree with statements that express moderately pro attitudes on some issue (e.g., moderate support for a political candidate), whereas person B tends only to agree with statements expressing an extremely anti attitude (e.g., extreme opposition to the candidate), then we're entitled to describe person A as *holding* a moderately pro or favorable attitude, and person B as *holding* an extremely anti attitude.

The next requirement, however, is to produce a numerical index to represent this attitude continuum, i.e., to translate

"moderately pro," "extremely anti," etc. into quantitative scores that can be used to compute the degree of difference between any two individuals, the average attitude of any group of individuals, and any other such operation that would be permitted in the case of similar (interval) measures. In psychophysics, one has the physical continuum of objective magnitudes as a benchmark, so could an analogue be found here? Thurstone's proposal was that one could derive more or less objective measurements of the favorability *expressed* by each statement by asking an independent group of "judges" to rate or sort them in terms of a scale from extremely unfavorable to extremely favorable on the issue.

The task fulfilled by these judges was quite specific: They were *not* being asked to say if they personally agreed or disagreed with each statement (the standard instructions stressed this point); instead, they were being asked what kind of attitude was being *expressed* by each statement. (Another way of presenting this, which we've used in our own work, is to ask judges to think how anti or pro would be the kind of person who would have made each statement.) Conventionally, Thurstone used an 11-point scale from extremely unfavorable to extremely favorable toward the issue, with none of the intermediate response categories carrying verbal labels. The median or mean rating by the group of judges of each statement was then taken as the "scale value" of that statement (or "item"). The judges' task was then done. Then the researcher, after excluding items that produced too variable a set of ratings and applying other procedures, would select a final series of items with scale values distributed evenly along the anti-pro continuum. These could then be presented to the respondents whose attitudes the researcher was trying to measure. These respondents then expressed their personal agreement/disagreement with each item. In the original method, each individual respondent's attitude (i.e., location on the continuum covered by the 11-point scale) could be computed from the mean scale value of the statements with which he or she agreed.

### Context effects in attitudinal judgment

While all this was going on, the possibility that context effects might also be apparent in judgments of social stimuli remained largely unexplored. Much of the responsibility for this, curiously, lies at the door of Thurstone—or at least those working under his influence—who, having once adapted psychophysical procedures for attitudinal judgment, seemed to constrain the kinds of questions that were then asked for about a quarter of a century. For Thurstone, the procedure of having judges rate attitude statements in terms of favorability toward an issue was merely a means to an end—the construction of a "valid" attitude scale. Furthermore, he argued, that, for a scale to be valid, "the scale values of the statements should not be affected by the opinion of the people who help to construct it" (Thurstone & Chave, 1929, p. 92). The way to test this, he suggested, was to present a common set of items to different groups of judges, known or presumed to have contrasting opinions on the issue in question (e.g., members of militarist and pacifist organizations)

and compare the set of scale values derived from their respective ratings. “If the scale values are practically the same in the two scales, the validity of the method will be pretty well established” (Thurstone & Chave, 1929, p. 93).

But what does “practically the same” mean? For measurement theorists, so used to establishing validity by looking for *correlations* between scale scores and an external criterion, this was interpreted as a demand that the scales constructed from the ratings given by different groups of judges should be highly correlated. And indeed, this was consistently found to be the case. So several “Thurstone scales” were happily developed to measure attitudes on a wide range of issues. But a moment’s thought is enough to remind us that two sets of scores can be highly correlated but extremely different in absolute terms. Measuring temperature in degrees Celsius and degrees Fahrenheit yields sets of numbers that are *perfectly* correlated but are definitely not “the same,” although they both measure the same physical property.

Nonetheless, the assumption that the scale values of attitude scale items were unaffected by judges’ attitudes remained essentially unchallenged until a classic study by Hovland and Sherif (1952). They were struck by “the paradox that some leading texts in social psychology say in the chapter on perception and judgment that judgments are greatly affected by the individual’s attitudes and motives while in the chapter on scaling methods they state that judgments of the meaning of items are unaffected by the positions of the judges who do the sorting” (p. 23). They therefore undertook a modified replication of a study by Hinckley (1932) that involved the construction of a scale to measure (to give it its original title) “attitude toward the social position of the Negro.” Where Hinckley, for a variety of reasons (see Eiser & Stroebe, 1972; Hovland & Sherif, 1952), concluded that judges’ own attitudes had no effect on their judgments, Hovland and Sherif demonstrated clear differences between the ratings of Hinckley’s original items given by groups of judges they describe as Anti-Negro white, “Average” white, Pro-Negro white and Negro. Most strikingly, the four groups in the above order rated an increasing proportion of the statements as expressing viewpoints antagonistic to Blacks. Hovland and Sherif interpreted this result as suggesting that judges were using their own opinion as a comparison standard or “anchor.” Groups whose own attitudes were more pro-Black would thus be using a more pro-Black standard (in AL theory terms, their ALs would be shifted toward the more pro-Black extreme). Hence, more items would fall on the unfavorable side of their standard (or AL) and so be judged as unfavorable (i.e., anti-Negro)—a simple example of judgmental *contrast*. Except that things weren’t quite that simple. Not all the items in the scale showed the same shifts as a function of judges’ attitudes. A minority of items near the favorable (i.e., pro-Black) extreme displayed an opposite trend, being rated as more extremely favorable by Black and pro-Black judges. Hovland and Sherif describe this effect as *assimilation*.

Following this study, the terms “assimilation” and “contrast” became increasingly established in the theoretical vocabulary

of social psychology. Indeed, the full title of Sherif and Hovland’s (1961) book was *Social judgment: Assimilation and Contrast Effects in Communication and Attitude Change*. The question of when assimilation and when contrast occurs resonates through social judgment research right up to the present day. However, it is worth pausing to point out some of the empirical and conceptual ambiguities that surrounded their introduction. Firstly, although Hovland and Sherif were keen to draw links with theories of psychophysical judgment, at the time there was overwhelming evidence for contrast effects, but little for assimilation effects. As Sherif notes in his preface to his (1961) book with Hovland, published shortly after Carl Hovland’s death earlier that year: “in the judgment literature, contrast effects were considered to be established, but, on the whole, assimilation effects either were ignored or, when noted, were often referred to by some such label as the ‘reverse of contrast’” (Sherif & Hovland, 1961, p. vi).

Nonetheless, there is a difference between demonstrating assimilation effects and establishing that they reflect the same kind of anchoring process presumably responsible for contrast effects. Even when they conducted their own psychophysical experiments using judgments of lifted weights, Sherif, Taub, and Hovland (1958) established only that, under certain conditions, the introduction of a comparison standard or anchor could lead to assimilation effects for a stimulus series *as a whole*. They failed to demonstrate with physical stimuli what appeared to be the case with attitude statements: that, *within the same series*, some stimuli would be assimilated and others contrasted. Nonetheless, by the time of Sherif and Hovland’s (1961) book, the view seemed well established that (in social judgment at least) items close to an anchor would be assimilated and those further away would be contrasted.

The second problem relates to the apparent imbalance in the strength of assimilation and contrast effects, and underlines the importance of looking critically at the content of measurement instruments (both here and more generally). The fact that fewer items were assimilated than contrasted could have just been due to there being fewer favorable than unfavorable items in the series presented for judgment. Of course, terms such as “favorable” and “unfavorable” are relative to their cultural and historical context—Hinckley (1932) would have merely been attempting to cover the range of discriminable attitudes on the issue at the time—but to more modern eyes and ears, the distribution seems extremely skewed. At the “unfavorable” extreme we find extraordinary items such as “No Negro has the slightest right to resent, or even question, the illegal killing of one of his race”—extraordinary not just because of what is said but because such a statement would not have made it through the piloting and selection process into the 1932 scale unless sufficient people had agreed with it. At the “favorable” extreme, however, we find at best cautious assertions of equal rights (“I believe that the Negro is entitled to the same social privileges as the white man”), with some clearly segregationist sentiments still attracting “moderately favorable” scale values.

So the point is that we cannot say, on the basis of the Hovland and Sherif's (1952) data, anything about the "natural" balance between the strength of assimilation and contrast effects.

## From perception to cognition

### *Categorization and accentuation*

My own early research in this field was focused on developing an alternative to the assimilation–contrast interpretation of the effects of judges' own opinions on their ratings of attitude statements. We (Eiser, 1971; Eiser & Stroebe, 1972) took as our starting point Tajfel's (1957, 1959) studies of the effects of value on perceptual judgments of size. Earlier work (Bruner & Goodman, 1947; Bruner & Postman, 1948) had suggested that valued stimuli, such as coins, tended to be overestimated in size compared with stimuli of the same physical dimensions, but no value—a so-called "perceptual overestimation" effect. Tajfel's insight was that the extent of overestimation was related to the extent of value for sets of objects where size and value were correlated, so that "superimposing" value on a physical dimension led not simply to overestimation of size, but an *accentuation of perceptual differences* between less and more valuable objects. He went on to argue that this was because participants would use value as a "cue" to size, and so (the "so" implying a notion that extra cues make discriminations easier) accentuate the differences between adjacent stimuli that differ in value as well as size. He extended this argument to claim that any arbitrary correlated cue or classification superimposed on a continuum of judgment ("correlated" meaning, for example, that all small stimuli were labeled as belonging to one class and all large stimuli to another; see Tajfel & Wilkes, 1963) would lead to an accentuation or polarization of judged differences between stimuli distinguishable from each other in terms of the superimposed cue or classification. This, he went on to argue, constitutes much of the cognitive basis for prejudice and intergroup discrimination (Tajfel, 1969, 1982; Tajfel, Sheikh & Gardner, 1964). Thus, these studies form a bridge between the earlier paradigm of perceptual judgment and more general cognitive approaches to social judgment.

Within the narrower context of attitudinal judgment, Tajfel's accentuation principle offered an appealing alternative to Sherif and Hovland's (1961) emphasis on anchoring process to explain assimilation–contrast effects. It allowed for judgmental shifts in opposing directions (polarization toward the scale extremes) within the *same* stimulus series. It also made no assumption concerning the balance of strength of so-called "assimilation" and "contrast" effects across different stimulus distributions. However, it did require firstly that a superimposed cue or classification could produce accentuation effects for attitude statements, and secondly that there was something within the typical attitudinal judgment paradigm that fulfilled the requirements of a superimposed cue.

The first requirement was easy to settle empirically. I presented student participants with 64 statements on the

nonmedical use of drugs to be rated on a scale labeled "extremely permissive" to "extremely restrictive" (Eiser, 1971). In the control condition, participants were told that the statements were drawn from newspapers. In the experimental condition, all the statements were identified as quotes from one of two fictitious newspapers, with the 32 more prodrug items attributed to one paper and the 32 more antidrug items to the other. As predicted, in the experimental as compared with the control condition, judgments of the attitude statements were significantly more polarized; that is, prodrug items were rated as more "permissive" and antidrug items as more "restrictive."

Evidence satisfying the second requirement had almost been already supplied by Sherif and Hovland (1961), although they did not quite recognize or interpret it as such. As they developed their assimilation–contrast model, they addressed the issue of where (in terms of distance of a judged stimulus from an anchor) assimilation switched over to contrast. When a person's own opinion was the anchor, they argued that statements falling within the person's "latitude of acceptance" would tend to be assimilated, whereas those falling within the "latitude of rejection" would be contrasted (these being separated by a "latitude of non-commitment"). The "latitude of acceptance" is simply the range of opinions that, even if not exactly coinciding with the judge's own opinion, would nonetheless be acceptable, and so can be easily identified for the statements with which the judge personally agrees. So what Sherif and Hovland are effectively saying is that judges should assimilate statements they agree with and contrast statements they disagree with. In other words, but not the words Sherif and Hovland use, judges' agreement/disagreement with the statements fulfills the requirement of a (subjectively) superimposed cue and should lead to an accentuation of judged differences between accepted and rejected statements. For more extreme judges, who agree with statements toward one end of the scale and disagree with statements toward the other end, this should lead to more extreme or polarized judgments. This should apply *both* to judges with extremely pro and to those with extremely anti attitudes on any issue.

So is this what happens? Yes and no. In the Hovland and Sherif (1952) study, as well as subsequent American studies using the same issue of racial attitudes (Selltitz, Edrich & Cook, 1965; Zavalloni & Cook, 1965), participants with more pro-Black or egalitarian attitudes gave more polarized ratings than more "neutral" judges, as predicted. But the reverse was true for the most anti-Black participants: They clumped more of their ratings close to the middle categories of the rating scale, avoiding the extreme categories. A similar apparent anomaly occurred in my own (Eiser, 1971) study: More prodrug judges (as identified by their own agreement/disagreement with the items) gave more polarized ratings than neutral judges, but anti-drug judges gave the least polarized ratings of all.

So what is going on with the anti judges? Perhaps part of their difficulty, as suggested by Romer (1983), might stem from semantic confusion, since the typical judgment task requires them to judge pro items as "favorable" (toward the issue) when



their own attitude toward these items is unfavorable. An additional possibility is that such individuals are reluctant to effectively declare that they are racist (say) by labeling the statements they agree with as “unfavorable toward Blacks.” A separate argument is needed to explain the asymmetry in my own drug-attitudes study, but here again some kind of social desirability process might be operating. This study was conducted within a social context (London in the late 1960s) where libertarian attitudes among young people, and students in particular, were very much the norm.

### **Judgmental language**

To get a handle on such possible processes, we conducted a series of studies that broke with the standard Thurstone tradition by requiring judges to rate the same items on multiple scales (Eiser, 1973; Eiser & Mower White, 1974, 1975; Eiser & van der Pligt, 1982). These scales were deliberately chosen to manipulate the so-called “value connotations” of the judgment scale labels. Thus, when the issue used was that of adolescents’ attitudes to adult authority (Eiser & Mower White, 1974, 1975), on some scales (P+), the descriptively proauthority end was labeled more positively and the antiauthority end more negatively (e.g., uncooperative–cooperative), whereas on other (A+) scales (e.g., independent–dependent) the direction of implicit value was reversed. The main finding from these studies was that polarization of judgment depended on an interaction between judges’ own opinion and the value connotations of the rating scale. On P+ scales, pro judges gave the most polarized ratings and anti judges least, whereas on A+ scales the contrary trend was found. Hence, for personal agreement/disagreement with attitude statements to lead to increased polarization according to the accentuation principle, the presented rating scale must be evaluatively consistent (or at least not inconsistent) with judges’ own position (see Eiser, 1990 for a fuller review of these studies).

Our demonstrations of the importance of value connotations in the language judges are required to use coheres with a more general insight concerning what judgment experiments involve. Providing a rating on a scale (any scale) is not a private act, say, of introspection; it is a *communicative* act between the participant and the researcher. A questionnaire is, quite literally, a set of *questions*, and participants’ responses a set of *answers*. Social judgment studies in particular rely heavily on natural language for presenting objects of judgment (attitude statements being the prime example), for identifying the attributes of these objects to be differentiated and/or categorized, and for expressing these differentiations and categorizations. Behind any suggestion of the role of social desirability factors in judgment is the implication that participants are mindful of how their responses may be interpreted by the researcher. However, the argument can go even broader. Filling in a questionnaire is like engaging in a conversation (albeit at a distance) and thus may be influenced by many of the same concerns and norms that shape more everyday conversations.

This insight has been developed extensively by Schwarz (e.g., 1994) with reference to the effects of a number of simple variations in questionnaire format including question order and the labeling of response scales. Schwarz invokes Grice’s (1975) maxims of conversational logic to explore some of the consequences of respondents trying to infer what researchers expect and want to know, while, for instance, avoiding redundancy and ambiguity in their answers, as far as possible. One line of work considers how participants treat questions of potentially overlapping reference, such as ones asking about their general life satisfaction on one hand and satisfaction with specific aspects of their life (e.g., job, marriage) on the other. In simplified terms, when participants evaluate specific aspects before rating their life in general, their ratings of specific and general satisfaction correlate more highly than when the general question precedes the more specific ones (Schwarz & Strack, 1991; Schwarz, Strack, & Mai, 1991b; Strack, Martin, & Schwarz, 1988). The argument here is that, when participants feel they have already provided information about their general life satisfaction, they try to avoid redundancy in their subsequent more specific answers by instead communicating how satisfied they are with, for example, their job as distinct from other aspects of their lives.

Another set of effects has strong resonances with Parducci’s (1963) range–frequency compromise, and relates to the suggestive effects of response scales where continuous numerical quantities are presented in terms of a limited number of response categories covering different ranges. For example, Schwarz, Hippler, Deutsch, and Strack (1985) asked participants to estimate their average daily TV viewing in terms of six categories, ranging either from “up to half-an-hour” through to “more than two-and-a-half hours,” or from “up to two-and-a-half hours” to “more than four-and-a-half hours.” Thus, the top category of the first (low-frequency) scale is equivalent in absolute terms to the combined top five categories of the second (high-frequency) scale. If the format of the response scale had no effect, the absolute estimates of TV watching derived from the two scales should be comparable. In fact, however, more than twice as many participants (37.5% vs. 16.2%) indicated that they watched more than two-and-a-half hours when presented with the high-frequency scale. The argument is that the scale categories conveyed to participants the range of answers (they thought) the researchers expected. (See also Eiser & Hoepfner, 1991, for a demonstration of this effect in relation to health and environmental attitudes.)

Such findings take us well beyond the early emphasis on so-called “distortions” or “shifts” in judgment of attitude statements that were the early focus of attitudinal judgment research. Indeed, the relevance of such linguistic factors was identified, though not pursued, by Sherif and Hovland as a potential area for future research.

An important area for research lies in relating the study of judgment to that of language. Language constitutes a framework within which concepts develop. Studies in the area of social judgment frequently use words as stimuli to be

categorized. What are the effects of different verbal forms on our interpretation of the stimuli and judgment of them? What constitutes verbal ambiguity? The extent to which distortions in interpretation result from ambiguity may be closely related to the type of language structure utilized.

(Sherif & Hovland, 1961, p. 201)

This area remains one with underexplored potential. If language is a framework for the development of concepts in general, there is an even stronger case for arguing that it helps shape the development and formation of attitudes in particular. Although the Eiser and Mower White (1974, 1975) studies were designed to reinterpret assimilation–contrast effects in terms of an accentuation principle, we attempted to draw some broader implications for attitude theory. The finding that judges give more polarized ratings on scales labeled so that their “own end” (i.e., the extreme closer to their own position) was evaluatively positive suggests that individuals *prefer to construe and represent* attitude issues in ways that enable them to maintain a positive view of their own position, and hence of themselves. Alternative forms of judgmental language thus offer alternative modes of representation, any of which will be seen as more relevant, applicable, and *salient* by some people, and less so by others (polarization of judgment being an indicator of salience).

This notion of “dimensional salience” (van der Pligt & Eiser, 1984) helps us address arguably one of the most fundamental, but also simplest to state, questions for attitude theory: Why do *different* people hold *different* attitudes on any given issue? Part of the answer has to be that individuals with different attitudes selectively regard different dimensions of the issue as salient. Everyday examples of this are easy to find. One issue we have studied in our own research is that of nuclear energy, where potential danger is clearly a more salient aspect for opponents than supporters (Eiser, Spears & Webley, 1989; Eiser & van der Pligt, 1979; van der Pligt, van der Linden, & Ester, 1982). The versatility of language allows for the same issue to be described in ways that emphasize different aspects, and convey very different value implications, and hence alternative modes of linguistic representation can help construct, as well as express, attitudinal differences (Eiser & Pancer, 1979; Eiser & Ross, 1977).

### Why categorize?

I noted earlier that there is a tension between (and within) theoretical accounts that emphasize the functions served by judgmental processes and those that seem more oriented toward explaining errors and biases. It is difficult to characterize Sherif and Hovland’s work as falling wholly within either camp. They clearly saw processes of evaluation and categorization as fulfilling important functions, although they also viewed assimilation and contrasts as “displacements” brought about by the influence of motivation. For other early researchers (e.g., Bruner, Goodnow & Austin, 1953; Campbell, 1956; Tajfel, 1959), a major concern was how individuals might form *simplified* concepts or representations of complex (i.e., multiattribute)

objects. So, within the priming paradigm, contrast could be interpreted as an example of interclass accentuation if the target and prime belong to different classes, whereas assimilation could reflect a reduction of intraclass differences if they belong to the same class. However, it is no longer so clear what motivational functions (in terms of simplicity or saved cognitive effort) are being served by such judgmental shifts.

Simplification is at least partly achieved by using additional cues as sources of information concerning the objects that are in some way deemed to be predictive of the attribute being judged, i.e., “peripheral” as distinct from “focal” attributes (Eiser & Stroebe, 1972). For Campbell (1956), this linkage between peripheral (predictive) and focal (judged) attributes relied explicitly on associative learning—as he termed it, the formation of a “composite habit.” Tajfel engaged less with learning theory as such, but in Tajfel & Wilkes (1963) and subsequently, participants were presented with novel stimuli and superimposed cues, and so had to learn an association between these during the course of the experiment. However, there is a difference conceptually between using extra information to simplify the demands of the judgment task and using it to help form simplified representations of the objects of judgment.

Intuitively it is easy to appreciate that categorization provides greater cognitive simplicity by relieving perceivers of the requirement to process each new stimulus from scratch, rather than relying on its similarity to previously encountered objects. But why should this necessitate an exaggeration of differences between categories? Another way in which cognitive complexity could be reduced could be through selectivity: that is, by limiting the *range* or scope of objects being seen as relevant to any judgment or decision. Back within the psychophysical paradigm, Brown (1953) showed that stimuli needed to be seen as relevant to the judgment task to influence AL. An especially interesting set of ideas is offered by Kahneman and Miller (1986). According to their norm theory, events are judged relative to some “norm” that is appropriate to the particular context. Thus, experiences such as surprise indicate that an event is discrepant from a relevant norm. This could sound quite like a reworking of AL theory, except for the fact that Kahneman and Miller propose that “norms are computed after the event, rather than in advance” (1986, p. 136). In other words, individuals do not always have ready-made standards available but construct these once they are required to judge or evaluate an event. This can involve associative memories, as well as counterfactual thinking of what alternative events or outcomes might have occurred.

A series of studies by Manis and colleagues attempted to reconcile these approaches. The basic design involved experimentally superimposing a classification on a stimulus series and then looking at judgments of stimuli at the borderline between the two categories. In Manis, Paskewitz and Cotler (1986), the stimuli in question were supposed responses by psychiatric and general hospital patients to tests designed to measure psychopathology. In the first (induction) phase of the experiment, these test responses were presented as coming from patients from two hospitals (“Central” and “Metropolitan”),

with those test responses prescaled as indicative of greater psychopathology being all attributed (in the critical condition) to Central patients, and those less disturbed to Metropolitan patients. The purpose of this phase was to “induce” an association between the hospital names and the perceived psychopathology of the patients. Participants were then presented with a series of paired comparisons between midscale items indicative of moderate or borderline disturbance (within each pair, one item was attributed to a Central and the other to a Metropolitan patient) and had to rate which was the more disturbed. The crucial finding was that those participants who had been induced to see “Central” as more disturbed and “Metropolitan” as less disturbed overall tended to rate the borderline Central patients as *less* disturbed than the borderline Metropolitan patients in the paired comparison task. In a follow-up, Manis, Nelson, and Shedler (1988) confirmed this finding when the items in the induction phase were extreme (in terms of disturbance vs. normality), but not when they were moderate.

Manis and colleagues interpret these effects as indicative of *intraclass contrast*, that is, a shift in judgment of the borderline (target) items away from the norm for their class. At face value, this appears to pose problems both for accentuation theory notions of categorization (Eiser, 1971; Tajfel, 1959) and inclusion–exclusion models of priming, to be discussed shortly (Martin, 1986; Schwarz & Bless, 1992). However, there are difficulties with this interpretation. Apart from the fact that the effects barely reach conventional significance levels, the inferred contrast process is never directly tested. Specifically, intraclass contrast should mean that the borderline items shift *away* from the rest of their class; i.e., the borderline Central patients should become seen as less disturbed, whereas the remaining “induction” Central patients remain seen as disturbed or even more disturbed than before. But no ratings of the induction items are presented.

We therefore conducted a modified replication of the Manis et al. experiments, using the same stimuli (Eiser, Martijn & van Schie, 1991). Crucially, we also required participants to rate the induction items. Compared with a control condition without any hospital names, participants in the experimental conditions rated the borderline items much as in the Manis et al. studies, that is, overall the borderline Central patients were seen as less disturbed and the borderline Metropolitan patients as more disturbed, contrary to what accentuation theory would predict. However, the remaining Central patients (those forming the induction series) were *also* seen as less disturbed (and the remaining Metropolitan patients as more disturbed) in the experimental compared with control condition. In other words, the borderline items did *not* shift away from the induction items from the same hospital; instead, the classes of items shifted as a whole in the direction of greater overlap between the perceived distribution of psychopathology in the two hospitals. What appeared to be an effect of intraclass contrast, therefore, was in fact an effect of interclass assimilation. It is probably not worth dwelling for too long here on reasons for what is an unusual effect, but it is still worth noting the fact that, under particular

conditions, superimposed classifications do not always lead to increased polarization and interclass accentuation. We reasoned that a limiting condition for accentuation might be the distinctiveness of any superimposed classification. If the experimentally defined classes are less easy to distinguish, or are seen as overlapping, attention to the class membership of the items presented for judgment may even lead to an increase in this perceived overlap, rather than a sharpening of the separation between the classes. The potential implications of this for stereotype reduction are intriguing. Regarding social judgment processes, however, the main message is that the effects of superimposed cues may result in different kinds of judgmental shifts, depending on their learnt associations.

### The influx of new paradigms

Around the early 1980s, there was a sharp shift in research emphasis away from the traditional concerns exemplified by Sherif and Hovland’s early work and the various responses to it. There were several reasons for this. Partly there was a sense in which the attitudinal judgment paradigm had run out of questions needing answering—at least in the form in which they were originally posed in relation to the assimilation–contrast model. Remember that the home base for this research was within the Yale Communication and Attitude Change Program, founded by Hovland. So this work was intended not just to throw light on processes of judgment, but to use these to draw out broader implications for attitude change and, in passing, attitude measurement.

In both regards, however, the base for drawing such broader implications was too narrow. I have said little here about the assimilation–contrast model of attitude *change*, largely because I see a clear discontinuity between how Sherif and Hovland use the concepts of anchoring and latitudes of acceptance and rejection to predict shifts in judgment of attitude statements and other objects (the topic of this chapter) and how they use these concepts to predict attitude change as a function of discrepancy between positions held by an audience and those advocated by a communicator. The latter falls outside the remit of this chapter, but Eagly and Chaiken (1993)—who, interestingly call *this*, and not Hammond’s (e.g., 1955) approach, “social judgment theory”—attribute the demise of the assimilation–contrast model of attitude change to it having a small “quarry,” in that “few variables other than source credibility have been researched or even identified” (p. 379).

As for attitude *measurement*, again only a fairly small corner of the field is covered. Simply stated, nobody was particularly interested any more in the cumbersome process of constructing Thurstone equal-appearing interval scales. Measuring the locations of individuals on a single dimension of favorability toward a single issue came to be seen as less interesting theoretically than looking for interrelationships between different attitude constructs. The business of combining separate items to form a single scale score, where required, became a simpler and speedier empirical matter, with the advent of statistical software

to perform reliability analyses and calculate Cronbach's alpha. At a more applied level, there remains a huge demand for attitude measurement in the sense of attitude surveys, but the focus here is generally on representing attitudes at the level of the group or population, rather than at the level of the individual. There continues to be excellent theoretical and methodological research by social psychologists into the judgmental processes affecting survey responses, with that by Norbert Schwarz (Schwarz, Groves, & Shuman, 1998) and Jon Krosnick (Krosnick, 1999) particularly worthy of mention.

And what of work on the accentuation principle as an alternative to the assimilation–contrast model? Within social judgment “proper,” the issue here became more one of understanding more general processes of categorization, as well as attitude polarization, than anything closely tied to the rating of attitude statements. Although my application of this principle was derived from Tajfel's (1957; Tajfel & Wilkes, 1963) work on perceptual judgments, the main aim of subsequent work by Tajfel and his other associates (as well as his own original inspiration) was to understand the cognitive underpinnings of phenomena such as racial prejudice, intergroup discrimination, and social identity (e.g., Tajfel, 1982). Other researchers with a background in attitudinal judgment such as Manis (1960, 1961) also turned their attention to stereotype research (see Manis et al., 1986, 1988, as described above). So there was a fair amount of diaspora in terms of social judgment researchers moving, at least partly, into other areas. (My own interests broadened to include attitude formation as well as judgment, especially in the context of health and environmental issues.)

Still less could social judgment depend on any historical links with psychophysics, which itself was changing rapidly in the direction of ever more fine-grained analysis of perceptual mechanisms, with the advent of modern neuroscience being just around the corner. If constructing Thurstone attitude scales was yesterday's news, sets of beautifully engineered brass cylinders for lifted-weights experiments were, quite literally, museum pieces. And so the field was left open to the influx of new paradigms, or conceptual orientations. These, to some extent, vied for supremacy with each other, before finding some common ground and the potential for compromise. As far as all of these were concerned, their cause had already been facilitated by a shift away from thinking of judgment in terms of metaphors derived from sensory perception toward thinking of it as a *cognitive* process. Even so, what *kind* of cognitive process or processes it involved remained unresolved. These paradigms (or orientations) respectively emphasized the links between judgment and memory, decision-making and learning. Of these, the “friendliest” was that concerned with *memory* processes. Research within this tradition could almost be characterized more as immigration than invasion, since many of the concepts developed within earlier social and attitudinal judgment work were adopted or modified to fit a new range of research questions. Research on *judgment and decision-making* posed a prospect more “threatening” to the continuity of work on the topics discussed so far in this chapter. Here was a more self-reliant and

cohesive research community, with strong external backers (linkages with other disciplines and areas of policy), with a proprietary claim on the term “judgment” but less history of obvious interest in traditional social judgment issues such as assimilation–contrast, attitudes, and categorization. The final orientation to consider is that of judgment as *learning*. Research on learning had become so deeply unfashionable with social psychologists during the so-called “cognitive revolution” around the 1960s (at the time it seemed to imply a far too deterministic view of human nature) that its continued relevance to, and presence within, theories of social judgment went almost unnoticed. The history of learning–theoretical perspectives on social judgment is thus less one of invasion so much as of “infiltration” and a gradual rehabilitation of a minority that continues to the present day.

### Judgment as memory

The first of these new paradigms, then, was that characterized by a more direct concern with memory and its effect on a range of cognitive and behavioral processes. This involved a kind of alliance or confluence of two traditions. The first is what used to be called “person perception” or “impression formation,” and focused on how people formed judgments of the attributes of other individuals on the basis of specific cues (e.g., Anderson, 1962, 1967, 1981; Tagiuri & Petruccio, 1958). The second, of more recent origin, took as its starting-point the question of how (social) information is represented in, and retrieved from, memory. One of the important things about memory is thus that it allows for judgments to be influenced by information that is not at the forefront of a person's conscious mind, and thus opens the door to a consideration of more “automatic” and less conscious processes.

One of the main entry points for this “immigration” into social judgment was the notion that judgments could be relative not simply to anchors provided by the experimenter, but to “subjective standards” constructed by the participant on the basis of imagination or prior experience (see Eriksen & Hake, 1957, for an early recognition of the role of subjective standards). Where might such subjective standards come from? Surely, from remembered experiences and associations. As far as attitudes are concerned, these too are a kind of memory and their effects on judgment should be interpreted in terms of more general memory processes. Perhaps attitudes are not so easy to manipulate, but other memories may be, and so the impact of memory on judgment should be amenable to experimental investigation. Matching psychological to physical magnitudes was likewise no longer a concern. In fact, it was viewed as an impediment, since it demanded both a precise specification of a dimension along which a series of stimuli was distributed *and* enough stimuli that the parameters of their distribution could be described. Far simpler to choose single objects of judgment of more intrinsic interest (e.g., real or hypothetical persons) and see when these might be assimilated to or contrasted from some standard in memory.

This broadening of theoretical perspectives to include the role of memory was definitely a positive move, but, in hindsight, the expansion was narrower than it appeared at the time—even just with respect to memory processes, let alone a range of other social cognitive processes less represented in this self-styled “social cognition” genre. The main limitation was that the focus was strongly on accessing or retrieving information from memory, and far less, if at all, on how memories might be acquired in the first place. Even so, parking this criticism for the time being, the introduction of the concept of memory accessibility opened up a whole new range of possibilities of studying the influence of stored representations and memories that might fulfill the role of subjective standards. One might not be able to manipulate individuals’ personal histories that gave rise to their memories, but one could manipulate which of their memories individuals were likely to *use* in specific contexts.

### Priming

The key notion here is that of associative memory. In simple terms, separate memories are associated with each other to varying extents. According to the idea of “spreading activation” (e.g., Bower, 1981), “activation” of one memory can “spread” to other memory with which it is “associated.” A “memory” can stand for anything—a linguistic concept, a remembered episode, a perceptual sensation, a moral value, an attitude, a stereotype, an action tendency or intention. This breadth of definition is both the strength and the potential weakness of this approach, since the key concepts (deliberately in inverted commas) deserve more precise specification. Yet the basic idea is simple enough. A *prime (noun)* is any piece of information, word or stimulus, typically with symbolic meaning, presented to an individual that can set in train, i.e., *prime (verb)*, a string of associations so that other concepts, thoughts or memories are more likely to come to mind and/or be acted upon.

This is certainly not the place to attempt to review the large literature on priming effects (see DeCoster & Claypool, 2004, for a meta-analysis), but it bears on the topic of this chapter since one of the principal questions has been when priming manipulations lead to assimilation effects, and when to contrast. Nonetheless, at the risk of oversimplification (since there are several procedural factors that can influence the strength and reliability of any effects), the typical outcome of priming is assimilation. Early demonstrations involved rating the characteristics of a hypothetical target person on the basis of an ambiguous behavioral description, after priming with trait constructs such as hostility (Srull & Wyer, 1979) or of contrasting evaluation (e.g., adventurous vs. reckless; Higgins, Rholes, & Jones, 1977). Following such manipulations, the target person would be rated as more hostile, adventurous, etc. In other words, the ambiguous behavioral information would be interpreted as consistent with, or confirming, the previously activated concept. Crucially, this occurs *without* participants recognizing any connection between the priming and rating tasks, and even when the priming stimulus is presented

subliminally to ensure that participants are unaware of its occurrence (e.g., Bargh & Pietromonaco, 1992).

What makes these demonstrations interesting is precisely the subtlety of the manipulations: Social perceptions, preferences and behavior appear to be influenced by information processed below the level of conscious awareness. So does this mean that, if participants are made aware of the prime, such effects will be stronger? In fact, no. Again at the risk of oversimplification, if the priming manipulation is made blatant, so that participants are fully aware of it, the typical effect appears to be one of contrast. There have been several attempts to explain the moderating role of awareness. Higgins (1996) favors a kind of misattribution interpretation for the effects of subtle primes: Participants treat any reactions or associations to the prime as attributable instead to the target stimulus, so that if, e.g., the concept of adventurousness has been activated by the prime, they will rate the target as more adventurous. Under conditions where participants realize that their thoughts of adventurousness emanate from a different source than the target, they are less likely to make such a misattribution. Indeed, they may then attempt actively to correct for the effects of any such irrelevant influence and so rate the target as even less adventurous, say, than in the absence of the prime—a contrast effect (Martin, 1986; see also Wegener & Petty, 1995, for a broader discussion of how participants may attempt to correct for assumed biases).

It would seem a shame, however, to dismiss such contrast effects to blatant or consciously recognized primes as merely an aspect of the microculture of the social psychological laboratory, with participants trying to second-guess the intentions of the experimenter. This would be beyond what Martin (1986) and Higgins (1996) wish to imply. Contrast is an extremely general phenomenon, as we have seen, and more the rule than the exception in other kinds of social judgment tasks. Perhaps the problem comes from treating “blatant primes” as exceptions to the typical priming paradigm, rather than viewing the priming paradigm itself as a rather particular—even peculiar—procedure for manipulating judgmental context. From such a perspective, introducing a blatant prime could be merely another way of adding a background stimulus to the judgmental context.

### Concepts, comparison standards and categories

But what kind of background stimulus? Before priming studies became the vogue, the terms “assimilation” and “contrast” were defined consensually as shifts in judgment of a target stimulus toward or away from a comparison standard. As we have seen, this definition (as in Sherif and Hovland, 1961) implied assumptions about the underlying processes that were open to challenge. There was always a certain circularity in attempting to explain observed shifts in judgment in terms of assimilation or contrast relative to a subjective standard (such as judges’ own attitude) and at the same time using such judgment shifts as evidence that judges were bringing such a subjective standard into play. Although it is straightforward to identify judges’ own

position on an attitude continuum, this by itself does not establish that judges use this as a reference point or anchor in the manner suggested by the assimilation–contrast model. Our alternative accentuation theory (Eiser, 1971, 1990; Eiser & Stroebe, 1972) also, of course, makes assumptions about underlying processes, but the crucial idea that judges categorize items according to their personal acceptability (or latitudes of acceptance–rejection in Sherif and Hovland’s terms) can be established simply from their recorded levels of agreement or disagreement with each statement, independently of any shifts (e.g., increased polarization) on any given scale of judgment hypothesized to follow from the use of subjective categories.

Within many priming experiments, the target stimulus is clearly identified (and unlike in most social judgment research, there is usually only one); however, it is often unclear what comparison standard is being used by participants. The “priming stimulus” is often not an object of judgment in its own right, but a means of rendering some concept, such as a personality trait, more accessible. With this increase in accessibility, it is assumed that the primed concept becomes part of the representation of the target stimulus. Thus, if the priming manipulation successfully increases the accessibility of the trait of adventurousness, the target person should be seen as more adventurous. In simpler terms, the prime influences the *interpretation* of information about the target.

Sometimes, however, experimenters have sought to increase the accessibility of a trait concept by having participants think about a prototypical *exemplar* of the trait in question, such as Hitler for the trait of hostility (Herr, 1986), Niki Lauda vs. Bill Clinton for (moderate) athleticism vs. nonathleticism and Frank Zappa vs. Steffi Graf for high vs. low use of drugs (Mussweiler & Strack, 2000). This same procedure can be generalized beyond judgments of persons. Herr, Sherman, & Fazio (1983) primed the attribute of ferociousness vs. gentleness in animals by having participants rate extremely ferocious (grizzly bear, tiger) and extremely gentle (dove, kitten) exemplars or animals’ size after rating, e.g., a whale or a flea. It seems plausible that the introduction of exemplars provides participants with a standard, leading more typically to contrast in comparative judgment.

But here again there are moderating factors. In the Herr et al. (1983) study, participants rated both real and imaginary animals (the assumption being that the latter would be more ambiguous). Ratings of (the ferociousness or size of) imaginary animals were assimilated to the attributes of primes that were moderate on the relevant dimension. However, more extreme primes produced contrast. If assimilation depends on a process of interpretation, it makes sense that this should have more influence on ratings of more ambiguous targets. But why should more moderate (and hence presumably weaker) primes apparently lead to *more* assimilation? The answer seems to be that, if assimilation depends on *inclusion* of attributes of the prime in the representation of the target, this will in turn depend on the target and prime being similar enough to each other to be seen as belonging to the *same category* (Martin,

1986; Schwarz & Bless, 1992). If the target and prime (or standard) can be *included* in the same category, this tends to lead to assimilation, whereas *exclusion* tends to lead to contrast. According to this argument, target–standard discrepancy along the dimension of judgment is thus essentially a proxy for category membership, in a way that echoes our own (Eiser & Stroebe, 1972) reinterpretation of Sherif and Hovland’s (1961) notions of latitudes of acceptance and rejection.

Once again, though, we need to be careful to specify what kind of judgment we are talking about. Just as in classical psychophysics it proved necessary to distinguish judgmental responses conceptually from psychological magnitudes, so ratings and representations are not always identical. Mussweiler and Strack (2000; see also Mussweiler, 2003) demonstrated that the same priming or anchor stimuli could produce contrast in “subjective” comparative judgments but assimilation in terms of “objective” representation of the target (for example, the self). For example, participants indicated first whether their own drug and alcohol consumption was more or less than that of Zappa or Graf (anchoring manipulation), then estimated their own monthly frequency of alcohol/drug consumption (“objective” judgment) and finally rated the “extensiveness” of their own consumption (“subjective” judgment). This manipulation demonstrated assimilation in objective judgment but contrast in subjective judgment. In other words, after comparing themselves with the high anchor Zappa, participants reported *more* consumption but rated their consumption as *less* extensive.

### Judgment as decision-making

While social judgment (in the sense I’ve used the term so far) was following the general trend of social cognition in incorporating increased attention to memory processes, another independent tradition was about to come into closer contact. The field of “judgment and decision-making” or, more broadly, “decision science” has always had a more applied focus than other areas described in this chapter. This focus has included analyzing decision problems and dilemmas, devising procedures for eliciting reliable preferences, and sometimes offering training programs in more effective decision-making. It thus has strong links to areas of management science and economics.

One of the main reasons why real-life decisions can be difficult is that the options we face typically involve multiple attributes or criteria. Maximizing one’s outcomes in terms of one criterion can therefore often involve accepting suboptimal outcomes in terms of another. Thus, the selection of an optimal solution to a multiattribute decision problem requires attention to how different attributes or criteria interrelate, and how important or valuable they are to the decision-maker. Hammond’s (1955; Hammond et al., 1975) social judgment theory provided a conceptual framework for this kind of work by directly applying Brunswik’s (1943) lens model of perception. (He was thus offering another take on the “judgment as perception” theme.) According to Brunswik, perception is organized through the incorporation of information from multiple cues. Hammond

argued that the same is true of evaluations and preferences, and so one of the main tasks of decision science is to identify the relative importance of specific cues (or criteria) and use these to predict preferences (typically through the use of multiple regression techniques). There are resonances here with our notion of “dimension salience” (van der Pligt & Eiser, 1984). However, although Hammond implies that one cause of decision conflict could be that different individuals assign different weights to different cues, he seems less concerned with trying to relate such differences to more general attitude processes.

The broader question of how individuals make choices between alternative decision options has received at least as much attention within economics as within psychology, since much of it provides one of the most powerful critiques of the “rational choice” approach to classical economics. In highly simplified form, this approach assumes that decision-makers compare the prospects of alternative actions in terms of two attributes: the benefits or costs of each possible outcome and the probability of each outcome. The product of the benefit (or cost) and probability then defines the “expected value” (EV) of each outcome, and it is assumed that this, *and this alone*, determines *preference*. In other words, a “rational” decision-maker will always prefer the option (or portfolio of options) with the most positive expected value. The fact that “rational” is a value-laden term leads to ambiguities about whether this theoretical approach constitutes a prescriptive guide to how decisions *ought* to be made, or whether it offers a descriptive account of how decisions *actually are* made. We do not need to worry too much here about its merits as a prescriptive guide. The empirical question is whether people’s actual preferences and decisions conform to the simple principle of “rational choice” thus defined, and the evidence is clear that they often do not. What is less clear is why any such departures from “rationality” occur, and whether they can be explained in terms of an alternative unified theory.

One of the earliest limitations to the notion of expected value was the recognition that “objective” value—for instance, amounts of money—is imperfectly (that is, nonlinearly) related to how much an outcome is actually *valued*, subjectively, by an individual (Bernoulli, 1738/1954). Suppose you won \$1,000 in a lottery. How happy would you feel? Probably pretty happy. Now suppose instead you won \$2,000. You would still feel happy, indeed possibly rather happier, but *twice* as happy? Maybe not. Now suppose instead that you had been told that your winnings were going to be \$100,000. You would no doubt be ecstatic. Then you learn that the exact sum had been recalculated as \$101,000. A bit better, for sure, but at that level the extra \$1,000 is neither here nor there. In other words, it is a common intuition that there is a kind of “law of diminishing returns” when it comes to estimating the desirability of, or pleasure received from, gains of different magnitude. This distinction between “objective” (e.g., monetary) and subjective value is expressed by using the term “utility” to refer to the latter. The relationship of utility to (objective) value in any given context may then be defined by an empirically derived “utility

function,” highly similar to the function identified by Fechner (1860) to relate perceptual to physical magnitudes.

The next limitation is that, in many instances, the actual probability of particular outcomes is not known, so decisions are guided instead by people’s subjective expectations. Hence, for most real-life decisions, the concept of expected value is replaced by that of “subjective expected utility” (SEU). The underlying assumption, however, is relatively unchanged: It is still assumed that individuals should prefer prospects, or choose courses of action, that maximize their net SEU in any given situation. This idea is deeply embedded in several familiar theoretical approaches within social psychology, most prominently Fishbein and Ajzen’s (1975) theory of reasoned action (TRA) and Ajzen and Fishbein’s (1980) theory of planned behavior (TPB). These and similar models assume that a main predictor of attitudes and hence (along with other comparable variables) behavioral intentions is the sum of an individual’s “evaluative beliefs” about an object or action. These “evaluative beliefs” are calculated as sums of products of subjective expectancies and utilities, just as in SEU theory. What is less easy to tie down in these theories is whether this calculation (of sum of products) is supposedly performed consciously as part of the “planning” or “reasoning” process, or whether individuals merely act *as if* such a calculation has been performed (somehow, somewhere). Naturally, this last question needs addressing if we are to know exactly what kinds of cognitive processes are being posited by TRA or TPB.

For the purposes of this chapter, however, the main relevance of the various critiques of the EV/SEU formulations of rational choice lies in (a) the fact that preferences can be influenced by how choices are presented or framed, even without any change in EV/SEU and (b) what is revealed about how individuals judge (and often misjudge) the likelihood of events. There are difficulties, however, in integrating much of this work into a historical account of social judgment in the broader sense, in that its main target lies elsewhere, in attempting to understand subjective probability estimates and preferences of the kind that can guide economic and similar behavior. This is far from a criticism (quite the contrary), but it does mean that it is a field with its own narrative, which in many respect proceeds parallel to, rather than in direct contact with, the rest of this chapter. If anything, the intellectual traffic has been in an opposite direction to that of the other themes considered here. Whereas social judgment has frequently drawn on insights from studies of perception, memory, etc., I do not see so many enduring signs of it drawing on studies of decision-making. Rather, decision theorists have drawn on insights from more general social judgment research and applied these within their own (sometimes very specific) paradigms. It is these that I shall now briefly describe.

### Framing

An axiom of classical EV theory is that preferences should reflect the *absolute* outcomes of choice alternatives and should be unaffected by how these outcomes are described or compared

to any standard. Kahneman and Tversky (1979, 1984) directly dispute this axiom. According to their prospect theory, outcomes are evaluated as gains or losses relative to a given reference point. This reference point could be the current state of affairs, an individual's expectations, or some standard defined by the experimenter. Judgments of outcomes or prospects, in other words, are *relative*. This in itself should come as no surprise to those familiar with the more general history of judgment research and the concept of frame of reference, but the strength of effects demonstrated as a consequence of quite simple manipulations can be dramatic.

For example, Tversky and Kahneman (1981) had participants imagine a choice between two preventive medicine programs to combat a serious epidemic, from which the expected death toll was 600. In one condition, the choice was between (A) a program that would “save 200” and (B) one that had a 1/3 probably of saving 600, but a 2/3 probability of saving nobody; in this condition 72% preferred A. In another condition, the same dilemma was presented as one between (C) where 400 would die and (D) where there was a 1/3 probability of nobody dying, but a 2/3 probability of 600 dying; in this condition, 78% preferred D. Note that the expected value of all these options is the same, but how they are judged is very different. Tversky and Kahneman express this by proposing that individuals tend to be “risk-averse” (i.e., prefer certain over uncertain outcomes) for gains, but “risk-seeking” for losses. However, what makes an outcome a gain or a loss is not its absolute value, but the standard against which it is compared. The “lives saved” frame implies a comparison with 600 deaths, whereas the “lives lost” frame implies a comparison with 0 deaths.

This concept of gain–loss framing has been extended to a variety of fields, including health communication (Rothman & Salovey, 1997). Evidence suggests that messages intended to encourage individuals to adopt behaviors to prevent the onset of disease (e.g., using sunscreen to avoid skin cancer; Detweiler, Bedell, Salovey, Pronin, & Rothman, 1999) are more effective when framed in terms of gains. However, messages focused on detection of a symptom or condition can be more effective if framed in terms of possible losses if the behavior is not followed (Rothman, Martino, Bedell, Detweiler, & Salovey, 1999). These findings suggest that thinking in terms of gains vs. losses can have important consequences for behavioral motivation, a theme taken up and further developed by Higgins (1998) in his regulatory focus theory.

## Heuristics

Another major challenge to the “rational choice” model of decision-making is the large body of evidence that individuals are generally rather poor at systematically processing statistical information. If we can't even use probabilities reliably, how can we be expected to make decisions based on multiplying these probabilities by the value of outcomes? The early work of Tversky and Kahneman (1973, 1974) demonstrated that people typically are vulnerable to several sources of error in their

estimates of frequency and probability, and proposed that this reflected an overreliance on simplified decision rules or *cognitive heuristics*. For example, the “availability heuristic” involves overestimating the probability of events that are more easily remembered or imagined, the “representativeness heuristic” reflects an overreliance on the typicality of attributes within a target category, whereas the “anchoring and adjustment heuristic” refers to the ease with which numerical estimates can be biased towards some arbitrary starting point offered by the experimenter. Taken as a whole, this research shows that many people have difficulty in making reliable judgments of probability or frequency, and tend to process information selectively while appearing to rely on simplified decision rules.

I say “appearing” to rely since it is less clear whether these “rules” constitute a typology of underlying cognitive processes, or more a descriptive typology of the effects themselves. Tversky and Kahneman tell us a great deal about *when* such effects occur. They also offer a general answer to the question of *why* they occur: heuristics offer a less effortful way of handling complex and uncertain information. However, the early demonstrations of these biases included less discussion about *how* they occur. Do heuristics reflect the operation of more general judgment processes under specific conditions, or do they point to new principles not previously discovered? There is good reason to favor the former interpretation. Indeed, slightly later work by Kahneman and Tversky (1979) and Kahneman and Miller (1986) constitute sophisticated applications of the principle of the relativity of judgments to subjective standards. “Anchoring and adjustment” effects—reflecting the sensitivity of judgments to numerical comparison standards provided by the experimenter—seem part and parcel of the broader story that question and response scale formats (including the ranges covered by scale categories) may suggest expectations about the responses the researcher anticipates from the participant (e.g., Schwarz et al., 1985). “Representativeness” likewise seems to reflect more general processes of categorization, as discussed earlier in this chapter, as well as holding possible implications for the broader literature on stereotype use.

Work on the “availability” heuristic stands slightly apart from that on other heuristics in that this did seem to have a direct influence on the development of more general social judgment research. In simple terms, this heuristic assumes that ease of recall (or imagination) shapes judgments of probability. The “common sense” basis for this heuristic is the assumption that events that actually occur more frequently will be easier to recall (in general terms), so that ease of recall is a reasonably valid cue to actual frequency. Tversky and Kahneman's (1973) paper came out just at the start of the shift in orientation of social judgment research away from the idea of “judgment as perception” toward a view of “judgment as memory”—or, more precisely, how judgment reflects the use of information from memory. The clearest link here is to the previously described work on priming and memory *accessibility*. The change in nomenclature reflects the argument that information may be “available” in memory (in the sense that it's there somewhere)



but not necessarily easy to retrieve or access. Since the “availability heuristic” relies on *ease* of recall, the term “accessibility” has become widely adopted as a more appropriate name for this construct.

Priming procedures were thus introduced as means of manipulating accessibility through associative memory. As we have seen, this work led quickly to considerations of the conditions under which primes produced assimilation or contrast, as well as moderating variables such as conscious awareness. For a while, the original concern with *ease* of recall (as distinct from memory *content*) became somewhat obscured. However, an important set of studies shows that *perceived* ease of recall can not only influence judgment, but itself be influenced by judgmental context and question format. Schwarz et al. (1991a) report three experiments in which participants rated their own assertiveness after being instructed to recall either 6 or 12 examples of when they themselves had behaved either assertively or nonassertively (i.e., a  $2 \times 2$  design). The reasoning behind this design is that, because it is more difficult to recall 12 examples (of anything) than 6 examples, individuals asked to recall examples of their own assertiveness would find it more difficult to think of themselves as assertive (i.e., rate themselves as *less* assertive) when trying to recall 12 examples than 6. Conversely, if asked to recall nonassertive behaviors, participants should rate themselves as *more* assertive when asked to recall more examples. This indeed was what was found (the effect being moderated by other manipulations designed to influence participants’ expectations and attributions for their success or failure at recalling such behaviors). Note that this result runs counter to a simple version of a priming (or accessibility) interpretation that would imply that self-ratings should follow the quantity of recalled information pointing in a given direction—e.g., that 12 examples of recalled assertiveness should lead to higher self-ratings of assertiveness than 6 examples.

Such findings have implications for self-reports of health status that similarly depend on recalling instances from memory. Rothman and Schwarz (1998) asked male participants to recall either eight or three examples of behaviors that either increased or decreased their own risk of heart disease (again, a  $2 \times 2$  design), after first reporting whether or not they had a family history of heart disease. The assumption was that those individuals with a family history of heart disease would be more motivated to think systematically about the issue and hence be less likely to base their ratings on perceived ease of recall. Those with such a family history rated themselves as less vulnerable, and needing to change their lifestyle, when they recalled more healthy and fewer unhealthy behaviors. In other words, their self-ratings suggest a kind of self-perception (Bem, 1972) process. In contrast, those without such a family history saw themselves as less vulnerable, or needing to change, when required to recall either more examples of risk-increasing or fewer examples of risk-decreasing behaviors. A related consideration is that recalling the *presence* of instances in memory may be easier, and/or have a greater influence on subsequent judgments, than recalling their *absence*. Fazio, Sherman, and

Herr (1982) term this the “feature-positive” effect. This again has implication for health self-reports. I asked student participants to rate their health status either by indicating which of a list of 30 symptoms they *had* experienced over either the past month or the past year, or by crossing out those symptoms they had *not* experienced (Eiser, 2000). Participants in the latter condition reported 70% more symptoms. However, after adjusting for the number of symptoms reported, participants in the former condition rated their own health more negatively.

### Decisions from experience vs. decisions from description

Despite the interest of social judgment researchers in memory processes and hence in the availability/accessibility heuristic in particular, there has been a continuing debate over the broader implications of the cognitive heuristics tradition. Does this work mainly show how clever Tversky, Kahneman and colleagues were in producing dramatic effects under experimental conditions, or does it mainly show how stupid most people are for much of the time? The latter alternative in particular has prompted more critical analysis of the implications of such experimental research. Nisbett and Ross (1980) quote the reaction of a colleague to their account of such findings: “If we’re so dumb, how come we made it to the moon?” It’s a fair question. One line of argument, offered by Hogarth (1981), is that yes, we rely on heuristics, but most of the time such heuristics are a pretty good guide to reality. They aren’t perfect, but they can mean that we get things right more often than not. This raises two important questions:

- 1 What is the function of heuristics and judgment processes? (Answer: to guide behavior rather than get full marks on statistics tests).
- 2 How closely do the experimental tasks used in this field of research correspond to the decision problems faced by people in everyday life? (Answer: only somewhat).

Perhaps because of the huge breadth and depth of Kahneman and Tversky’s contributions to the science of decision-making, this last point has received rather intermittent attention. Yet it is vital to consider if we are to assess the relevance of this part of their work to the more general history of social judgment. What is the dependent variable? Estimates of the absolute or relative likelihood or frequency of particular events. What is the main finding? These estimates often deviate from the answers that would be given if proper rules of statistical reasoning were followed. But how do we know that such estimates are incorrect statistically? *Because the instructions to participants typically provide statistical information that participants then fail to use appropriately.* Does this aspect of methodology matter? If the target of this program of research is to challenge the dominant “rational choice” view of decision-making, possibly not, since it shows that, even when people are given all the statistical information needed to make a “rational” choice, they are still

prone to error. If, however, the aim is to understand more generally how people predict and make sense of events that happen to them in real time, it can matter a very great deal.

To make this point in more concrete terms, let's return to Tversky and Kahneman's (1981) example of lives lost or gained as a result of adopting alternative prevention programs. The probabilities are provided to the (hypothetical) decision-makers *in advance*. But in real life, how would these ever be known? And what would constitute evidence for, for example, a program having a 1/3 probability of saving everyone and a 2/3 probability of saving nobody (and nothing in between)? How many previous instances of the use of such a program would be required before such probabilistic predictions could be offered with any reliability? None of this detracts from the robustness of the effects actually demonstrated, but this demonstration seems entirely contingent on the peculiar cultural context of a psychological experiment, in which participants take the information provided to them for granted, and do not ask whether such information is "really" true. But if we step outside this context, the question needs to be asked whether decision-makers often, or indeed ever, know the exact probabilities they are dealing with for sure, and if not, what do they do about it?

This concern may be expressed by saying that the large part of experimental work on cognitive heuristics and such like requires participants to make "decisions from description" whereas a less restricted paradigm would also allow the study of "decisions from experience." The latter occur in situations where individuals cannot know all relevant probabilities in advance, but have to work these out for themselves through sampling their environment and receiving feedback, over time, from the consequences of their own decisions and those of others. There is an increasing body of literature showing that individuals will often make very different choices under these contrasting conditions, even when "descriptions" convey accurately in advance the information that could eventually be gained experientially from complete sampling of the evidence available (Barron & Erev, 2003; Hertwig, Barron, Weber, & Erev, 2004). This appears to apply particularly to assessment of rare events that might not be directly experienced if sampling of potentially available data is incomplete (Rakow & Newell, 2010).

This raises the more general point that incomplete feedback can lead to incomplete or biased learning. For example, much excessive risk-taking might be partly attributable to the fact that real-life (negative) reinforcement schedules are often partial, inconsistent, and/or delayed. Not all instances of dangerous driving, imprudent investment, or even criminal activity lead certainly or immediately to injury, loss, or punishment. If, on the basis of a limited sample, perpetrators get away with it, they may underestimate the risk and even overestimate their own skills. Conversely, if feedback is contingent on sampling of the environment where such sampling is itself risky (e.g., animal foraging behavior), this may lead to excessive avoidance of objects or places that are in fact beneficial rather than dangerous, an effect demonstrated experimentally by Fazio, Eiser, and Shook (2004).

More important than the question of which paradigm can lead to which kinds of errors, however, is that of the processes that underlie these different kinds of judgment, irrespective of whether these judgments are "correct" from a prescriptive point of view. The "decisions-from-description" paradigm, on which much of the cognitive heuristics research has been based, essentially challenges participants' capacity for abstract reasoning with statistical concepts. By contrast, the "decisions-from-experience" paradigm appears much more representative of how we have to deal with information and uncertainty in real life, where there can often be a cost to delay or indecision. Judgment in such contexts hence becomes a dynamic process, evolving over time and shaped by learning.

### Judgment as learning

Once any research field starts to focus on apparent contradictions between the findings of different experiments, or on the limiting conditions controlling relatively subtle effects, there is a risk of losing sight of the broader picture. Social judgment research is grounded historically in fundamental questions concerning the workings of the human mind, but along the way such questions have become narrowed down. This has happened for the best of pragmatic reasons. However, as a consequence, a number of assumptions have become generally accepted, some of which have been less critically examined than others.

This much (at least) we can say with confidence:

- that the judgment of any target object is always *relative* to its context, by which is meant (a) any other objects—remembered, imagined or actually presented with which the target object may be compared, and (b) any other features of the stimulus environment that may influence participants' attention to particular attributes of the target object and/or their interpretation of the judgment task;
- that (especially in social judgment) such target objects are typically *multiattribute*, so that participants may attend more to some attributes than to others;
- that judgment requires *selective* attention to, and processing of, several features of the context and attributes of the target objects according to their perceived relevance;
- that often this need for selectivity is served by a form of *categorization*;
- and that the expression of judgment within experiments is a *communicative* act performed within the constraints of the language or response modality made available to participants.

But beyond this, several questions still remain. The priming literature in particular talks of retrieving concepts from memory, but what kind of process of retrieval is being assumed? One influential notion is the "storage bin" model proposed by Srull and Wyer (1979). According to this metaphor, constructs are stacked in separate "bins" in long-term memory. Constructs that have been

more frequently and/or recently used will tend to be found near the top of their relevant stack, and so will be more easily retrieved. By contrast, Higgins (1996; Higgins, Bargh, & Lombardi, 1985) has proposed that retrieval depends on the transfer of “excitation” between constructs according to their associations in memory, with accessibility depending on the number and strength of such associations. The metaphor here is more one of electronics or neural activity, and so is at least one step closer to a model about which it makes sense to start asking how such processes might be embodied in the human brain. Nonetheless, the model still deals with constructs at a highly symbolic level.

Learning particularly is another process that is often assumed, but rarely explicitly defined. It is therefore worth revisiting a few of the findings reviewed in this chapter and considering what kind of learning might be involved. If memory retrieval depends on associations, these associations will need to have been acquired at some stage. However, within the priming paradigm, experimenters typically rely on associations that are already “given” in conventional language use, or on nonlinguistic stimuli of strong affective value within the culture. Within more traditional judgment paradigms, the process of learning is at least more visible since participants typically are not informed about the range and distribution of stimuli they have to judge, but have to work this out for themselves. An early demonstration of such learning in operation comes from Parducci (1956; Parducci & Hohle, 1957), who showed that participants were more responsive (as evidenced by shifts in judgment) to an extension of the stimulus range than to a restriction of the range through the elimination of previously presented stimuli at the extremes.

### Associative learning and categorization

Associative learning is clearly implied in the early experiments involving superimposed classifications (e.g., Eiser, 1971; Tajfel & Wilkes, 1963), since participants were not told in advance which stimuli fell into which class. The categorization process seems to involve learning an association between individual stimuli (or stimulus magnitudes) and (the labels denoting) class membership. But here again interpretation of experimental effects relies to quite an extent on intuitions, rather than precise analysis, concerning the underlying cognitive process. The dominant idea, from Tajfel (1959), is that superimposed cues somehow make the task of differentiation “easier.” This, of course, only applies if the cues (e.g., class labels) are informative of the positions of the stimuli along the dimension of judgment, which is why a number of studies incorporate control conditions in which class membership is randomly associated with stimulus positions. One way of making judgment easier could be to simply respond to the class label, and ignore any individuating characteristics of the stimuli. In the context of interpersonal judgment, this would amount to stereotyping. This should clearly lead to a reduction of intraclass differences and hence to a perceived separation (accentuation of differences) for “midrange” stimuli at the borderline between the classes.

However, it is far less clear why a superimposed classification should lead (as is typically the case) to increased polarization, i.e., an accentuation of interclass differences. Perhaps easier judgments can be made with more confidence, and confidence translates itself into greater use of the judgment scale extremes. This is very plausible, but it invokes a process that has very little to do with associative learning or categorization *per se*, rather than conventions about how rating scales are used and interpreted. If categorization depends on learnt associations between individual stimuli and class labels, at some stage we have to define how the class labels are themselves to be mapped onto the scale of judgment. If, for instance, a class label functions merely as shorthand for the mean or central tendency of the stimuli belonging to its class, then reliance on the class label (to the relative neglect of individuating stimulus information) should lead to a reduction of intraclass differences but *not* an increase in interclass differences. To account for the latter effect, we would need to assume that class labels were taken to signify positions (prototypes, perhaps) that were more extreme than the central tendencies of the stimuli in their respective classes. Again, this is quite plausible, but even so we are still faced with the problem that we often find significant increases (Eiser, 1971) and occasionally decreases (Eiser et al., 1991) in interclass differences *without* reliable reductions in intraclass differences. So something else must be going on, over and above any direct association between class labels and individual stimulus attributes.

### The mind as a dynamical system

I asserted at the beginning of this chapter that the processes underlying social judgment phenomena are extremely general. I also drew attention to the philosophical axioms concerning the unity of the psychological and the physical that inspired Fechner’s psychophysics. But are these just fond statements of principle, or can they be cashed in to buy us a better understanding of judgmental effects? If we take as a starting point that the mind is a physical system, where might that lead us? One direction is to ask what can be said, at a high level of abstraction, about physical systems in general. Here one lesson that could be drawn is that we are looking not just at some organization of matter, but at what can also, or alternatively, be defined as an organization of energy. Physical systems, therefore, cannot be properly understood without consideration of their *dynamics*. What we take as objects (our own bodies, for instance) are more precisely states at a given point in time, all of which are liable to change over time, albeit with huge variability in the rate and predictability of such change. What is true of physical states may also be said of psychological ones, which of course are physical states too (Eiser, 1994, 1997; Port & van Gelder, 1995; Vallacher & Nowak, 1994; van Gelder, 1998). As expressed most radically by Hume (1740/1911, p. 239), even the impressions that give rise to a concept of self “are nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement.”

This is all a far cry from the idea of mental constructs being stacked up in “storage bins” (Srull & Wyer, 1979). Nonetheless, dynamic systems exhibit not only change but patterning and organization, and here a highly significant and hopeful conclusion can be drawn: that even highly complex patterns and organizations can be attributed to quite simple “rules” or algorithms, reapplied over several iterations. Such “emergent complexity” or “self-organization” can be witnessed throughout natural phenomena, such as cell division, how leaves grow on trees, and the coordinated movements of shoals of fish or flocks of birds. For us psychologists, the fundamental question of how thought is possible in a physical system thus translates itself into the perhaps somewhat less daunting question of how thought, consciousness, etc., might emerge from the interaction and iteration of brain activity at a simpler level.

Which takes us to the second lesson. Everything we have been considering is dependent on the workings of the human brain, so, in building our theories, *some* acknowledgment of how (we think) the brain works might not go amiss. This doesn’t mean at all that we need to wait for neuroscientists to provide us with suitable theories: They won’t. I’m not even sure what it would really mean theoretically if an imaging researcher had participants perform a social judgment task with their heads in a scanner, and found that different parts of the brain lit up when participants showed assimilation than when they showed contrast. No doubt somebody will do something like this very soon. Of course it is worth discovering what functions are primarily served by particular structures or projections in the brain, but the knowledge so gained may not directly answer the questions of most importance in our field. I am suggesting something much simpler than this: that we should seek to frame our theories in terms that have at least some plausibility at a physical level.

This may seem uncontroversial, but in fact it challenges the way a number of familiar viewpoints are currently formulated. Notions such as “spreading activation” and “excitation transfer” appeal to the idea of the brain as an interconnected network, but we are left asking what activation is spreading *between*. The answer offered is that activation spreads between concepts or constructs. The words we use matter here. It may be difficult to avoid metaphors entirely, but at least we should try not to mix them. It is fine to say that *thinking* of concept A makes it more likely that one will *think* of concept B. This is not a new idea. Hume again:

the true idea of the human mind is to consider it as a system of different perceptions or existences, which are linked together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other . . . One thought chases another, and draws after it a third, by which it is expelled in its turn.

(1740/1911, p. 247)

Remember, though, that Hume is talking about thoughts as *events* (“in perpetual flux”), so the kind of system he sees the mind as being is clearly a dynamic one. If we shift to the

terminology of brain activity, it is fine to say that activation can spread between different neurons, or larger regions of the brain. We now have the technology to measure such activation and observe it spreading. But none of this entitles us to say that activation spreads between *constructs*.

### Connectionism: A physical(ist) view of psychological processes

A possible way forward is offered by the conceptual framework known as *connectionism* (Ellis & Humphreys, 1999; McLeod, Plunkett, & Rolls, 1998). This is not (or not necessarily) a theory of specific forms of brain activity, but one that claims (or aspires to) “neural plausibility.” In other words, it is formulated in terms that *could* have meaning at a neural level. The starting point is that the brain consists of a vast number of interconnected neurons—“interconnected” meaning that (electrical) activation can be transmitted across the synapses between neurons. The dendritic (i.e., “tree-like”) structure of neurons means that they can receive and send activation from and to several other neurons simultaneously. This leads to the concept of a “neural network,” both as a way of describing actual configurations of neurons and (more often) as an analogue of such configurations at a more abstract level in computer simulations. Beyond these “architectural” features, the key assumption is that the *links between neurons in a network are strengthened by learning*. Stronger associations mean stronger connections between neurons, and since these are continually being updated through learning, the system they constitute is essentially a dynamic one. Much connectionist research is concerned with the capacities and limitations of different learning algorithms that control such changes in “connection weights,” but such details need not detain us here. The more relevant point is that the concepts and methods of connectionism have now been applied successfully to a wide range of not only cognitive, but social psychological, phenomena (Eiser, 1994; Read & Miller, 1998; Smith, 1996; Van Overwalle, 2007).

One reason for the appeal of connectionism is that it offers a parsimonious alternative to theories that are framed in terms of vaguer linguistic concepts and/or rely on assumptions about symbolic processing. Connectionism may, for instance, be particularly appropriate for the study of more automatic processes. However, its relevance is not confined to such phenomena. Social judgment, right from the beginning, has had problems conceptualizing the distinction between subjective representations of objects and the act of rating them on a response scale. Despite the field’s antidualist origins in Fechner’s psychophysics, several researchers slip into a way of talking, and perhaps thinking, about judgment that implies some kind of inner subjective (Cartesian) theatre where mental representations, or memories, of objects are paraded, observed and compared. This implies that when judges make a rating, what they are rating is not the object itself, but their mental representation of it. But if this mental representation involves comparison with other mental representations, we are left

having to account for how this comparison is itself represented mentally, so we are trapped in an infinite regress.

According to the connectionist perspective, however, there is no such inner theater. Concepts, representations, and memories are embodied in distributed patterns of connections between units (or neurons) in the network. All the information acquired through learning is stored in the connection weights, which are constantly being updated through new learning. Thus there is no distinction between memory and cognitive processing in connectionist systems (McLeod et al., 1998). In other words, connectionist systems have no need of a distinct repository for stored memories. It is the pattern of connection weights themselves that controls what earlier theorists (e.g., Bower, 1981) called “spreading activation” and constrains perceptual and behavioral responses to particular inputs.

Within judgment tasks, one such behavioral response or output will be the choice of response category on the rating scale. Like all outputs of a connectionist network, this choice will depend on the satisfaction of multiple constraints, that is on the presence of several kinds of information and contextual features: the presence of superimposed cues, comparison standards, multiple stimulus attributes, and frequency counts of previous rating choices. Although class labels may come to be associated with other attributes of the target stimuli, we do not need to assume that these, and only these, associations are what determines the strength of any polarization effects. *Any* feature of the stimulus environment (input pattern) can be directly or indirectly connected to, and hence pass activation to, *any* kind of output. Some such connections may be “hard-wired” neuro-anatomically, but many will depend on experience and learning. Likewise, some may involve conscious representation of the target object being judged, but many will not.

### Conclusion: Judgment as construction

For well over half a century, social judgment has been one of the main conduits for the interchange of theories and methods between social and general cognitive psychology. Although I use the term “interchange” deliberately to imply that our field is in no way secondary, the fact remains that most of the trade in ideas has been asymmetrical. Social judgment theorists have been more content than many other social psychologists to appeal directly to concepts and models drawn from more general psychological research on processes of perception, memory and learning. The history of social judgment is, to a great extent, characterized by shifts in emphasis between these processes.

Social judgment, nonetheless, has some important distinctive features. The objects of judgment with which we are concerned are rich in meaning, derived from social relationships, interactions and values, and frequently represented in language. At the very least, social objects are typically far more complex than those used in more narrowly defined laboratory research on cognitive processes. This complexity poses extra challenges, both for theorists and for the judges themselves. Judges are

presented with multidimensional objects, but have to express their impressions of these objects in terms of far fewer dimensions, and often only one. Many of the experimental effects described in this chapter reflect the need for judges to be selective in the information, or dimensions, to which they give greatest priority and attention, as well as their sensitivity to covariation between subsets of information or dimensions. In short, to make sense of their task, judges *construct* patterns, or configurations, out of the complex information presented to them.

In the preface to their collection *The Construction of Social Judgment*, Martin and Tesser (1992) expressed the conviction that social judgment was at a turning point. Their optimism was not misplaced, but perhaps has yet to be wholly fulfilled. We build our theories with words, but the same words can sometimes suggest varied, and even contrary, ideas. In everyday speech, “construction” can refer both to the *process* of building something out of its constituent parts, and to a finished *product* or “structure.” We tend to think of structures as inert and relatively permanent. Judgments assuredly are not. The products of the judgment process are choices, decisions, and (within most of the experimental paradigms described) more mundanely, the selection of a response on a rating scale. Such products are temporary and transient, precisely because they are context-dependent. The focus therefore needs to be on construction as a *process*. In much the same way, Lichtenstein and Slovic (2006) have more recently used the title *The Construction of Preference* to alert decision researchers (and economists) to the ways in which expressions of choice and preference, even involving monetary value, are extremely sensitive to context and method of measurement.

The centrality of this constructive process is underlined by the universal theme of all judgment research: Judgments are relative to context. I have spent much of this chapter charting the history of debates concerning whether this relativity manifests itself in contrast or assimilation effects, or an increase or decrease in judged differences between classes. These are not “mere” details, but they are details, nonetheless, compared with the overall message that context always matters. It matters not just because it provides comparison standards or “anchors” that become incorporated into judges’ configurations of the judgmental frame of reference. It matters because this context is always changing. Just as we cannot step into the same river twice, we never can repeat the same stimulus experience exactly. Every new experience affects our learning history, and this learning history sets the context for our subsequent experiences. Helson (1964) suggested that this process could be thought of as one of weighted averaging. This now seems too limited to account for changes that involve not just the adoption of new comparison standards, but also the development of new associations. But Helson recognized, perhaps more than many others, the essentially dynamic nature of judgment in context. The core aim of social judgment theory is still to make sense of these dynamics and of how we handle complexity in our engagement with the objects that constitute our social world.

## References

- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice Hall.
- Anderson, N. H. (1962). Application of an additive model to impression formation. *Science*, *138*, 817–818.
- Anderson, N. H. (1967). Averaging model analysis of set size effect in impression formation. *Journal of Experimental Psychology*, *75*, 158–165.
- Anderson, N. H. (1981). *Foundations of information integration theory*. New York, NY: Academic Press.
- Bargh, J. A., & Pietromonaco, P. (1992). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology*, *43*, 437–449.
- Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, *16*, 215–233.
- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1–62). New York, NY: Academic Press.
- Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk (L. Sommer, trans.). *Econometrica*, *22*, 23–36. (Reprinted from *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, *1738*, *5*, 175–192.)
- Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*, 129–148.
- Brown, D. R. (1953). Stimulus-similarity and the anchoring of subjective scales. *American Journal of Psychology*, *66*, 199–214.
- Bruner, J. S., & Goodman, C. C. (1947). Value and need as organizing factors in perception. *Journal of Abnormal and Social Psychology*, *42*, 33–44.
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York, NY: Wiley.
- Bruner, J. S., & Postman, L. (1948). Symbolic value as an organizing factor in perception. *Journal of Social Psychology*, *27*, 203–208.
- Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychological Review*, *50*, 255–272.
- Campbell, D. T. (1956). Enhancement of contrast as composite habit. *Journal of Abnormal and Social Psychology*, *53*, 350–355.
- Churchland, P. C. (1986). *Neurophilosophy: Toward a unified science of the mind/brain*. Cambridge, MA: MIT Press.
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason and the human brain*. New York, NY: Putnam.
- DeCoster, J., & Claypool, H. M. (2004). A meta-analysis of priming effects in impression formation supporting a general model of information biases. *Personality and Social Psychology Review*, *8*, 2–27.
- Dennett, D. C. (1991). *Consciousness explained*. New York, NY: Little, Brown.
- Detweiler, J. B., Bedell, B. T., Salovey, P., Pronin, E., & Rothman, A. J. (1999). Message framing and sunscreen use: Gain-framed messages motivate beach-goers. *Health Psychology*, *18*, 189–196.
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL: Harcourt Brace Jovanovich.
- Eiser, J. R. (1971). Enhancement of contrast in the absolute judgment of attitude statements. *Journal of Personality and Social Psychology*, *17*, 1–10.
- Eiser, J. R. (1973). Judgement of attitude statements as a function of judges' attitudes and the judgemental dimension. *British Journal of Social and Clinical Psychology*, *12*, 231–240.
- Eiser, J. R. (1990). *Social judgment*. Buckingham, UK: Open University Press.
- Eiser, J. R. (1994). *Attitudes, chaos and the connectionist mind*. Oxford, UK: Blackwell.
- Eiser, J. R. (1997). Attitudes as attractors: More than a metaphor? *Psychological Inquiry*, *8*, 119–124.
- Eiser, J. R. (2000). The influence of question framing on symptom report and perceived health status. *Psychology and Health*, *15*, 13–20.
- Eiser, J. R., & Hoepfner, F. (1991). Accidents, disease, and the greenhouse effect: Effects of response categories on estimates of risk. *Basic and Applied Social Psychology*, *12*, 195–210.
- Eiser, J. R., Martijn, C., & van Schie, E. (1991). Categorization and interclass assimilation in social judgement. *European Journal of Social Psychology*, *21*, 493–505.
- Eiser, J. R., & Mower White, C. J. (1974). Evaluative consistency and social judgment. *Journal of Personality and Social Psychology*, *30*, 349–359.
- Eiser, J. R., & Mower White, C. J. (1975). Categorization and congruity in attitudinal judgment. *Journal of Personality and Social Psychology*, *31*, 769–775.
- Eiser, J. R., & Pancer, S. M. (1979). Attitudinal effects of the use of evaluatively biased language. *European Journal of Social Psychology*, *9*, 39–47.
- Eiser, J. R., & Ross, M. A. (1977). Partisan language, immediacy and attitude change. *European Journal of Social Psychology*, *7*, 477–489.
- Eiser, J. R., Spears, R., & Webley, P. (1989). Nuclear attitudes before and after Chernobyl: Change and judgment. *Journal of Applied Social Psychology*, *19*, 689–700.
- Eiser, J. R., & Stroebe, W. (1972). *Categorization and social judgement*. London: Academic Press.
- Eiser, J. R., & van der Pligt, J. (1979). Beliefs and values in the nuclear debate. *Journal of Applied Social Psychology*, *9*, 524–536.
- Eiser, J. R., & van der Pligt, J. (1982). Accentuation and perspective in attitudinal judgment. *Journal of Personality and Social Psychology*, *42*, 224–238.
- Ellis, R., & Humphreys, G. (Eds.). (1999). *Connectionist psychology: A text with readings*. Hove, UK: Psychology Press.
- Eriksen, C. W. & Hake, H. W. (1957). Anchor effects in absolute judgments. *Journal of Experimental Psychology*, *53*, 132–138.
- Fazio, R. H., Eiser, J. R., & Shook, N. J. (2004). Attitude formation through exploration: Valence asymmetries. *Journal of Personality and Social Psychology*, *87*, 293–311.
- Fazio, R. H., Sherman, S. J., & Herr, P. M. (1982). The feature-positive effect in the self-perception process: Does not doing matter as much as doing? *Journal of Personality and Social Psychology*, *42*, 404–411.
- Fechner, G. T. (1860). *Elemente der Psychophysik*. Leipzig, Germany: Breitkopf & Hartel.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- Grice, P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41–58). New York, NY: Academic Press.
- Hammond, K. R. (1955). Probabilistic functioning and the clinical method. *Psychological Review*, *62*, 255–262.
- Hammond, K. R., Stewart, T. R., Brehmer, B., & Steinmann, D. O. (1975). Social judgment theory. In M. F. Kaplan & S. Schwartz (Eds.), *Human judgment and decision processes* (pp. 271–312). New York, NY: Academic Press.

- Helson, H. (1947). Adaptation-level as frame of reference for prediction of psychophysical data. *American Journal of Psychology*, *60*, 1–29.
- Helson, H. (1964). *Adaptation-level theory*. New York, NY: Harper & Row.
- Herr, P. M. (1986). Consequences of priming: Judgment and behavior. *Journal of Personality and Social Psychology*, *51*, 1106–1115.
- Herr, P. M., Sherman, S. J., & Fazio, R. H. (1983). On the consequences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology*, *19*, 323–340.
- Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*, 534–539.
- Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York, NY: Guilford Press.
- Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, (Vol. 30, pp. 1–46). San Diego, CA: Academic Press.
- Higgins, E. T., Bargh, J., & Lombardi, W. (1985). The nature of priming effects on categorization. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *11*, 59–69.
- Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, *13*, 141–154.
- Hinckley, E. D. (1932). The influence of individual opinion on construction of an attitude scale. *Journal of Social Psychology*, *3*, 283–296.
- Hogarth, R. M. (1981). Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics. *Psychological Bulletin*, *90*, 197–217.
- Hovland, C. I., & Sherif, M. (1952). Judgmental phenomena and scales of attitude measurement: Item displacement in Thurstone scales. *Journal of Abnormal and Social Psychology*, *47*, 822–832.
- Hull, C. L. (1952). *A behavior system* (Vol. 1). New Haven, CT: Yale University Press.
- Hume, D. (1740/1911). *A treatise of human nature*. London, UK: Dent.
- Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*, 136–153.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.
- Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*, 341–350.
- Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, *50*, 537–567.
- Lichtenstein, S., & Slovic, P. (Eds.). (2006). *The construction of preference*. Cambridge, UK: Cambridge University Press.
- Manis, M. (1960). The interpretation of opinion statements as a function of recipient attitude. *Journal of Abnormal and Social Psychology*, *60*, 340–344.
- Manis, M. (1961). The interpretation of opinion statements as a function of message ambiguity and recipient attitude. *Journal of Abnormal and Social Psychology*, *63*, 78–81.
- Manis, M., Nelson, T. E., & Shedler, J. (1988). Stereotypes and social judgment: Extremity, assimilation and contrast. *Journal of Personality and Social Psychology*, *55*, 28–36.
- Manis, M., Paskewitz, J., & Cotler, S. (1986). Stereotypes and social judgment. *Journal of Personality and Social Psychology*, *50*, 461–473.
- Martin, L. L. (1986). Set/reset: Use and disuse of concepts in impression formation. *Journal of Personality and Social Psychology*, *51*, 493–504.
- Martin, L. L., & Tesser, A. (Eds.). (1992). *The construction of social judgments*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- McLeod, P., Plunkett, K., & Rolls, E. T. (1998). *Introduction to connectionist modelling of cognitive processes*. Oxford: Oxford University Press.
- Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, *110*, 472–489.
- Mussweiler, T., & Strack, F. (2000). The “relative self”: Informational and judgmental consequences of comparative self-evaluation. *Journal of Personality and Social Psychology*, *79*, 23–38.
- Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice Hall.
- Parducci, A. (1956). Incidental learning of stimulus frequencies in the establishment of judgment scales. *Journal of Experimental Psychology*, *52*, 112–118.
- Parducci, A. (1963). Range-frequency compromise in judgment. *Psychological Monographs*, *77* (2, Whole No. 565).
- Parducci, A. (1984). Value judgments: Towards a relational theory of happiness. In J. R. Eiser (Ed.), *Attitudinal judgment* (pp. 3–21). New York, NY: Springer.
- Parducci, A., & Hohle, R. (1957). Restriction of range in the judgment of single stimuli. *American Journal of Psychology*, *70*, 272–275.
- Port, R. F., & van Gelder, T. (Eds.). (1995). *Mind as motion: Explorations in the dynamics of cognition*. Cambridge, MA: MIT Press.
- Rakow, T., & Newell, B. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making*, *23*, 1–14.
- Read, S. J., & Miller, L. C. (Eds.). (1998). *Connectionist models of social reasoning and social behavior*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Romer, D. (1983). Effects of own attitude on polarization of judgment. *Journal of Personality and Social Psychology*, *44*, 273–284.
- Rothman, A. J., Martino, S. C., Bedell, B. T., Detweiler, J. B., & Salovey, P. (1999). The systematic use of gain- and loss-frame messages on interest in and use of different types of health behavior. *Personality and Social Psychology Bulletin*, *25*, 1355–1369.
- Rothman, A. J., & Salovey, P. (1997). Shaping perceptions to motivate healthy behavior: The role of message framing. *Psychological Bulletin*, *121*, 3–19.
- Rothman, A. J., & Schwarz, N. (1998). Constructing perceptions of vulnerability: Personal relevance and the use of experiential information in health judgments. *Personality and Social Psychology Bulletin*, *24*, 1053–1064.
- Schwarz, N. (1994). Judgment in a social context: Biases, shortcomings, and the logic of conversation, In M. Zanna (Ed.) *Advances in experimental social psychology* (Vol. 26, pp. 123–162). New York, NY: Academic Press.
- Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217–245). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991a). Ease of retrieval as information: Another

- look at the availability heuristic. *Journal of Personality and Social Psychology*, 61, 195–202.
- Schwarz, N., Groves, R. M., & Shuman, H. (1998). Survey methods. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 143–179). New York, NY: Oxford University Press.
- Schwarz, N., Hippler, H. J., Deutsch, B., & Strack, F. (1985). Response scales: Effects of category range on reported behavior and subsequent judgments. *Public Opinion Quarterly*, 49, 388–395.
- Schwarz, N., & Strack, F. (1991). Evaluating one's life: A judgment model of subjective well-being. In F. Strack, M. Argyle & N. Schwarz (Eds.), *Subjective well-being: An interdisciplinary perspective* (pp. 27–47). Oxford, UK: Pergamon.
- Schwarz, N., Strack, F., & Mai, H. P. (1991b). Assimilation and contrast effects in part-whole question sequences: A conversational logic analysis. *Public Opinion Quarterly*, 55, 3–23.
- Selltiz, C., Edrich, H., & Cook, S. W. (1965). Ratings of favorableness about a social group as an indication of attitude toward the group. *Journal of Personality and Social Psychology*, 2, 408–415.
- Sherif, M., & Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication and attitude change*. New Haven, CT: Yale University Press.
- Sherif, M., Taub, D., & Hovland, C. I. (1958). Assimilation and contrast effects of anchoring stimuli on judgment. *Journal of Experimental Psychology*, 55, 150–155.
- Smith, E. R. (1996). What do connectionism and social psychology offer each other? *Journal of Personality and Social Psychology*, 70, 893–912.
- Srull, K. T., & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology*, 37, 1660–1672.
- Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, 64, 153–181.
- Stevens, S. S. (1966). *Psychophysics: Introduction to its perceptual, neural and social prospects*. New York, NY: Wiley.
- Strack, F., Martin, L. L., & Schwarz, N. (1988). Priming and communication: The social determinants of information use in judgments of life-satisfaction. *European Journal of Social Psychology*, 18, 429–442.
- Tagiuri, R., & Petruccio, L. (Eds.) (1958). *Person perception and interpersonal behavior*. Stanford, CA: Stanford University Press.
- Tajfel, H. (1957). Value and the perceptual judgment of magnitude. *Psychological Review*, 64, 192–204.
- Tajfel, H. (1959). Quantitative judgement in social perception. *British Journal of Psychology*, 50, 16–29.
- Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues*, 25, 75–97.
- Tajfel, H. (Ed.). (1982). *Social identity and intergroup relations*. Cambridge, UK: Cambridge University Press.
- Tajfel, H., Sheikh, A. A., & Gardner, R. C. (1964). Content of stereotypes and the inference of similarity between members of stereotyped groups. *Acta Psychologica*, 22, 191–201.
- Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgment. *British Journal of Psychology*, 54, 101–114.
- Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, 33, 529–554.
- Thurstone, L. L., & Chave, E. J. (1929). *The measurement of attitude*. Chicago, IL: University of Chicago Press.
- Tresselt, M. E. (1948). The effect of experience of contrasted groups upon the formation of a new scale. *Journal of Social Psychology*, 27, 209–216.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207–232.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453–458.
- Upshaw, H. S. (1962). Own attitude as an anchor in equal-appearing intervals. *Journal of Abnormal and Social Psychology*, 64, 85–96.
- Vallacher, R. R., & Nowak, A. (Eds.). (1994). *Dynamical systems in social psychology* (pp. 197–218). San Diego, CA: Academic Press.
- van der Pligt, J., & Eiser, J. R. (1984). Dimensional salience, judgment and attitudes. In J. R. Eiser (Ed.). *Attitudinal judgment* (pp. 161–177). New York, NY: Springer-Verlag.
- van der Pligt, J., van der Linden, J., & Ester, P. (1982). Attitudes to nuclear energy: Beliefs, values and false consensus. *Journal of Environmental Psychology*, 2, 221–231.
- van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, 21, 615–665.
- Van Overwalle, F. (2007). *Social connectionism: A reader and handbook for simulations*. New York, NY: Psychology Press.
- Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgment: The role of naïve theories in corrections for perceived bias. *Journal of Personality and Social Psychology*, 68, 36–51.
- Zavalloni, M., & Cook, S. W. (1965). Influence of judges' attitudes on ratings of favorableness of statements about a social group. *Journal of Personality and Social Psychology*, 1, 43–54.