CHAPTER 3

ENCODING, REPRESENTING, AND ESTIMATING EVENT FREQUENCIES: A MULTIPLE STRATEGY PERSPECTIVE

NORMAN R. BROWN

Abstract

People use several formats to represent frequency information and they use several strategies to generate frequency judgements. These observations are at the core of an approach to event frequency called the Multiple Strategy Perspective *(MSP). This chapter provides an overview of the MSP and a review of pertinent research methods and findings. The overview catalogues the common frequency formats, estimation strategies, and performance patterns. It also identifies a set of systemic relations which first links encoding factors and event properties to frequency representations, then links these frequency representations to strategy selection, and finally links strategy selection to estimation performance (i.e. speed, accuracy, and bias). Strategy selection and implication for basic and applied research are considered in the concluding sections of the chapter.*

Introduction

How many sexual partners have you had in your life time?
During the past month, how often have you shopped at a department store?
How many times did the word METAL appear on the list that you just studied?

Questions like the first two are asked on surveys by sociologists, economists, and market researchers who hope to obtain an accurate information about the frequency of potentially important, but undocumented, behaviours; questions like the third are put to experimental participants by psychologists interested in understanding how repeated experiences affect memory and performance. Of course, in content, these questions could hardly be more different. Nonetheless, the processes used to answer them and the knowledge that these processes operate on have much in common. Sometimes people retrieve and count relevant instances; sometimes they recall stored tallies, rates or impressions; and sometimes they gauge frequency by assessing some aspect of memory performance or the memory trace. In other words, regardless of whether the to-be-estimated items are arbitrary words presented on a list or meaningful personal events, people represent frequency information in a number of different ways, and they use a variety of estimation strategies to generate frequency judgements. Moreover, event properties and encoding factors are systemically related to the way that frequency information is represented. Furthermore, the representation of frequency information and the strategies used to judge event frequencies are related in a systematic manner, as are strategy selection and estimation performance.

An approach called the *Multiple Strategy Perspective* (*MSP*) has been developed to summarize the regularities linking encoding, representation, process, and performance and to catalogue common frequency formats and estimation strategies. This chapter provides an overview of this approach and a selective review of pertinent research methods and findings. The main premises of the MSP are as follows:

(A) Information about event frequency can be represented in several different ways;

(B) People use multiple strategies to estimate event frequency;

(C) Encoding factors and event properties influence the way frequency is represented;

(D) The frequency-relevant contents of memory restrict strategy selection;

(E) Motivational factors and response context also affect strategy selection;

(F) Strategy selection affects estimation speed and accuracy;

(G) In turn, patterns of performance reflect strategy choice.

The sequence of stages implied by premises C, D, and F is sketched in Fig. 3.1. In this figure, the arrows connecting the stages are dashed to express the idea that there is variability, and at times redundancy, in the way that people encode frequency relevant information and that there also is certain amount of flexibility in the way that people respond when they are confronted with multiple sources of frequency-relevant information. In other words, the MSP is not strictly deterministic. Rather, it is assumed that a given set of encoding conditions may foster some types of frequency-relevant representations and not others, and that as a result some strategies will be applicable, but not others. The implication here is that a single set of frequency judgements may include estimates produced by several different estimation processes.

These premises, with appropriate substitution of terms, can be or have already been applied to a wide variety of estimation and problem-solving tasks (e.g. Siegler 1987). This is in keeping with the notion that human cognition is both flexible and opportunistic,

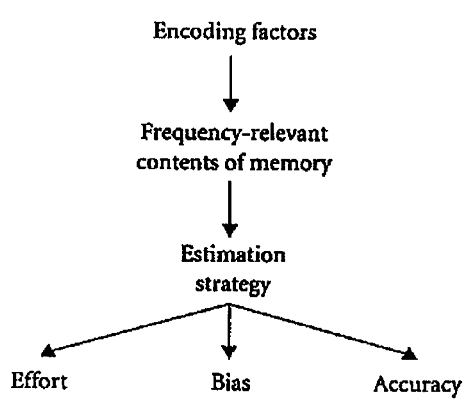

Figure 3.1 Overview of the core elements of the Multiple Strategy Perspective (MSP).

but that it is also constrained by the nature of accessible task-relevant information. What links this general view to the present topic is an empirically grounded catalogue of estimation strategies and a growing understanding of the ways that encoding, representation, process, and performance are related in this domain. Figs. 3.2 and 3.3 summarize this information. Fig. 3.2 presents a process-based taxonomy of the commonly observed strategies, and Fig. 3.3 lists several of these strategies along with the encoding factors, representational formats, and performance patterns that accompany their use.

The material presented in these summary illustrations is discussed at length in the next sections. Before turning to this material, some background is necessary. There are two related points; both concern the nature of the research that led to the development of the MSP, and the differences between this line of research and conventional laboratory-based work on event frequency. The first point is that MSP studies have often dealt with real-world events rather than word lists. These events were emphasized because many of the researchers interested in this topic were specifically concerned with understanding how survey respondents answer *behavioural frequencies questions*. These questions require the

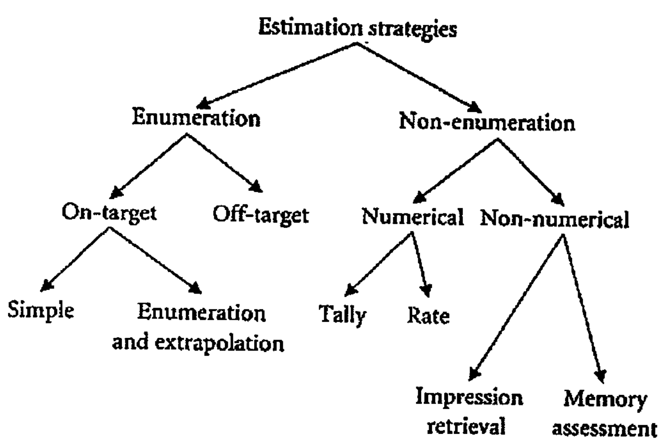

Figure 3.2 A taxonomy of common frequency estimation strategies.

Encoding	Content	Strategy	Performance
'Memorable' events	'On-target' instances	On-target enumeration	▪ RT ↑ freq ▪ underestimation
Regularity	Rate	Rate retrieval	▪ Fast, flat RT ▪ heaping
Intention and time	Tally	Tally retrieval	▪ Fast, flat RT ▪ accurate(?)
Frequent presentation	Vague quantifiers	Impression retrieval	▪ Fast, flat RT ▪ overestimation
Similar instances	Coherent traces	Memory assessment	▪ Fast, flat RT ▪ overestimation
Encoding/test mismatch	'Off-target' instances	Off-target enumeration	▪ SLOW, flat RT ▪ regressive estimates

Figure 3.3 Summary of the relation between encoding factors, frequency-relevant memory contents, and performance characteristics associated with several common frequency estimation strategies.

respondent to indicate how often he or she has engaged in a particular activity during a given reference period. For example, a consumer survey might ask respondents how often they have purchased laundry detergent in the past month, or a health survey might ask about the number of doctor's visits in the past year. The immediate goals here were to determine whether people could provide accurate unbiased answers to such questions and, assuming performance was less than perfect, to identify event properties and task conditions that affect estimation accuracy.

The second feature that differentiates event frequency studies that have contributed to the development of the MSP from those that have not is a fairly obvious one. In addition to measures that reflect the magnitude and accuracy of the obtained frequency judgements, all MSP studies have included at least one process-sensitive measure. Several different methods have been used. These include on-line measures such as response times (RTs) and concurrent verbal reports, and off-line measures such as immediate retrospective reports, effort ratings, and strategy menus. Although each method has its limitations, it is common for MSP studies to employ more than one measure and for these measures to converge on a conclusion.

With few exceptions (Hintzman & Gold 1983; Hintzman *et al.* 1981; Hockley 1984; Marx 1985; Voss *et al.* 1975), process-sensitive measures have not appeared in the mainstream frequency memory literature. This is not to say that process and representation have not been an issue. Indeed, memory researchers have long speculated that people may represent frequency information in several different ways and that they may use several different estimation processes to generate frequency judgements (Bruce *et al.* 1991; Hasher & Zacks 1984; Hintzman 1976; Howell 1973; Johnson *et al.* 1979; Watkins & LeCompte 1991). However, in retrospect, it seems that the almost exclusive reliance on accuracy measures, coupled with the relative simplicity of their test materials, made it difficult for these researchers to profit from their insights. In contrast, the catholic approach to methodology adopted by *Cognitive Aspects of Survey Methodology* (CASM) researchers and the complex nature of the real-world events that were of interest to the survey community established the conditions that made it possible to identify several estimation strategies and to associate strategy selection with event proprieties. Much of the research reviewed below demonstrates that the findings reported in the behavioural frequency literature could be corroborated and extended using experimental methods, and that the two approaches could be coordinated in a way that combined rigour of the laboratory with the insights gleaned from observing real-world performance.

A taxonomy of estimation strategies

Enumeration-based estimates

The main division in taxonomy presented in Fig. 3.2 is between strategies that involve enumeration and those that do not. An estimate is based on enumeration when individual items or events are retrieved and counted, and the count arrived at by this retrieval process serves as the basis for a frequency judgement (Barsalou & Ross 1986; Begg *et al.* 1986; Blair & Burton 1987; Brown 1995, 1997; Bruce *et al.* 1991; Burton & Blair 1991; Conrad & Brown 1994; Conrad *et al.* 1998; Conrad *et al.* 2001; Greene 1989;

Williams & Durso 1986). Enumeration is considered to be *on-target* when (almost) all of the retrieved items are members of the to-be-estimated event category; when a high percentage of retrieved items do not belong to the target category, enumeration is considered to be *off-target* (Conrad *et al.* 2001). In principle, the distinction between on- and off-target enumeration could be a fuzzy one. In practice, this is not a problem, at least in the laboratory; depending on the task, enumeration protocols and corresponding cued-recall tests indicate that intrusion rates are either very high, which means that enumeration is off-target, or very low which means it is on-target; there simply is no middle ground (see below).

There are two types of on-target enumeration: *simple enumeration* and *enumeration and extrapolation*. An enumerated response is assigned to the former category when the number of items retrieved during the estimation process and the magnitude of the frequency judgement that follows are equal, and to the latter category when estimated frequency is greater than the number of retrieved instances. Other things being equal, simple enumeration becomes less common and enumeration and extrapolation more common as presentation frequency increases (Brown 1995, 2001). Interestingly, participants do not extrapolate when they are off-target.

On-target enumeration

On-target enumeration has been observed in both laboratory and real-world settings. It is characterized by two distinctive features. First, the time required to decide on a specific numerical response increases with event frequency; second, people tended to underestimate event frequency, and this tendency is more pronounced for frequently presented items than for less frequently presented items. Fig. 3.4 provides examples of these tendencies. The data plotted in this figure are drawn from Brown (1995). In this study, participants were

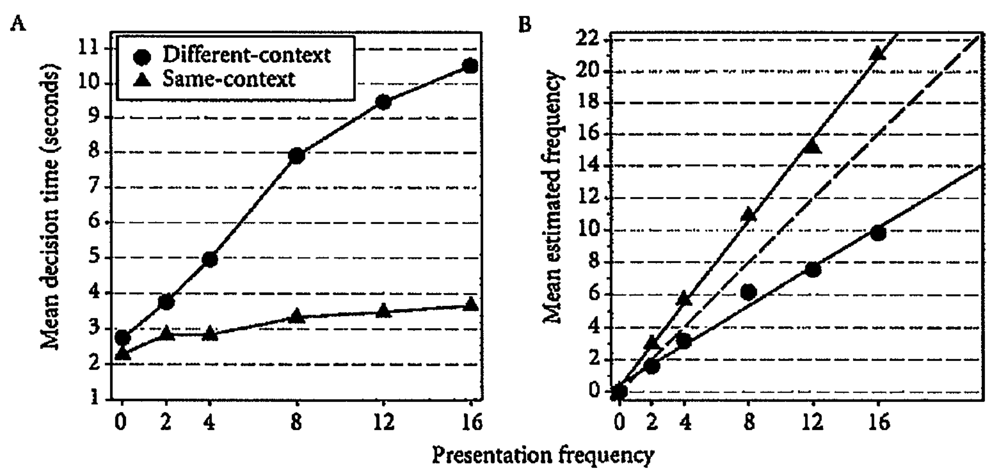

Figure 3.4 Mean response times (Panel A) and mean estimated frequencies (Panel B) and for the different-context group (solid circle) and the same-context group (empty circle). In Panel B, the solid lines represent the best linear fit for the means, and the dashed line represents the actual frequencies. These data are drawn from Brown (1995), Experiment 2.

presented with word pairs consisting of a target word (a category label, e.g. CITY) and a context word (a category exemplar, e.g. London). One group studied a *different-context* list and the other a *same-context* list. On the different-context list, target words were paired with unique context words on each presentation (e.g. CITY—Boston, CITY—Cleveland, CITY—London, etc.), and on the same-context list, a target word were always paired with same context word (e.g. CITY—London, CITY—London, CITY—London, etc.), and on both lists, presentation frequencies for the target words ranged from 2 to 16. Following the study list, target words (and non-presented category labels) were presented one at a time, and participants either thought aloud as they generated their estimates (Experiment 1) or were timed as they silently judged the list frequency of each (Experiment 2). The verbal protocols collected in Experiment 1 indicated that different-context participants frequently enumerated (57%) and the same-context participants never enumerated. The decision time data collected in Experiment 2 indicated that RT increased sharply with presentation frequency, but only in the different-context condition. In both experiments, participants who studied the different-context list (i.e. participants who relied on enumeration) tended to underestimate event frequencies and those who studied the same-context lists overestimated them.

This pattern of performance has been replicated many times (Brown 1995, 1997, 2001; Conrad & Brown 1994; Conrad *et al.* 2001) and is easy to understand. RT increases with presentation frequency because enumeration is a serial process. Assuming that the time to retrieve an instance is constant (or increases with each additional instance, Bousfield & Sedgewick 1944; Indow & Togano 1970) and that participants retrieve more instances when responding to high frequency items than to low frequency items, it follows that it should take more time to respond to the former than to the latter. Underestimation occurs because relevant instances may be permanently forgotten, because output interference causes some instances to become temporally inaccessible, and because retrieval efforts are sometimes terminated before all relevant instances have been recalled. Concurrent verbal protocols and immediate and delayed retrospective reports provide converging support for this interpretation. When participants are tested under conditions that produce underestimation and steep RT functions, inevitably, these process measures include information indicating that estimates were generated by retrieving instances one at a time and counting them up (Brown 1995, 2001; Conrad & Brown 1994; Conrad *et al.* 1998, 2001; Marx 1985).

An obvious precondition for enumeration is the availability in memory of relevant, readily accessible event traces. This implies that conditions that foster the encoding, storage, and retrieval of these traces should promote the use of enumeration-based strategies. Consistent with this claim several behavioural frequency studies have found a positive association between the rated distinctiveness of repeated instances of an event class and the likelihood that respondents enumerate when they judge the frequency of that class (Conrad *et al.* 1998; Menon 1993). Similarly, as noted above, laboratory studies demonstrate that enumeration is common when repeated events are unique, but completely absent when the repeated events were identical (Brown 1995). The usual explanation of this finding is that highly similar event instances are readily assimilated into a pre-existing scheme, and that as a consequence, access to the individual event traces is lost, whereas dissimilar instances maintain their individuality in memory (Menon 1993).

Subsequent research has demonstrated that distinctiveness is a necessary, but not a sufficient condition for enumeration. In a series of experiments, Brown (1997) presented participants with lists comprised of target-context pairs. The target words were repeated across the study list; however each context word appeared only once making each pair unique, and hence distinctive. After studying the lists, participants estimated the presentation frequency of the target words and then recalled context words.

Study time (6 sec per pair versus 2 sec per pair) and target-context relatedness (category labels paired with either category exemplars or unrelated words) were manipulated between experiments. Study-phase instructions were manipulated within experiment; participants in one group were instructed to commit the context words to memory; participants in a second group were instructed to focus on the frequency of the target words; and participants in a third group were instructed to study the word pairs for a memory test, but were not informed of the nature of the test.

This set of manipulations produced large differences in memory for the context words and large differences in the steepness of the functions that related RT to presentation frequency. More importantly, there was a strong systematic relation between context memory and response time; when context memory was best, estimation time increased sharply with presentation frequency, and the steepness of this estimation time-presentation frequency function decreased with context memory. More concretely, when participants were informed of the upcoming cued-recall test and were given 6 sec per pair to study the related context lists, RT increased from 3 sec to 11 sec as presentation frequency increased from 2 to 16; in a subsequent cued recall test, participants in this group recalled 46% of the context words. In contrast, RT increased from 3 sec to 4 sec across the same range when participants studied the unrelated-pair list under comparable encoding conditions. Cued recall for this group was 15%.

These results indicate that enumeration was common when instance memory was good and that the use of this strategy declined as relevant memory traces became increasing difficult to recall. More generally, this study demonstrated that factors which increase the memorability of event instances increase the use of enumeration and these factors encompass both encoding conditions (i.e. study time and study phase instructions) and event properties (i.e. similarity of event instances and memorability of differentiating contexts). Several other factors, unrelated to event memory, are also known to play a role in determining whether people enumerate. The influence of these factors on strategy selection is taken up in the last section of this chapter.

Off-target enumeration

Both off-target enumeration and on-target enumeration engage a serial retrieval process and produce frequency judgements that are based on the outcome of this process. Despite these similarities the two strategies are used under very different conditions and produce very different patterns of performances. A recent study by Conrad illustrates these points (Conrad *et al.* 2001). In this study, participants in one group, the *instance-plus-property* group, studied 109 word pairs consisting of a property label and a related noun (e.g. chocolate-BROWN, garbage-SMELLY, etc.). These pairs were constructed from published property norms (Battig & Montague 1969; McEvoy & Nelson 1982; Underwood & Richardson 1956). As in the different-context conditions described above, property labels were presented multiple times across the list, but each noun was presented only once. Participants in a second group, the *instance-only* group, studied the same set of nouns, but in the absence of the property labels. Following the study phase, all participants were presented with the property labels and were timed as they estimated how many of the studied nouns possessed the target property.

Participants in the two groups responded very differently during the frequency judgement task (see Fig. 3.5). In the instance-plus-property condition, RT increased with presentation frequency, frequency judgements were quite accurate and high frequency properties tended to be underestimated. For reasons outlined above, these results indicate that participants in this condition made heavy use of standard on-target enumeration-based estimation strategies. In contrast, in the instance-only condition, RTs were very slow and the speed of responding was unrelated to the presentation frequency. Moreover, these effortful responses were highly inaccurate; judged frequencies for property-bearing instances were uncorrelated with the normative property frequencies and participants tended to overestimate the frequency uncommon property nouns and underestimate the frequency of the common property nouns. Other researchers have looked at the property frequency task and have also found that that the instance-only condition yields very poor frequency discrimination (Barsalou & Ross 1986; Brooks 1985; Freund & Hasher 1989).

A protocol study (Conrad *et al.* 2001, Experiment 1) shed light on the origins of this 'worst of both worlds' performance observed in the instance-only condition. In this experiment, participants who had studied the instance-only list thought aloud as they estimated the frequency of the target properties. An analysis of their verbal reports indicated that these participants retrieved and counted list items on 80% of the trials and that the use of enumeration was unrelated to the presentation frequency. However, it was more common for the protocols to include more list items that were normatively *incorrect* than ones that were normatively correct; on average, 2 out of 3 enumerated

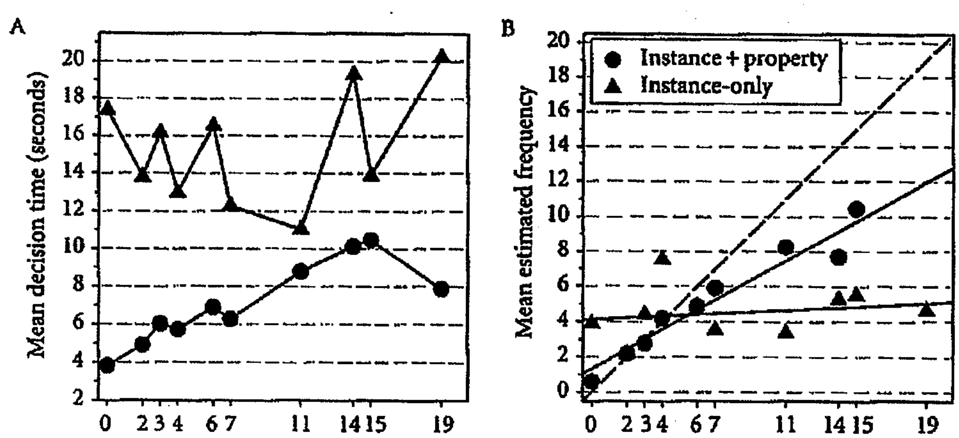

Figure 3.5 Mean response times (Panel A) and mean estimated frequencies (Panel B) and for the instance-plus-property group (solid circle) and the instance-only group (empty circles). In Panel B, the solid lines represent the best linear fit for the means, and the dashed line represents the actual frequencies. These data are drawn from Conrad *et al.* (2001).

items fell into the former category. This is why the application of enumeration under these circumstances is termed 'off-target'. Apparently, in this situation people recall instances in a more or less haphazard manner and use a liberal criterion to determine whether a retrieved item should be accepted as a member of the target category.

The use of on-target enumeration in the instance-plus-property condition and off-target enumeration in the instance-only condition results also rules out the possibility that it is the nature of the frequency judgement task (i.e. taxonomic frequency vs. property frequency) that determines whether enumeration is on- or off-target. Instead, it appears that people rely on off-target enumeration when there is a mismatch between the way that retrievable items are categorized at encoding and the way memory is subsequently probed. This mismatch can be avoided by indicating the appropriate category at the time of study or by probing memory in a way that captures the organization that people naturally impose on diverse instances. This suggests that it might be possible to use a frequency judgement task, supplemented with the appropriate processes measures, to determine whether or not a proposed organizational scheme is natural, i.e. employed without prompting at encoding to classify instances and at retrieval to access them. Natural schemes should support on-target enumeration, whereas people will have to rely on off-target enumeration when they are asked to respond to 'unnatural' schemes.

One final point. This line of research indicates that people do not naturally encode events by their properties, and suggests that under normal circumstances the assessment of property frequency is a futile endeavour. In frequentist models of judgement and decision making property frequencies are equivalent to base rates (Dougherty *et al.* 1999; Gigerenzer & Hoffrage 1995). Given that people are very poor at assessing property frequencies, it is not surprising that they usually neglect base rates when generating conditional likelihood estimates. On this view, base rate is neglected at least in part because it cannot be assessed with any certainty.

Nonenumeration strategies

In Fig. 3.2, the nonenumeration strategies are divided into two groups. The numerical strategies, *retrieved rate* and *retrieved tally*, are grouped together because, in both cases, prestored *numerical* information is retrieved and used as the basis for a frequency judgement. Access to prestored nonnumerical frequency information also plays a key role in the application of *impression retrieval* strategy. However, this strategy is grouped with memory assessment because both share a common response process, one that converts a frequency judgement from a relative or qualitative format to a numerical one and because it is difficult if not impossible to distinguish them using process measures (Conrad & Brown 1996).

Numerical strategies

Obviously, a prerequisite for the application of the retrieval strategies is access to facts that specify frequency information, and the absence of such facts precludes their use. Beyond this commonality, these strategies display substantial differences. The simplest of the retrieval strategies is retrieved tally. People who have stored and can recall a tally are able to respond to a frequency question by stating its value. Assuming that potentially relevant events are classified correctly and that a relevant counter has been updated with the occurrence of each new instance of the target category, these responses should be quite accurate.

Brown and Sinclair (1999) found that tally-based responses are fairly common when people are asked to estimate how many sexual partners they have had in their life time. Furthermore, Brown (1997) has used RT data to argue that people who were informed of an upcoming frequency test prior to exposure to a study list sometimes counted presentation frequencies and used these counts as the basis for subsequent frequency judgements (see below). Interestingly, the use of retrieved tallies has not been reported in other studies that have looked at behavioural or laboratory frequency. Taken together, these findings suggest that counts can be used to represent event frequencies, but that this is the expectation rather than the rule. Apparently, people sometimes count events, but only when the events are particularly important, interesting, or noteworthy. In addition, research participants may attempt to track event frequencies, but only when they are instructed to do so. Otherwise, it can be assumed that people do not keep count of most activities and hence are typically unable to use a retrieved-tally strategy when asked to estimate event frequencies.

Many behavioural frequency studies have documented the use of rate-based strategies (Burton & Blair 1991; Conrad *et al.* 1998; Means & Loftus 1991; Menon 1993). Typically, this strategy is used when a respondent is asked about a common activity that occurs on a regular basis. It makes sense that the regularity of an activity is associated with knowledge of its rate of occurrence. When the period that defines the rate and the period specified by the question are the same, the retrieved rate can be stated as an estimate. However, when there is a mismatch between the two periods, respondents must multiply (or divide) the rate by an appropriate value to arrive at a response.

It is this latter process that produces the spiked or 'clumped' pattern in the distribution of the estimated frequencies that is characteristic of rate-based responding. A study by Conrad *et al.* (1998) provides an example of this tendency. In this study, respondents in a telephone survey were asked to estimate the number of times they had participated in several common activities (shopped for groceries, purchased gasoline, etc.) during the past month. In addition, they were required to provide a retrospective strategy report following each estimate. These reports made it possible to construct the *response distributions* presented in Fig. 3.6. These distributions display the percentage of frequency judgements produced by a given strategy that were assigned a given value (also see Brown, Buchanan & Cabeza 2000; Hintzman & Curran 1995). As Fig. 3.6 (Panel A) indicates, the most frequent response made by people using a rate-based strategy was 4 (once a week). The only other common rate-based estimates were 2 (once every two weeks), 8 (twice a week), and 30 (every day). This pattern was in sharp contrast to ones produced from enumerated estimates and general impression estimates (i.e. estimates produced by a memory assessment strategy or qualitative retrieval); the former (Fig. 3.6, Panel B) displayed an exponential decrease across the low end of the range, and the latter (Fig. 3.6, Panel C) displayed a strong clumping of responses at 'prototypical' values (i.e. 10, 15, 20).

Nonnumerical strategies

As mentioned above, two processes are also required to produce a frequency judgement when a participant uses a qualitative retrieval strategy. The first is a memory retrieval

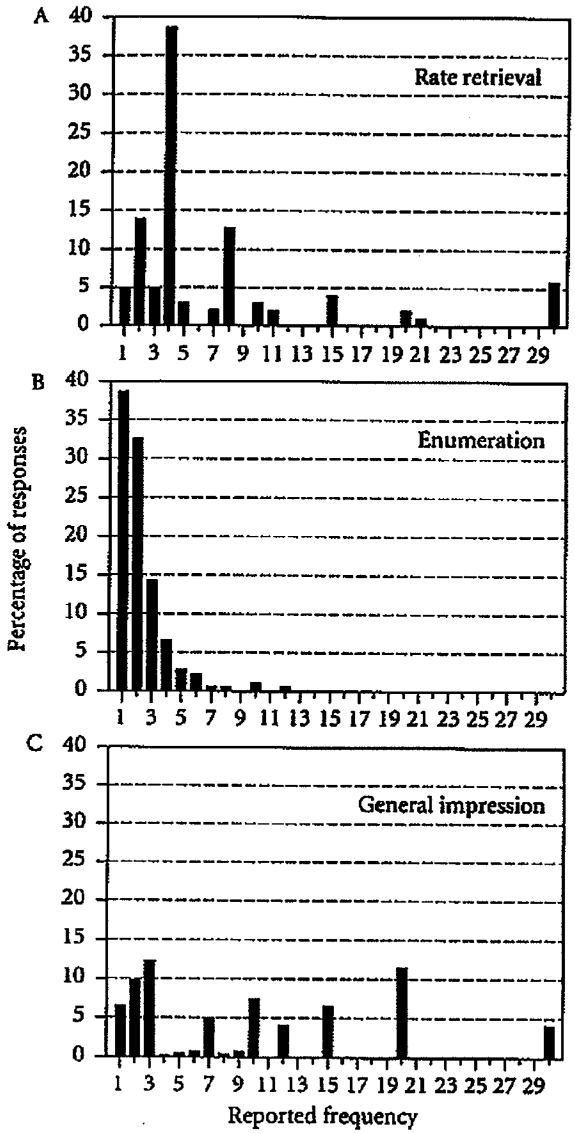

Figure 3.6 Distribution of frequency estimates as a function of estimation strategy. Panel A plots rate-based estimates, Panel B plots enumeration-based estimated strategy, and Panel C plots estimates that included general impression terms. These data are drawn from Conrad *et al.* (1998).

process that succeeds in accessing a fact or impression from memory that expresses frequency relevant information in a nonnumerical manner (e.g. 'The word CITY showed up *many times* on the preceding list' or 'I have *rarely* been cross-country skiing'). The second process involves selecting a numerical value to represent the quantity expressed by the retrieved fact.

One reason for proposing that people sometimes encode and later retrieve this type of nonnumerical information is the finding that vague quantifiers (terms like *a lot*, *many*, *very few*, etc.) are often included in verbal reports and that the use of these terms increases dramatically with presentation frequency (Brown 1995, Experiment 1). This is what one would expect if the repetition of an item or event increases peoples' awareness that it has been repeated and the likelihood that a fact conveying this awareness is stored in memory. In other words, in the same way that a seasoned academic eventually comes to *know* that he or she has attended many conferences, participants in a memory experiment, particularly one that repeats some items many times, eventually come to *know* that they have seen some of the words many times.

In the absence of verbal reports, it is difficult to distinguish between this direct retrieval strategy and the use of memory assessment. In part this is because both strategies yield fast flat RT functions of the sort observed when participants estimate the frequency of similar or identical instances (e.g. Fig. 3.4, Panel A, same-context condition). In addition, it appears that both strategies engage the same conversion process and that this process produces a characteristic pattern of performance, regardless of whether the input is a vague quantity retrieved from memory or an intuition about relative frequency based on an evaluation of some aspect of memory performance. In both cases, this process must map information about relative frequency on to a numerical response range. When people rely on this process, they often select round numbers as responses (Fig. 3.6, Panel C; see also Brown & Sinclair 1999; Conrad *et al.* 1998; Hintzman & Curran 1995) and they tend to overestimate event frequencies, particularly large ones (Brown 1995; Conrad & Brown 1994).

Overestimation is most pronounced when the upper bound of the response range is unspecified, but it can be eliminated or reversed by providing participants with information that restricts the range or by presenting them with a conservative set of response options. In contrast, enumeration-based and rate-based estimates are immune to range manipulations (Brown 1995, Experiment 3; Menon *et al.* 1995, Experiment 1). These findings suggest that overestimation and conversion are related because people must define a response range without the benefit of explicit numerical information and because the range can be grossly overestimated, but not grossly underestimated. This is not a problem when people base their estimates on information that is inherently numerical (i.e. when they enumerate or retrieve a tally or rate) because these processes do not engage a range-sensitive conversion process.

It should be acknowledged that the MSP does not, at present, specify the exact nature of the memory assessment process used to generate a qualitative evaluation of event frequency. However, it is assumed that each encoding of a repeated event has a monotonic effect on the contents of memory, that access or retrieval is influenced by these frequency-related changes in memory, and that people are sensitive to processing differences brought about by these frequency-related changes. This conceptualization leaves open the possibility that frequency is indexed by retrieval fluency (i.e. availability; Tversky & Kahneman 1973; Betsch & Pohl in this volume), trace strength (Hintzman 1969; Morton 1968), the similarity between a memory probe and the contents of episodic memory (Hintzman 1988; Jones & Heit 1993; Nosofsky 1988), associative strength (Sedlmeier in this volume) or a nonsymbolic mental magnitude (Gallistel this volume; Whalen, Gallistel & Gelman 1999). The difficulty in selecting between these possibilities parallels the more general difficulties encountered in distinguishing impression retrieval and memory assessment. This is another situation where process measures are likely to be uninformative. In particular, regardless of the nature of measure assessment process, the RT function should be fast and flat, and the conversion process should produce similar pattern of estimates regardless of origins of the frequency information.

Strategy selection

Strategy selection, one of the important issues raised by the MSP that has not yet been thoroughly investigated, is the last topic addressed in this chapter. As pointed out above, strategy selection is *restricted* by frequency relevant contents of memory; enumeration is impossible in the absence of retrievable event instances; direct retrieval strategies are impossible in the absence of prestored frequency information; and memory assessment strategies are ineffective when the probe fails to access a coherent representation. What is less obvious is how people select strategies when they have access to multiple sources of potentially relevant information.

One possibility is that people always select the least effortful of the available strategies. If this is position is correct, then it should be difficult to find evidence that people use enumeration, the most effortful of strategies, when they have access to other sources of frequency relevant information. However, the protocols reported for the 1995 article (Brown 1995, Experiment 1) provide two findings that are inconsistent with this prediction. First, enumeration and extrapolation was the most common strategy used in the different context condition, accounting for the 29% of all responses. Of course, in order to extrapolate, one needs to believe that the number of items retrieved is less than the number of items presented. It is possible that the knowledge used to justify extrapolation could also serve as the basis for a nonenumerated response. If so, the frequent use of this strategy is inconsistent with the notion that enumeration is used only as a last resort.

The second finding that argues against the least-effort hypothesis is the occasional appearance of general impression statements (i.e., vague quantifiers) in enumerated responses. For example, in the protocol study, at the highest presentation frequency, 13% of the enumerated estimates included one of these statements. If one assumes that general impression statements are made when a qualitative fact is retrieved from memory or when the product of a memory assessment process is verbalized, then the existence of these dual-strategy responses indicates that people sometimes enumerate even when they have accessed information that could have served as the basis for a less effortful estimate, one based on direct retrieval or memory assessment.

If the least-effort position fails to explain strategy selection, perhaps it can be explained by an enumeration bias. It could be that accessing retrievable traces triggers enumeration and that once enumeration has begun it continues at least until the first chunk or two have been recovered and unpacked. If enumeration is compulsive when retrievable traces are available, then only those factors that affect the availability of these traces should affect the use of the enumeration-based strategies. Counter to this prediction, the application of enumeration-based strategies is influenced by a number of factors that do not necessarily influence the accessibility of enumerable traces. These include: (a) actual frequency—beyond a certain level, enumeration decreases as actual frequency increases. (Note the drop in RT at presentation frequency 19 in Fig. 3.5, Panel A; also see, Bruce & van Pelt 1989; Burton & Blair 1991; Means & Loftus 1991); (b) response deadline—enumeration is less common when response time is restricted than when it is not (Burton & Blair 1991); (c) time frame—longer reference periods produce less enumeration than shorter reference periods (Blair & Burton 1987; Burton & Blair 1991);

(d) the presence of retrievable tallies—holding level of recall constant, the presence of a retrievable tally decreases peoples reliance on enumeration (Brown 1997).

There are good reasons then for rejecting both the least-effort hypothesis and the enumeration-bias hypothesis, though these arguments still leave the strategy selection issue unresolved. The findings outlined above, however, are consistent with the notion that people have a degree of control over strategy selection and that they attempt to balance perceived effort and perceived accuracy when conditions allow for strategy choice (Payne *et al.* 1993; Russo & Dosher 1983). The assumptions here are that people realize that enumeration is demanding and time consuming; that they prefer concrete information (in the form of instance counts or directly coded frequency facts) to the vague information (in the form of intuitions supplied by memory-assessment processes); and that they are willing to work quite hard to accumulate the former provided they have enough time to do so. As a result, people who have access to enumerable traces may forgo enumeration when the response time is restricted or when they believe that a large number of instances would have to be retrieved to provide a fair indication to the to-be-estimated quantity. Because this belief would have to be based on another source of frequency relevant information, i.e., a prestored or computed impressions of frequency, people who decide not to enumerate are still in the position to execute an alternative estimation strategy, albeit that one requires the conversion of a qualitative impression to a numerical value. The preference for retrieved tallies over enumeration reported in Brown (1997) indicates that when two sources of information are equally credible (enumerated counts and directly coded frequency facts), people prefer the more convenient one. In summary, these data suggest that perceived accuracy and convenience are evaluated when an estimation strategy is selected; that accuracy is given more weight when one source of information is clearly more credible; and that convenience is weighted heavily when competing sources are considered equally credible or when the more credible strategy is deemed to be too demanding.

Conclusion

People use different formats to represent event frequency and they use several different strategies to generate frequency judgements. These observations, which are at the core of the MSP, present both challenges and opportunities to researchers in the area. One challenge is to establish an exhaustive catalogue of relevant frequency formats, estimation processes, along with a specification of their prerequisites, conditions of use, and consequences. A second is to develop tasks that isolate estimation processes and methods for relating judgement to process on a case-by-case basis. A third challenge is to develop or modify models of memory and judgement that incorporate or accommodate the central tenets of the MSP.

I believe that the research reviewed above indicates that progress has been made in meeting the first two challenges, but only the first two. Nonetheless, it should be possible to profitably extend the MSP in several directions. One way to extend this approach is to use it to understand (or re-evaluate) problematic findings in the frequency literature (e.g. the diverse impact of encoding contexts and processes on judgements of frequency; cf. Begg *et al.* 1986; Hintzman & Stern 1978; Jonides & Naveh-Benjamin 1987; Rose 1980; Rowe 1973). The MSP should also be useful to memory psychologists interested in investigating a *single* representation or strategy. The contribution here would be to alert such researchers to the possibility that a single set of frequency judgements may contain estimates that are produced by different strategies and based on different frequency representations. At the same time, if researchers are aware of this pitfall and are willing to take advantage of the available processes-sensitive measures, it should be possible to study a given strategy or format in isolation. Finally, this general approach may provide guidance in designing behavioural frequency questions and interpreting biased or inconsistent responses to such questions (Brown & Sinclair 1999).

In summary, there is now a good deal of evidence indicating that people represent and estimate event frequency in variety of ways and that encoding, representation and performance are linked in a systemic manner. It is at least possible that the MSP will be extended to address a variety of basic and applied issues in memory. If nothing else, this line of research demonstrates the benefits of coordinating real-world and laboratory research efforts and the crucial role that complex decision processes play in memory performance.

Acknowledgement

This project has been supported by the author's NSERC operating grant. I would like to thank Fred Conrad for innumerable discussions of the topics covered in this chapter. These discussions have contributed greatly to the development of the MSP. I would also like to thank Peter Lee for his comments on this chapter.

References

Barsalou, L. W. & Ross, B. H. (1986). The role of automatic and strategic processing in the sensitivity of superordinate and property frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*:116–134.

Battig, W. P. & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph, 80* (3, Part 2): 1–46.

Begg, I., Maxwell, D., Mitterer, J. O. & Harris, G. (1986). Estimates of frequency: Attribute or attribution? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*:496–508.

Blair, E. & Burton, S. (1987). Cognitive processes used by survey respondents to answer behavioral frequency questions. *Journal of Consumer Research, 14*:280–288.

Bousfield, W. A. & Sedgewick, C. H. (1944). An analysis of sequences of restricted associative responses. *Journal of General Psychology, 30*:149–165.

Brooks, J. E. (1985). Judgements of category frequency. *American Journal of Psychology*, 98: 363–72.

Brown, N. R. (1995). Estimation strategies and the judgement of event frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*:1539–1553.

Brown, N. R. (1997). Context memory and the selection of frequency estimation strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*:898–914.

Brown, N. R. (2001). Exemplar generation and the use of enumeration-based frequency estimation strategies. Manuscript in preparation.

Brown, N. R., Buchanan, L. & Cabeza, R. (2000). Estimating the frequency of nonevents: The role of recollection failure in false recognition. *Psychonomic Bulletin and Review, 7*:684–691.

Brown, N. R. & Sinclair, R. C. (1999). Estimating number of lifetime sexual partners: Men and women do it differently. *Journal of Sex Research, 36*:292–297.

Bruce, D., Hockley, W. E. & Craik, F. I. M. (1991). Availability and category-frequency estimation. *Memory of Cognition, 19*:301–312.

Bruce, D. & van Pelt, M. (1989). Memories of a bicycle tour. *Applied Cognitive Psychology, 3*:137–156.

Burton, S. & Blair, E. (1991). Task conditions, response formulation processes and response accuracy for behavioral frequency questions in surveys. *Public Opinion Quarterly, 55*:50–79.

Conrad, F. G. & Brown, N. R. (1994). Strategies for estimating category frequency: Effects of abstractness and distinctiveness. *American Statistical Association, Proceedings of the Section on Survey Methods Research* (pp. 1345–1350). Alexandria, VA: American Statistical Association.

Conrad, F. G. & Brown, N. R. (1996). Estimating frequency: A multiple strategy perspective. In D. Herrmann, M. Johnson, C. McEvoy, C. Hertzog and P. Hertel (eds) *Basic and Applied Memory: Research on Practical Aspects of Memory*, vol. 2 (pp. 167–178). Hillsdale, NJ: Erlbaum.

Conrad, F. G., Brown, N. R. & Dashen, M. (2001). Estimating the frequency of events from unnatural categories. Manuscript submitted for publication.

Conrad, F. G., Brown, N. R. & Cashman, E. (1998). Strategies for answering behavioral frequency questions. *Memory, 6*:339–366.

Dougherty, R. P., Gettys, C. F. & Ogden, E. E. (1999). MINERVA-DC: A memory process model of judgements of likelihood. *Psychological Review, 106*:180–209.

Freund, J. S. & Hasher, L. (1989). Judgements of category size: Now you have them, now you don't. *American Journal of Psychology, 102*:333–353.

Greene, R. L. (1989). On the relationship between categorical frequency estimation and cued recall. *Memory and Cognition, 17*:235–239.

Hasher, L. & Zacks, R. T. (1984). Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist, 39*:1372–1388.

Gigerenzer, G. & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*:506–528.

Hasher, L. & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*:356–388.

Hintzman, D. L. (1969). Apparent frequency as a function of frequency and the spacing of information. *Journal of Experimental Psychology, 80*:139–145.

Hintzman, D. L. (1976). Repetition and memory. In G. H. Bower (ed.) *The Psychology of Learning and Motivation*, vol. 10 (pp. 47–91). New York: Academic Press.

Hintzman, D. L. (1988). Judgements of frequency and recognition memory in a multi-trace memory model. *Psychological Review, 95*:528–551.

Hintzman, D. L. & Curran, T. (1995).When encoding fails: Instructions, feedback, and registration without learning. *Memory and Cognition, 23*:213–226.

Hintzman, D. L. & Gold, E. (1983). A congruity effect in the discrimination of presentation frequency: Some data and a model. *Bulletin of the Psychonomic Society, 21*:11–14.

Hintzman, D. L., Grandy, C. A. & Gold, E. (1981). Memory for frequency: A comparison of two multiple-trace theories. *Journal of Experimental Psychology: Human Learning and Memory, 7*:231–240.

Hintzman, D. L. & Stern, L. D. (1978). Contextual variability and memory for frequency. *Journal of Experimental Psychology: Human Learning and Memory, 4*:439–549.

Hockley, W. E. (1984). Retrieval of item frequency information in a continuous memory task. *Memory and Cognition, 12*:229–242.

Howell, W. C. (1973). Representation of frequency in memory. *Psychological Bulletin, 80*:44–53.

Indow, T. & Togano, K. (1970). On retrieving sequence from long-term memory. *Psychological Review, 77*:317–331.

Johnson, M. K., Raye, C. L., Wang, A. Y. & Taylor, T. H. (1979). Fact and fantasy: The role of accuracy and variability in confusing imaginations with perceptual experiences. *Journal of Experimental Psychology: Human Learning and Memory, 5*:220–240.

Jones, C. M. & Heit, E. (1993). An evaluation of the total similarity principle: Effects of similarity on frequency judgements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*:799–812.

Jonides, J. & Naveh-Benjamin, M. (1987). Estimating frequency of occurrence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*:230–240.

McEvoy, C. L. & Nelson, D. L. (1982). Category names and instance norms for 106 categories of various sizes. *American Journal of Psychology, 95*:581–634.

Marx, M. H. (1985). Retrospective reports on frequency judgements. *Bulletin of the Psychonomic Society, 23*:309–310.

Means, B. & Loftus, E. F. (1991). When personal history repeats itself: Decomposing memories for recurring events. *Applied Cognitive Psychology, 5*:297–318.

Menon, G. (1993). The effects of accessibility of information in memory on judgements of behavioral frequencies. *Journal of Consumer Research, 20*:431–440.

Menon, G., Raghubir, P. & Schwarz, N. (1995). Behavioral frequency judgements: An accessibility-diagnosticity framework. *Journal of Consumer Research, 22*:212–228.

Morton, J. (1968). Repeated items and decay in memory. *Psychonomic Science, 10*:219–220.

Nosofsky, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*:54–65.

Payne, J. W., Bettman, J. R. & Johnson, E. J. (1993). *The adaptive decision maker.* New York: Cambridge University Press.

Rose, R. J. (1980). Encoding variability, levels of processing, and the effects of spacing of repetitions upon judgements of frequency. *Memory and Cognition, 8*:84–93.

Rowe, E. J. (1973). Frequency judgements and recognition of homonyms. *Journal of Verbal Learning and Verbal Behavior, 12*:440–447.

Russo, J. E. & Dosher, B. A. (1983). Strategies for multiattribute binary choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*:676–696.

Siegler, R. S. (1987). The perils of averaging data over strategies: An example children's addition. *Journal of Experimental Psychology: General, 116*:250–264

Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 4*:207–232.

Underwood, B. J. & Richardson, J. (1956). Some verbal materials for the study of concept formation. *Psychological Bulletin, 53*:84–95.

Voss, J. F., Vereb, C. & Bisanz, G. (1975). Stimulus frequency judgements and latency of stimulus frequency judgements as a function of constant and variable response conditions. *Journal of Experimental Psychology: Human Learning and Memory, 3*:337–350.

Watkins, M. J. & LeCompte, D. C. (1991). Inadequacy of recall as a basis for frequency knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*:387–396.

Whalen, J., Gallistel, C. R. & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science, 10*:130–137.

Williams, K. W. & Durso, F. T. (1986). Judging category frequency: Automaticity or availability? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*:387–396.