CHAPTER 2

FREQUENCY PROCESSING: A TWENTY-FIVE YEAR PERSPECTIVE

ROSE T. ZACKS AND LYNN HASHER

Abstract

We first proposed that frequency of occurrence information is 'automatically' encoded in the context of a general theoretical framework relating attention and memory encoding (Hasher & Zacks 1979). This chapter begins with a description of the origins of that framework, focusing on earlier evidence indicating that people of all ages and under a very broad range of circumstances reliably and unintentionally encode information about the relative frequencies of events. Notwithstanding challenges to the automatic encoding view, we believe this empirical generalization remains valid today. Additionally, we describe recent examples of findings from research on language processing and statistical reasoning that add to earlier evidence of the critical contribution of frequency knowledge to cognitive and social functioning. Finally, we note that, in a number of respects, the broader intellectual climate in psychology today is more consistent with our approach to memory encoding than was the intellectual climate of the 1970s.

In the early 1970s, there was already substantial evidence that people are both acutely sensitive to frequency of occurrence information and that they also use that information to solve a wide range of cognitive and behavioural problems (cf. Hasher & Zacks 1979, 1984, for reviews). These observations, together with some data of our own, initially led us to speculate that frequency of occurrence information is encoded 'automatically'. In this paper, we consider some of this early evidence along with more recent findings; we briefly review the theoretical framework we developed (the Automatic and Effortful framework); and consider the value of frequency information for a wide array of behaviours. In a final section, we discuss the intellectual climate at the time we developed this framework and the relevance of its theoretical claims to contemporary issues. In concluding we note that the available evidence continues to support the observations we started with 25 years ago: frequency of occurrence is a fundamental aspect of the information that people code about their experiences in the world and that information in turn plays a fundamentally important role in a wide range of behaviours.

Encoding of frequency of occurrence information

Early data on processing of frequency of occurrence included a number of salient demonstrations of people's pervasive sensitivity to how often the same event was repeated both in their everyday life and in the laboratory.[1] For example, evidence was available that people could actually reliably (and rather surprisingly) rank order the frequencies of individual words, of single letters, and even of pairs of letters in English (see Hasher & Zacks 1979, p. 369). These findings demonstrated that information in memory tracked the frequencies of naturally occurring events, even for such apparently meaningless units as letter pairs. In the laboratory, several studies had shown that participants could not only accurately estimate the relative frequency of presentation of words in a list, they could also accurately keep track of how often the same words occurred in two successive lists (Hintzman & Block 1971) and of how often a sentence was repeated in a gist versus a verbatim form (Gude & Zechmeister 1975). Furthermore, frequency judgements were known to be independent of both temporal recency of presentation and duration of presentation (cf. Hintzman 1976). Additionally, other studies (e.g. Hasher & Chromiak 1977) consistently found that the accuracy of frequency judgements was no greater when participants were explicitly informed prior to list presentation that they would be tested on frequency than when they were given only general memory instructions. There was also evidence that frequency judgement accuracy was not improved by practice or by feedback on performance (Hasher & Chromiak 1977).

Data reported in our 1979 paper added to this general picture of pervasive encoding of frequency information. In one of two experiments that compared groups of different ages, the participants were children from kindergarten and grades 1, 2, and 3. Following a general instruction that their memory would be tested, the children were shown pictures of common objects, each of which occurred 1, 2, 3, or 4 times in a randomly ordered sequence. They then saw each picture again and were asked to say how often it had appeared in the preceding list. The means of these judgements as a function of actual presentation frequency for each age group are reprinted in Fig. 2.1. For all four age groups, judged frequency increased regularly with increasing presentation frequency. Indeed, the frequency judgement functions were very similar across ages. This absence of age differences across the kindergarten through third grade age span contrasts sharply with the typical finding of large improvements on most other explicit memory tasks over this age range. Furthermore, Hasher and Chromiak (1977) had

---

[1] We note that the research we describe here stems from an era that predates the rise of interest in implicit learning and memory. Consequently, despite our claim of automatic encoding of frequency information and our informal observations that people were generally unaware of the degree to which they had encoded frequency of occurrence data, our research on frequency, like that of others, used *direct* (explicit) measures of frequency knowledge. That is, we used tasks that asked participants to directly judge absolute or relative frequency (e.g. to rank order the frequencies of a set of items). To the extent that implicit measures of some types of knowledge are more sensitive than explicit measures (e.g. Nissen and Bullemer 1987), the older findings as well as ours can be seen, in retrospect, even more striking than they appeared at the time.

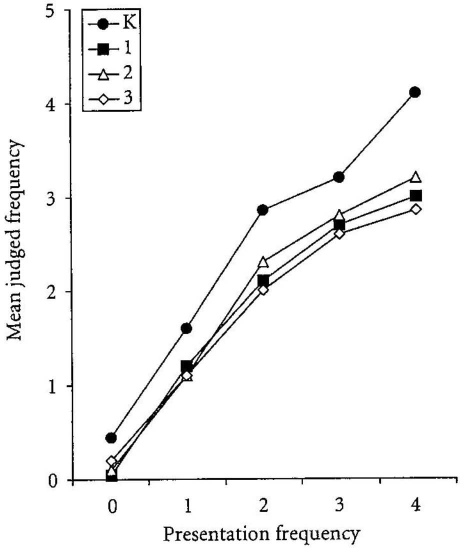

Figure 2.1 Estimated frequency of occurrence as a function of actual frequency of occurrence for children from kindergarten through third grade. Reprinted from Hasher and Zacks (1979). David Goldstein was a collaborator on this experiment.

found no age differences testing participants ranging in age between second graders and college students, suggesting that the ability to encode frequency information is fully functional quite early in life, as more recent work has demonstrated even more compellingly and at an even younger age (see the description below of the work of Saffran and colleagues). Finally, we also found no evidence of an age difference in the patterns of frequency judgements between college students and healthy older adults.

At this time, a particularly compelling finding for us emerged from additional analyses done on several sets of frequency judgements. We computed correlations between judged and actual frequency of occurrence for individual participants and found that the great majority of these correlations were strongly positive. For example, in one study (Hasher & Zacks 1979, Experiment 2) all but one (of 80) of the correlations was significant. Indeed, in that paper, the mean of the *individual* correlations was 0.77. Thus, the sensitive encoding of frequency of occurrence characterizes the performance of *each individual participant* and is not an artifact of averaging across participants, some of whom do process frequency and some of whom do not.

That ability differences also do not impact on frequency processing is suggested by findings from our final example of past research on processing of frequency of occurrence. One of the experiments reported by Zacks, Hasher, and Sanft (1982) included groups of college students who differed by approximately 140 points in verbal Scholastic Aptitude Test (SAT) scores.[2] Across a variety of instructional conditions, there was no hint that the higher and lower verbal SAT groups differed in memory for frequency of occurrence.

---

[2] The SAT is a widely administered test of verbal and mathematical reasoning skills that is used by many US colleges and universities to make admission decisions. This size difference is a difference of approximately 1.4 standard deviation units.

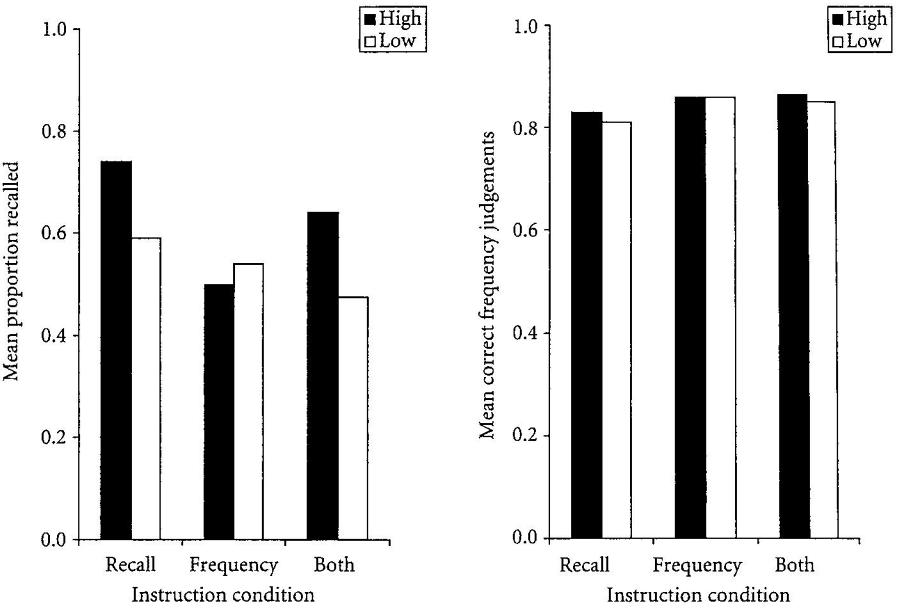

Figure 2.2 Effects of verbal ability and instruction on free recall (left panel) and relative frequency judgements (right panel). Groups of participants tested on both memory tasks were instructed prior to study that they would receive a recall test, a frequency test, or both. The filled bars are for the higher ability groups and the unfilled bars for the lower ability groups. Figure 2.2 is based on results reported in Zacks, Hasher, and Sanft (1982).

These examples indicated that people of a wide range of ages and abilities and under a wide range of conditions reliably encode a record of the frequency with which particular events occur. Importantly, the pattern of findings that we and others had obtained on memory for frequency (and to some degree for memory for spatial location and temporal order; see Hasher & Zacks 1979) is strikingly different than findings that have long been known to occur on other, more standard explicit memory tests including paired-associate, serial, and free recall, and recognition tests. On such tasks, age, verbal ability, instructions, practice, and feedback *do* make a difference (Hasher & Zacks 1979, 1984; Zacks *et al.* 1982). For example, the Zacks *et al.* (1982) experiment included groups receiving a free recall test rather than a frequency test. On that memory test, there were robust effects of ability and instructions—higher ability and test-appropriate instructions were associated with better free recall performance (see Fig. 2.2).

Automatic and Effortful encoding framework

Our 1979 and 1984 papers set out a theoretical framework that encompassed these strikingly divergent patterns of memory performance for frequency of occurrence information on the one hand and for item memory on the other. The primary distinction in the framework is between two types of encoding mechanisms, automatic and effortful. *Automatic* encoding mechanisms make minimal demands on attentional resources. They handle fundamental attributes of experience, including frequency, spatial location, and temporal order, and ensure that, for attended events,[3] these attributes are 'continually registered in memory, whatever the age, the ability, the education, or the motivation of an individual' (Hasher & Zacks 1984, p. 1372). (In a similar vein, Underwood [1969] had earlier argued that spatial, temporal, and frequency attributes are encoded as part of the memory of an event regardless of the type of encoding activity.) By contrast, *effortful* encoding mechanisms require considerable attentional capacity. They include such optional encoding operations as conscious rehearsal, organization, imagery, and semantic elaboration that are critical to good performance on most standard episodic memory tests. Because available attentional resources can vary across individuals and situations (Kahneman 1973), the use and effectiveness of effortful encoding operations and the resulting memory performance can vary with age, ability, educational experiences, and motivation.

A major focus of our work on this framework was on the encoding of frequency information (cf. Hasher & Zacks 1984). At the time, automatic processing was a lively topic in the cognitive literature, but with little consensus on exactly how to define automaticity (Shiffrin 1988). Along with most other theorists (but not all; see Allport 1989), we believed that *the* fundamental characteristic of automatic processing was the minimal demands such processing placed on our limited capacity attentional system. This notion, along with other considerations (e.g. the presumption that there are systematic differences in attentional capacity as function of age and other individual difference factors), allowed us to generate an operational definition of automatic encoding. This definition took the form of a set of criteria that together served as converging operations for concluding that a particular encoding operation was automatic. The specific criteria were that automatic encoding mechanisms (a) operate without intention; (b) do not benefit from intention to encode the particular attribute; (c) do not benefit from training to encode the attribute or from feedback; (d) show minimal individual differences; (e) show minimal age differences; and (f) show minimal impact of state (e.g. arousal level) and situational (e.g. divided attention demands) factors that otherwise impact on available attentional capacity. By contrast, effortful encoding operations were defined by the opposite pattern on each of these criteria.

In part because our main goal was to highlight the unique properties of encoding of frequency, spatial location, and temporal information, our original formulation of the Automatic and Effortful framework did not make specific claims about how these types of information were represented in memory. However, in our 1984 paper on frequency of occurrence, we did briefly review arguments favouring a multiple-trace view over alternatives such as strength-based or frequency counter views (Hasher & Zacks 1984, pp. 129–130). More specifically, it seemed to us and to others (Hintzman 1976; Greene

---

[3] Hasher and Zacks (1979) assumed that attending to an event was a precondition for the occurrence of automatic as well as effortful encoding operations. (This point has sometimes been ignored.) Automatic encoding operations require no further attentional resources, whereas effortful encoding operations do require additional resources.

1992) easier to reconcile the evidence that we continually and unintentionally store a fine-grained record of the repetition of individual events with a memory system in which each attended occurrence of an event results in the establishment of an independent memory trace rather than with a system in which each occurrence produces an increment in the strength of a single memory trace or a counter. Nonetheless, we recognized that such arguments are not definitive and that alternatives to the multiple-trace view could not be entirely ruled out. In fact, Alba, Chromiak, Hasher, and Attig (1980) suggested that encoding of superordinate category frequency involved incrementing a frequency tag or counter for the category label as each new category instance was presented. Contributions to the current volume demonstrate that this issue remains unresolved. The positions of different authors on how frequency information is mentally represented include both strength-type (Whalen, Gallistel & Gelman 1999) and multiple-trace-type (Fiedler, this volume) views. In addition, a number of the current authors (e.g. Brown, this volume; Haberstroh & Betsch, this volume) propose combined views in which frequency information is assumed to be represented in multiple ways in memory. Given that years of research have yielded at least partial support for several different proposals about how frequency is coded in memory, this latter approach may have some face validity. But adoption of a multiple representation approach would not be the end of the matter. As authors like Brown (this volume) and Haberstroh and Betsch (this volume) are well aware, other questions would arise, including ones about how situational factors influence reliance on different frequency representations.

Automatic encoding criteria: need for revision?

We have already described several examples of the earlier (before 1984) research that supported our contention that encoding of frequency information conformed to each of our automatic encoding criteria. In our view, more recent research continues to demonstrate people's striking and largely unintentional sensitivity to frequency of occurrence information. Nonetheless, findings have been reported that, by at least some accounts, call into question the validity of one or more of our automatic criteria with respect to frequency encoding. For example, the literature contains findings of significant individual and developmental differences in memory for frequency, including cases in which older adults were found to be slightly less proficient at encoding frequency of occurrence than younger adults (e.g. Kausler, Lichty & Hakami 1984). There are also reports of significant effects of encoding variables such as level or processing and, occasionally, benefits of prewarning participants of a memory test (for a summary of such findings, see Greene 1992, pp. 143–145).

Additional questions have been raised about the *accuracy* of frequency judgements, particularly as measured in tasks that require participants to assign numerical frequency estimates to items varying in actual presentation frequency. Although, in general, these estimates accurately track the *relative* frequencies of the items (i.e. judged frequency increases regularly with increasing actual frequency), the *absolute* accuracy of the provided estimates is subject to various distortions. In particular, low frequencies tend to be overestimated and high frequencies tend to be underestimated (see Fig. 2.1); and the estimates are affected by such factors as item retrievability (the 'availability heuristic') and specific characteristics of the judgement task (e.g. whether or not a range of frequencies is indicated and if it is, the specific low and high values; cf. Schwarz & Wänke, this volume; Zechmeister & Nyberg 1982).

That factors influencing accuracy and bias in frequency judgements continue to be of interest to cognitive and social cognition researchers is clear from a number of contributions to this volume. One example is the work described by Schwarz and Wänke relating to the availability heuristic. Their findings indicate that the subjective 'ease' of retrieval of target instances can influence frequency estimates independently of amount recalled, but their findings also show that relevant situational and instructional cues can lead to discounting of the ease factor. Another example is Fielder's research (this volume) showing a strong effect of retrieval cue structure on the accuracy of frequency judgements.

With respect to the hypothesis of automatic encoding of frequency information, one critical point illustrated by the work just described is the fact that multiple factors—not just stored information—influence frequency judgements. That is, to understand how people arrive at frequency estimates under even relatively sterile laboratory conditions, let alone in the complexity of real life, requires consideration of different judgement heuristics, retrieval factors, contextual variables, and so on. Admittedly then, a framework such as ours that concerned only *encoding* mechanisms provides, at best, only a partial account of how people arrive at the frequency estimates they produce on frequency judgement tasks. On the other hand, the multifaceted nature of frequency estimation processes permits the suggestion that inaccuracies and biases in frequency judgements, developmental differences, and so on are not necessarily contrary to the automatic *encoding* view. At least some of these effects could be a consequence of postencoding factors. Indeed, as suggested above, it appears that relative accuracy measures (e.g. rank order correlations between judged and actual frequency) are less likely to show effects not predicted by the automaticity view than are absolute accuracy measures (e.g. mean deviation of judged from true frequency). Arguably, this pattern is due to the lower sensitivity of the relative measures to potential sources of postencoding bias such as anchoring effects and scaling distortion in number assignment. In general, we remain convinced that the bulk of the literature indicates a robust sensitivity to the relative frequency of events across a wide range of circumstances and subject variables. Perhaps even more convincing on this point is evidence that demonstrates sensitivity to frequency of occurrence information through indirect or implicit measures, as is the case in the studies described in the next section.

The uses of frequency information

In addition to establishing the principle of automatic encoding of frequency (and of spatial location and temporal order) information, a major goal of our 1979 and 1984 papers was to consider the purposes to which automatically encoded information might be put. That is, having suggested that frequency, spatial location, and temporal order information were automatically encoded, we wanted to explore their contributions to cognition. At a general level, it seemed reasonable to consider the possibility that frequency, space, and time were fundamental organizing dimensions of behaviour.

Rather than attempting a broad discussion of the more specific uses of all types of automatically encoded information, the 1984 paper (Hasher & Zacks 1984) focused on the case of frequency. In particular, that paper summarized research demonstrating that frequency knowledge plays a critical role in event memory, in the representation and acquisition of new knowledge and skills, in decision making, and in cognitive and social development. We noted, for example, the considerable evidence indicating that category representations are based, in part at least, on stored information about feature frequency. We also noted evidence suggesting that children's sensitivity to the frequency with which members of their own versus the opposite sex display particular attitudes and behaviours contributes to their acquisition of gender-appropriate social roles. As well, we found that one source of beliefs is the frequency with which assertions are made (Hasher, Goldstein & Toppino 1977). Such findings supported our conclusion that frequency information was critical to at least cognitive functioning but probably was useful across a broad range of behavioural functions. More recent findings, including ones described throughout this volume, place our conclusion about the importance of frequency information to behaviour on even firmer grounds now than in 1984. Consider some dramatic examples from the recent literature.

Our first example comes from research on early language acquisition. Before a child learns the meanings of words, or the syntactic roles words play in sentences, the child must be able to break the speech record up into words. But how does a child pick out individual words from a speech stream that can be characterized as 'mostly continuous, without consistent pauses or other acoustic cues marking word boundaries' (Saffran, Newport, Aslin, Tunick & Barrueco 1997, p. 102)? An impressive series of studies by Saffran and colleagues (Aslin, Saffran & Newport 1998; Saffran, Aslin & Newport 1996; Saffran, Newport & Aslin 1996; Saffran *et al.* 1997) indicates that sensitivity to the frequency with which different sounds follow each other in speech helps babies (and children and adults) to segment the speech stream into words.

To arrive at this conclusion, Saffran and colleagues use a paradigm in which participants as young as eight months of age are exposed to brief samples of an artificial language consisting of several three-syllable nonsense 'words' (e.g. *bupada* and *tubitu*). Random repetitions of the words are strung together in a continuous auditory stream containing no acoustic or prosodic cues to the boundaries between words. In this situation, the primary cue to word boundaries is the greater frequency of two-syllable sequences that occur within a word (e.g. *bupa* and *pada*) as compared to two-syllable sequences that occur between two words (e.g. *datu* and *tubu*). Use of this frequency cue as a basis for word segmentation has been demonstrated in infants using a selective attention paradigm (Aslin *et al.* 1998; Saffran, Aslin & Newport 1996) and in 6–7 year old children and adults (Saffran, Newport & Aslin 1996; Saffran *et al.* 1997) using forced-choice recognition procedures. Other important findings are that children and adults are equally adept at this kind of learning (Saffran, Newport & Aslin 1996; Saffran *et al.* 1997) and that intentional and incidental learning of the words are equally effective (Saffran *et al.* 1997). The apparent age invariance in these results, the lack of a benefit from intentional encoding, and most importantly, the demonstration of the exquisite sensitivity to frequency information (including under the impoverished conditions of this paradigm) are all aspects of the Saffran *et al.* findings that echo the sorts of observations that originally got us interested in the processing of frequency—and, that confirm the suggestion that frequency serves as a fundamental attribute for organizing the physical world.

It can be noted that in contrast to the highly artificial speech samples used by Saffran and colleagues, normal speech provides additional frequency-based cues that people use to help them pick out word boundaries (Kelly & Martin 1994). For example, it has been shown that word segmentation is influenced by the knowledge that certain syllables (e.g. *the, to*) are more frequently followed by the start of a new word than others (e.g. *a, in*) and by the knowledge that words more frequently begin with a syllable containing a fully realized vowel than with a syllable containing a reduced vowel.

Segmentation of words is not the only area of language to which frequency information makes a contribution. For example, Hasher and Zacks (1984) reviewed results suggesting that sensitivity to 'single-letter positional frequency' (the frequency with which particular letters occur in specific positions within words) plays an implicit role in written wording decoding (see also Sedlmeier, Hertwig & Gigerenzer 1998.) Indeed, it is this knowledge that may contribute to the well-known word superiority effect, by which a letter within a word is more easily identified than is a single letter presented alone.

More recently, important claims have been made about the role of frequency information in syntactic parsing, that is, in the determination of the syntactic structure of a sentence. So-called 'constraint-based views' of syntactic parsing (e.g. Trueswell 1996) claim that parsing is guided by multiple types of information, including knowledge of the frequencies with which particular content words, especially verbs, occur in different syntactic roles. These influences can be demonstrated with respect to the resolution of temporary syntactic ambiguities. Consider, for example, sentences containing verbs (e.g. *searched, accepted*) that are the same whether they are functioning as a past tense verb (e.g. *The thief searched the room*) or as a past participle (e.g. *The thief searched by the police was indicted*). As a result of this overlap in function, some verbs create sentences that are temporarily ambiguous until after a critical word is presented (e.g. 'by' in the second example sentence above). Trueswell (1996) showed that the amount of difficulty people have in arriving at a correct interpretation of such sentences depends in part on the frequency with which a particular verb occurs as a main verb versus as a participle. In particular, sentences containing the less preferred (and overall, less frequent) participle form produce relatively little difficulty for comprehenders when the verb involved has a high participle frequency (e.g. *accepted*) as compared to cases in which the verb involved has a low participle frequency (e.g. *searched*).

Turning now the importance of frequency information in a nonlanguage domain, we consider work on statistical reasoning. In sharp contrast to results deriving from the 'heuristics and biases' approach initiated by Kahneman and Tversky (1973), several studies published in the1990s have demonstrated that people possess good statistical intuitions and that they display these intuitions when the relevant information is presented to them as frequencies rather than as probabilities. In particular, 'base rate neglect' is a prominent feature of reasoning on Bayesian reasoning problems when a probabilistic format is used to present the relevant data. That is, people are generally insensitive to the prior probability or base rate of a particular hypothesis (e.g. that a woman has breast cancer) when they estimate the probability of that hypothesis given that a particular fact is known (e.g. she has a positive mammogram test). When base rates are low, as they are for most specific diseases, gross overestimates are made of the probability that a positive test result means a positive diagnosis. By contrast, Cosmides and Tooby (1996) and Gigerenzer and Hoffrage (1995) (see also Hertwig & Hoffrage, this volume) have found that the majority of participants (up to 92% in Cosmides & Tooby's data) do take base rate appropriately into account when a frequency format is used for the presentation of the critical information (for more details, see Chapter 18). Cosmides and Tooby argue that such results support the 'frequentist hypothesis' which states that 'some of our inductive reasoning mechanisms do embody a calculus of probability, but they are designed to take frequency information as input and produce frequencies as output' (1996, p. 3). Likewise, Gigererenzer and Hoffrage (1995, p. 697) have argued that 'the (human) mind is attuned to frequencies'. And we would of course agree.

Descriptions of the uses of frequency information in many additional domains are found throughout this book. To take just one example, in chapter 16 Alba presents a picture of pervasive effects of frequency information in consumer decision making. A particularly striking instance of these effects comes from studies suggesting that the relative frequency of lower prices for particular products is a determining factor in people's judgements that one store is overall cheaper than its competitors. More specifically, when asked to pick the store that is cheaper on a market basket of items, there is a strong tendency to pick the store that is cheaper on *more* of the items. This is true even when the other store has a large price advantage on the items on which it is cheaper, with the result that the total cost of the market basket is actually the same across the two stores.

In summary, like the claim that frequency of occurrence is continually encoded into memory, the claim that frequency information serves a wide range of cognitive and social functions remains on firm ground. If anything, evidence such as that just reviewed places our claim about the uses of frequency information on broader and firmer ground than it was 25 years ago. In fact, very similar claims were recently highlighted in a paper by Kelly and Martin (1994). These authors argue that animals and humans share a 'fine-grained sensitivity to probabilistic (i.e., relative frequency) patterns in their environment' (Kelly & Martin 1994, p. 105, parenthetical comment added), and that this sensitivity facilitates numerous behaviours in both animals (e.g. conditioning, foraging) and humans (e.g. depth perception, language processing).

The ideas outlined in Hasher and Zacks (1979, 1984) have proved their staying power. Of course, our ideas developed in a particular intellectual climate in (mostly) experimental psychology and in the next section of the paper we explore the relation between the Automatic and Effortful framework and the intellectual climate in psychology 25 years ago and now.

Historical context

At the time our first paper was published, research in human memory was firmly in the grasp of the 'cognitive revolution' (e.g. Lachman, Lachman & Butterfield, 1979), a presumed advance over the associationist/functionalist approach to memory (known then as 'verbal learning'[4]) that had dominated the field of human experimental psychology in North America for 40–50 years. Despite the revolution, some critical intellectual characteristics (including pre-theoretical assumptions) of the earlier tradition remained intact in the new cognitive approach to the study of memory. Several of these characteristics are worth noting because our own ideas either explicitly or implicitly rejected them. In particular, our viewpoint disagreed with the continued pursuit of 'general laws' of learning, memory, and cognition, with the continued study of memory as independent from other cognitive processes, particularly attention, and with the focus on deliberate forms of learning. We develop these points in the following paragraphs.

A primary goal of the associationist tradition, whether humans or other animals were the focus of study, was the pursuit of general laws of learning and memory. In our view, an overly stringent application of this goal led to a number of shortcomings, including the failure to note that behaviour has an ecological and evolutionary component to it. Because the goal of classic experimental psychology was (and of course to some degree, still is) to establish general laws of learning, differences among species, their development, and their ecological niche were largely considered nuisance factors (or at least ones to be looked at after the general laws had been specified) rather than issues to be studied in their own right. The result was an absence of diversity in the behaviours and subjects studied: In human learning and memory laboratories, young college students (often sophomores, mostly attending elite research institutions) were the focus of research. In animal laboratories, it was rats and pigeons. In both types of laboratories, the range of learning paradigms was similarly restricted.

The limitations of so narrow a focus became apparent in some animal laboratories in the late 1960s and early 1970s with the recognition of critical boundary conditions for basic learning in nonhuman animals. For example, one standard principle of classical conditioning was that acquisition of an association between the conditioned and unconditioned stimuli (CS and US, respectively) is strongly dependent on the CS-US interval. Much evidence had shown that if more than a few seconds separates the CS and the US, conditioning proceeds very slowly, if at all (Kimble 1961). However, research in the 1960s (e.g. Garcia, Ervin & Koelling 1966) indicated that this general rule does not apply in the case of a type of learning that has survival value for the organism, specifically, learning to avoid food that has in the past made one sick, the 'bait shyness' effect. In research that brought such learning into the lab, rats (and other organisms) were found to readily acquire an aversion to a taste (e.g. the taste of saccharine-sweetened water) that was paired with sickness (e.g. from exposure to X-ray radiation) even though the delay between the CS (sweet taste) and US (sickness) was an hour or more. Furthermore, such taste aversion learning was found to require only a single learning trial, and aversions were not formed to nontaste stimuli (e.g. a flashing bright light or a clicking noise) even after many pairings with sickness (Garcia *et al.* 1966).

This and other work was making it apparent at that time that members of a species are 'prepared' by evolutionary forces to learn some associations easily but not others (e.g. Seligman 1970). The laws of learning established for unprepared associations do not necessarily apply to prepared ones. Moreover, because prepared learning mechanisms presumably evolved to solve the particular problems that a species has in meeting basic needs (food, water, safety, procreation), some of these will be species specific, dependent on the particular ecological niche that the species occupies.

By the early 1970s, evolutionary considerations and notions of adaptive function had not much filtered into the study of human learning and memory (or other cognitive functions for that matter). Our own views were heavily influenced by the animal research findings. We believed that adaptive function was an important consideration for human as well as nonhuman research. Similar to Seligman's (1970) notion of preparedness, we specifically proposed that humans have an innate ability to automatically encode frequency, spatial location, and temporal order because these are fundamental attributes of experience that define the flow of events in our environment and that enable us to solve critical problems (e.g. see Hasher & Zacks 1979; p. 360). Our views represent an early example of what has since come to be known as evolutionary psychology (cf. Barkow, Cosmides & Tooby 1992). Our views also antedate subsequent work with nonhuman animals that also suggests the fundamental importance of, and sensitivity to frequency (Shettleworth 1998, pp. 363–377).

In addition to the lack of attention to adaptive function, the pursuit of general laws of learning restricted research to a limited number of subject types—in human research, primarily the college sophomore. Prompted by research on memory development in animals (e.g. Campbell & Spear 1972), by the then more limited research on young children and older adults, and by some of our own data comparing individuals of different ages and abilities (see above), we came to believe that much could be learned by using developmental and individual difference approaches to test our theoretical notions. In this view, we were particularly influenced by Underwood's (1975) notion that consideration of individual differences can play a beneficial role in theory development.

A second property of North American psychology in the 1970s that we disagreed with was its tendency to treat learning and memory as topics that were easily separable from other cognitive domains such as language, problem solving, and attention. Our own view was (and remains) that memory is central to the study of any number of allied cognitive fields, in that both its implicit and explicit components limit behaviour. However, it was (and remains) our view that memory and attention are actually integral aspects of each other and neither is likely to be successfully studied without explicit recognition of the other (see Hasher, Zacks & May 1999, for our more recent views on this topic). To a large degree, the development of our views on this point coincided with the publication of Kahneman's book, *Attention and effort*, in 1973. In that book, Kahneman forcefully argued that cognitive processes have varying attentional resource demands and that although attentional capacity is limited, the limit is not a rigid one but varies significantly both between and within individuals (e.g. as a result of mood, arousal, etc.). These ideas were central to the Automatic and Effortful framework.

Finally, there was at the time what we saw as an exaggerated emphasis on the importance of intentional learning (though see Postman's [1964] work on incidental memory) and associated memory strategies (such as rehearsal, imagery mnemonics, etc.) for guaranteeing that experienced information left a residue in memory. Several books and papers outside of the human memory tradition had made the point in one way or another that much of behaviour is actually either acquired implicitly or used implicitly, or both. For example, Goffman (1963) suggested that social and cultural rules were acquired in this way. Skinner (1971) argued that much of human behaviour was under the control of stimuli that people were unaware of. And there were a few startling findings in the human literature, notably the early suggestion of implicit retrieval by amnesics and their controls of lists of words they could not recall explicitly (Warrington &

Weiskrantz 1968), that seemed quite telling with respect to the overemphasis on deliberate processes for learning and memory. With respect to frequency information, the evidence that people knew the relative frequencies of single letters, pairs of letters, syllables and other units of experience all suggested to us that much information could be acquired without the deliberate intention of the learner. Indeed, we were impressed that participants in our studies inaccurately denied their ability to judge frequency reliably, especially in the case of incidentally encoded information. Our reviews of the ways in which frequency knowledge was used (e.g. to establish the truth of assertions, to make decisions, and so on) was in keeping the implicit memory findings of Warrington and Weiskrantz (1968) and the incidental memory findings of Postman (1964), and of course is in keeping with the current wide interest in implicit acquisition and utilization of knowledge.

In many ways, cognitive psychology in the early 2000s fits more tightly with the ideas raised by the Automatic and Effortful framework. In particular, much research is directed at the study of implicit knowledge and at its implicit utilization. There is direct consideration in the literature of adaptive functions of cognition, and there are burgeoning literatures on the development of memory and on its pattern of decline in normal and abnormal aging. There are also many who are now concerned with the integral relationship between attention and memory.

Summary and reflections

From the vantage point of hindsight, it is clear that we might have done a few things differently. First, our choice of the term 'automatic' created a number of problems that stemmed from confusions between our use of the term (as described above) and its quite different use in several widely cited papers (Posner & Snyder 1975; Shiffrin & Schneider 1977). For example, Shiffrin and Schneider used the term 'automatic' processing to describe the result of extensive training under consistent mapping conditions, whereas we used the same term to refer to a largely innate mechanism that results in the inevitable encoding of certain attributes of attended events. We would have avoided some confusion and the excess theoretical baggage that tended to get attached to our view if we had used a term like 'obligatory' encoding.

Secondly, we may have diverted attention from our central points (sensitivity to and importance of frequency information) by trying to specify the criteria for automatic encoding. Our goal was definitional precision, but this effort fostered an emphasis in subsequent research on whether or not *all* the criteria applied to a particular encoding process. This focus on the criteria probably detracted from the basic idea of incidental, obligatory encoding of critical aspects of experience. Additionally, we might well have thought more about variables that influence levels of accuracy in frequency judgements and what these variables meant for our notion of automaticity. Likewise, we acknowledge our minimal contribution to debates on the underlying representation of frequency information. (Both issues have justifiably received considerable attention, but judging from other contributions to this volume, answers remain incomplete.)

These considerations notwithstanding, we believe that the Automatic and Effortful view was basically correct in its proposals about the encoding of frequency of occurrence information (and to a lesser degree, temporal order and spatial location). That is, we still believe the following:

1. The encoding of frequency information is an inevitable consequence of attending to events and, in that sense, is obligatory;
2. Frequency knowledge plays a critical role in many implicit cognitive functions;
3. Research limited to one sample type (in North America, the college sophomore) is likely to be less informative than is research that adds age as a variable and that takes into consideration nonnormal samples and factors such as mood and arousal changes; and
4. Memory and attention are interdependent cognitive functions that cannot be adequately understood in isolation from each other.

We close with an observation that we made in the course of the reading we did to prepare this chapter. We believe there is (e.g. Kelly & Martin 1994) and has been for some time (e.g. Hintzman 1976; Underwood 1969), considerable consensus about our acute sensitivity to frequency information and about the critical roles that frequency information plays in cognition. Surprisingly, this knowledge is unlikely to be conveyed to students of cognitive psychology—an informal recent survey of cognitive psychology textbooks suggests that few of them pay more than passing attention to the processing of frequency information. This seems curious to us based on our reading of the literature as well as other chapters in this volume.

Acknowledgements

The preparation of this chapter was supported by Grant Number AGO 4306 from the National Institute on Aging. We express our continued gratitude to David Goldstein for his assistance with our early work with children.

References

Alba, J. W., Chromiak, W., Hasher, L. & Attig, M. S. (1980). Automatic encoding of category size information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 4*:370–378.

Allport, A. (1989). Visual attention. In M. I. Posner (ed.) *Foundations of cognitive science* (pp. 631–682). Cambridge, MA: MIT Press.

Aslin, R. N., Saffran, J. R. & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*:321–324.

Barkow, J. H., Cosmides, L. & Tooby, J. (1992). *The adapted mind: Evolutionary psychology and the generation of culture.* New York: Oxford University Press.

Campbell, B. A. & Spear, N. E. (1972). Ontogeny of memory. *Psychological Review, 79*:215–236.

Cosmides, L. & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgement under uncertainty. *Cognition, 58*:1–73.

Garcia, J., Ervin, F. R. & Koelling, R. A. (1966). Learning with prolonged delay of reinforcement. *Psychonomic Science, 5*:121–122.

Gigerenzer, G. & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*:684–704.

Goffman, E. (1963). *Behaviour in public places: Notes on the social organization of gatherings*. New York: Free Press.

Greene, R. L. (1992). *Human memory: Paradigms and paradoxes*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gude, C. & Zechmeister, E. B. (1975). Frequency judgements for the 'gist' of sentences. *American Journal of Psychology, 88*:385–396.

Hasher, L. & Chromiak, W. (1977). The processing of frequency information: An automatic mechanism? *Journal of Verbal Learning and Verbal Behaviour, 16*:173–184.

Hasher, L., Goldstein, D. & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behaviour, 16*:107–112.

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

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

Hasher, L., Zacks, R. T. & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In A. Koriat and D. Gopher (eds.) *Attention and performance XVII. Cognitive regulation of performance: Interaction of theory and application* (pp. 653–675). Cambridge, MA: MIT Press.

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. & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology, 88*:297–306.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D. & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*:237–251.

Kausler, D. H., Lichty, W. & Hakami, M. K. (1984). Frequency judgements for distractor items in a short-term memory task: Instructional variation and adult age differences. *Journal of Verbal Learning and Verbal Behaviour, 23*:660–668.

Kelly, M. H. & Martin, S. (1994). Domain-general abilities applied to domain-specific tasks: Sensitivity to probabilities in perception, cognition, and language. *Lingua, 92*:105–140.

Kimble, G. A. (1961). *Hilgard and Marquis's conditioning and learning*. New York: Appleton-Century-Crofts.

Lachman, R., Lachman, J. L. & Butterfield, E. C. (1979). *Cognitive psychology and information processing: An introduction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Nissen, M. J. & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*:1–32.

Posner, M. I. & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (ed.) *Information processing and cognition: The Loyola symposium* (pp. 55–85). Hillsdale, NJ: Lawrence Erlbaum Associates.

Postman, L. (1964). Short-term memory and incidental learning. In A. Melton (ed.) *Categories of human learning* (pp. 146–201). New York: Academic Press.

Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*:1926–1928.

Saffran, J. R., Newport, E. L. & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language, 35*:606–621.

Saffran, J. R. Newport, E. L., Aslin, R. N., Tunick, R. A. & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science, 8*:101–105.

Sedlmeier, P., Hertwig, R. & Gigerenzer, G. (1998). Are judgements of positional frequencies of letters systematically biased due to availability? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*:754–770.

Seligman, M. E. P. (1970). On the generality of the laws of learning. *Psychological Review, 77*:406–418.

Shettleworth, S. J. (1998). *Cognition, evolution, and behaviour.* New York: Oxford University Press.

Shiffrin, R. M. (1988). Attention. In R. C. Atkinson, R. J. Herrnstein, G. Linszey, and R. D. Luce (eds.) *Steven's handbook of experimental psychology: vol. 2. Learning and cognition* (pp. 739–811). New York: John Wiley and Sons.

Shiffrin, R. M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic encoding, and a general theory. *Psychological Review, 84*:127–190.

Skinner, B. F. (1971). *Beyond freedom and dignity.* New York: Alfred A. Knopf.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language, 35*:566–585.

Underwood, B. J. (1969). Attributes of memory. *Psychological Review, 76*:559–573.

Underwood, B. J. (1975). Individual differences as a crucible in theory construction. *American Psychologist, 30*:128–134.

Warrington, E. K. & Weiskrantz, L. (1968). A new method of testing long-term retention with special reference to amnesic patients. *Nature, 217*:972–974.

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

Zacks, R. T., Hasher, L. & Sanft, H. (1982). Automatic encoding of event frequency: Further findings. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*:106–116.

Zechmeister, E. B. & Nyberg, S. E. (1982). *Human memory: An introduction to research and theory.* Monterey, CA: Brooks/Cole Publishing Co.