Chem. Senses 23: 309–326, 1998

Odor Identification: Perceptual and Semantic Dimensions

William S. Cain, René de Wijk, Christine Lulejian[1], Franc Schiet[1] and Lai-Chu See[1]

Chemosensory Perception Laboratory, University of California, San Diego, La Jolla, CA 92093-0957 and [1]John B. Pierce Laboratory, New Haven, CT 06519, USA

Correspondence to be sent to: Dr William S. Cain, Chemosensory Perception Laboratory, Department of Surgery (Otolaryngology), University of California, San Diego, La Jolla, CA 92093-0957, USA. e-mail: wcain@ucsd.edu

Abstract

Five studies explored identification of odors as an aspect of semantic memory. All dealt in one way or another with the accessibility of acquired olfactory information. The first study examined stability and showed that, consistent with personal reports, people can fail to identify an odor one day yet succeed another. Failure turned more commonly to success than vice versa, and once success occurred it tended to recur. Confidence ratings implied that subjects generally knew the quality of their answers. Even incorrect names, though, often carried considerable information which sometimes reflected a semantic and sometimes a perceptual source of errors. The second study showed that profiling odors via the American Society of Testing and Materials list of attributes, an exercise in depth of processing, effected no increment in the identifiability/accessibility beyond an unelaborated second attempt at retrieval. The third study showed that subjects had only a weak ability to predict the relative recognizability of odors they had failed to identify. Whereas the strength of the feeling that they would 'know' an answer if offered choices did not associate significantly with performance for odors, it did for trivia questions. The fourth study demonstrated an association between ability to discriminate among one set of odors and to identify another, but this emerged only after subjects had received feedback about identity, which essentially changed the task to one of recognition and effectively stabilized access. The fifth study illustrated that feedback improves performance dramatically only for odors involved with it, but that mere retrieval leads to some improvement. The studies suggest a research agenda that could include supplemental use of confidence judgements both retrospectively and prospectively in the same subjects to indicate the amount of accessible semantic information; use of second and third guesses to examine subjects' simultaneously held hypotheses about identity; use of category cuing or similar techniques to discover the minimum semantic information needed to precipitate identification; some use of subjects trained in quantitative descriptive analysis to explore whether such training enhances semantic memory; and judicious use of mixtures to explore perceptual versus semantic errors of identification.

Introduction

This set of five studies delves into aspects of semantic memory for odors, or what some would prefer to call generic knowledge of odors, in particular odor identifica- tion (Crowder and Schab, 1995). Odor identification often seems subjectively unstable and unpredictable. Despite its limitations, though, a form of odor identification serves as a tool in clinical evaluation (Cain *et al.*, 1983, 1988; Cain and Gent, 1986; Lawless and Zwillenberg, 1983; Doty *et al.*, 1984; Wright, 1987). In that role, it works well when patients have choices to guide them. Otherwise, the unpredictability of performance defeats its diagnostic value (Sumner, 1962; Cain and Krause, 1979; Cain, 1982b). No studies have characterized the variability well, no less explained its meaning.

A need to instruct people about odors offers additional practical incentive to understand the features of semantic memory for odors (Cain, 1979, 1980). It deals in the main with knowledge acquired naturalistically rather than via instruction. It also forms the foundation for all subsequent odor learning, much of which may not actually comprise new learning so much as stabilization of and access to material already learned.

The 'student' for odor instruction ranges from the householder who needs to identify the smell of gas (see Cain and Turk, 1985) to the aspiring perfumer. The more one knows about how people identify odors, the better one might improve performance. Nevertheless, the transmission of veridical information through olfaction also deserves attention for its own sake. Some have argued, for example, that odor memory stands apart from sensory memory in other modalities (Herz and Eich, 1995; Herz and Engen, 1996). A defining issue concerns the role of semantic retrieval in odor memory.

Five topics relevant to semantic memory for odors receive attention: (i) day-to-day variation in performance; (ii) attribute ratings as a possible means to precipitate identification; (iii) the predictive validity of feelings of knowing about odor identity; (iv) the relation between identification

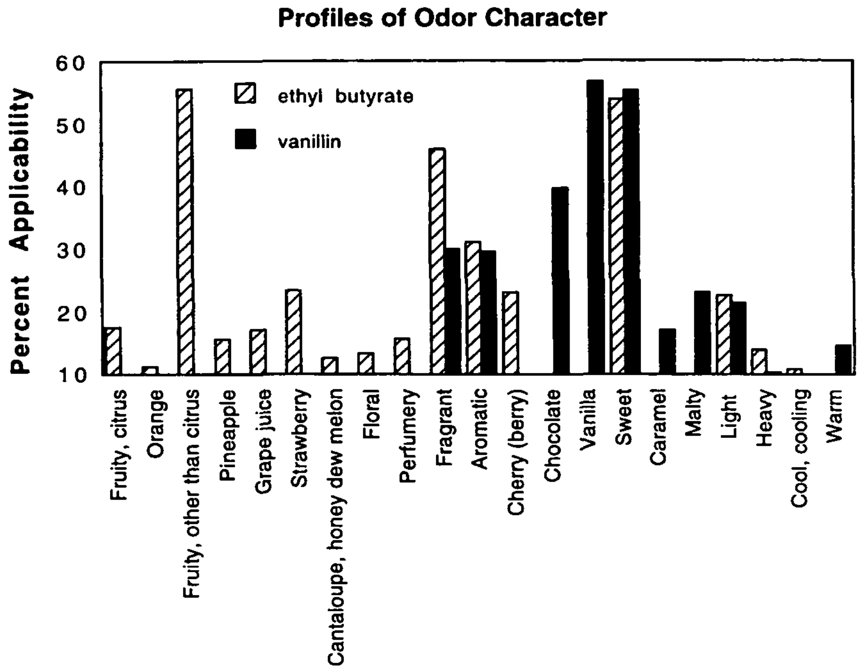


Figure 1 Odor profiles for tutti-frutti and vanillin expressed in terms of percent applicability by Dravnieks's (1985) formula for values above 10% applicability.

and quality discrimination; and (v) odor-specific versus general practice for improvement of identification.

Fluctuations in performance

If asked whether they could identify by smell such items as banana, chocolate, lemon and cinnamon most people would probably say yes (Cain, 1982a). When asked to do so, however, most would identify somewhat fewer than half (Desor and Beauchamp, 1974; Cain, 1979; Cain, 1982a). Errors would comprise near misses, such as calling garlic by the name onion, and far misses, such as calling garlic by the name mustard (specific, but wrong) or by the name spice (too generic). Subjective experience during the task commonly suggests that failures to identify might prove temporary. The name of the item might seem almost accessible and hence likely to become available at another time (Lawless and Engen, 1977). Do such subjective feelings reflect the true state of affairs? The first study addressed the issue.

The issue hardly pertains only to odors. As Bahrick and Hall (1991) noted: 'Access to large portions of our knowledge is unstable, so that recall fluctuates depending upon momentary context' (p. 1). Research outside olfaction has dealt with stabilization of 'unstable access' of previously acquired material, e.g. foreign vocabulary, the analogue of the issue in olfaction.

Attribute rating as a key to identification

Dravnieks (1982, 1985) devised a general way to describe what he called odor character (Cain, 1988). The scheme requires ratings of odor samples on a scale of 0 to 5 against 146 attributes (e.g. fragrant, sweaty, almond-like, burnt–smoky) chosen by a process that began with the compilation of 800 odor-relevant names followed by the elimination of attributes that seemed redundant or too obscure for general use. To illustrate the outcome of such profiling, Figure 1 depicts profiles for the odors ethyl butyrate (the tutti-frutti odor or Juicy Fruit Gum flavor) and vanillin (vanilla extract odor) from data gathered by Dravnieks (1985). (Table 1 shows an alphabetical list of the descriptors.) Whereas the fruity-smelling substance appropriately had representation among many fruit-relevant descriptors and such general descriptors as fragrant, aromatic, sweet and light, the smell of vanillin had representation among confection-relevant descriptors, including vanilla itself, chocolate candy, which contains vanilla flavoring, and general descriptors. Because odor profiling requires subjects to approach odors more analytically and perhaps to process them more deeply (Craik and Lockhart, 1972) than the subjects might do spontaneously, it would seem a possible way to precipitate actual identification. This matter received attention in the second study.

Whereas repeated attempts at identification have led to incremental stabilization of unstable access, intervention in which an experimenter supplies new information (e.g. corrective feedback) has usually led to more dramatic increases in performance. Short of offering subjects a list of veridical alternatives as an intervention to stabilize access to acquired knowledge of odors, supplying the list of 146

Table 1 Alphabetical list of descriptors used for profiling (adapted from Dravnieks, 1985)

1 alcohol-like	38 cooked vegetables	75 laurel leaves	112 raw cucumber-like
2 almond-like	39 cool, cooling	76 lavender	113 raw potato-like
3 animal	40 cork-like	77 leather-like	114 rope-like
4 anise (licorice)	41 creosote	78 lemon (fruit)	115 rose-like
5 apple (fruit)	42 crushed-grass	79 light	116 rubbery (new rubber)
6 aromatic	43 crushed-weeds	80 like ammonia	117 sauerkraut-like
7 bakery (fresh bread)	44 dill-like	81 like blood, raw meat	118 seasoning (for meat)
8 banana-like	45 dirty linen-like	82 like burnt paper	119 seminal, sperm-like
9 bark-like, birch bark	46 disinfectant, carbolic	83 like cleaning fluid (carbona)	120 sewer odor
10 bean-like	47 dry, powdery	84 like gasoline, solvent	121 sharp, pungent, acid
11 beery (beer-like)	48 eggy (fresh eggs)	85 like mothballs	122 sickening
12 bitter	49 etherish, anaesthetic	86 malty	123 soapy
13 black pepper-like	50 eucalyptus	87 maple (as in syrup)	124 sooty
14 burnt candle	51 fecal (like manure)	88 meaty (cooked, good)	125 soupy
15 burnt milk	52 fermented (rotten) fruit	89 medicinal	126 sour milk
16 burnt rubber-like	53 fishy	90 metallic	127 sour
17 burnt, smoky	54 floral	91 minty, peppermint	128 spicy
18 buttery (fresh)	55 fragrant	92 molasses	129 stale
19 cadaverous, like dead animal	56 fresh green vegetables	93 mouse-like	130 stale tobacco smoke
20 camphor-like	57 fresh tobacco smoke	94 mushroom-like	131 strawberry-like
21 cantaloupe, honey dew melon	58 fried chicken	95 musk-like	132 sulphidic
22 caramel	59 fruity (citrus)	96 musty, earthy, moldy	133 sweaty
23 caraway	60 fruity (other)	97 nail polish remover	134 sweet
24 cardboard-like	61 garlic, onion	98 nutty (walnut, etc.)	135 tar-like
25 cat-urine-like	62 geranium leaves	99 oak wood, cognac-like	136 tea-leaves-like
26 cedarwood-like	63 grainy (as in grain)	100 oily, fatty	137 turpentine (pine oil)
27 celery	64 grape-juice-like	101 orange (fruit)	138 urine-like
28 chalky	65 grapefruit	102 paint-like	139 vanilla-like
29 cheesy	66 green pepper	103 peach (fruit)	140 varnish
30 chemical	67 hay	104 peanut butter	141 violets
31 cherry (berry)	68 heavy	105 pear (fruit)	142 warm
32 chocolate	69 herbal, green, cut grass	106 perfumery	143 wet paper-like
33 cinnamon	70 honey-like	107 pineapple (fruit)	144 wet wool, wet dog
34 clove-like	71 household gas	108 popcorn	145 woody, resinous
35 coconut-like	72 incense	109 putrid, foul, decayed	146 yeasty
36 coffee-like	73 kerosene	110 raisins	
37 cologne	74 kippery (smoked fish)	111 rancid	

descriptors has some aspect of an intervention to stabilize access.

Feeling of knowing versus recognition

When unable to identify particular odors, people might comment spontaneously that they would recognize the names if allowed to choose from alternatives. Recognition indeed makes lighter demands than identification subjectively and usually leads to better performance (e.g. Tulving, 1983; Schab and Crowder, 1995). Various researchers have shown that people, including children, can make reasonable, if imperfect, judgements of how well they would recognize the answer to a question they failed to answer by recall (e.g. Nelson *et al.*, 1984; Schneider, 1985; Metcalfe, 1986). Stimuli have included, among other types, pictures and text. When given a trivia question, such as to the name of the northernmost point of South America, a person may fail to recall the name but may recognize it from among alternatives. Whether impressions of incipient success presage accurate recognition is demonstrable. With trivia questions, for example, the impressions, indicated by ratings or ranking, have proven themselves reasonably accurate (Metcalfe, 1986). The third study examined whether that would hold true for odors and how it compared with performance on trivia.

In so far as feelings of familiarity indicate relative accessibility of semantic information about odors, then actual recognition should exhibit a significant association with expected performance. In so far as the feelings may represent perceptual vagueness (e.g. perception of a stimulus as, say, a confection, but not exactly which confection), actual recognition may prove unpredicted from expected performance.

Identification versus discrimination

When given corrective feedback over time, subjects can increase their ability to identify common odors to levels compatible with limitations based on discriminability among stimuli (Cain, 1979). The use of odor identification in clinical evaluation implicitly builds upon a presumed association between discriminability of quality and sensitivity of olfaction *per se*. The use assumes that persons with poor sensitivity will discriminate less well and will therefore identify fewer odors (de Wijk and Cain, 1994). Scores in odor identification should therefore reflect the condition of the sensory modality *per se*.

The correlation coefficients (r) between odor identification and the absolute threshold were in the range –0.8 to –0.9 when applied to groups of both normals and persons with olfactory disorders (Doty *et al.*, 1984; Cain and Rabin, 1989). Such an outcome legitimizes the use of odor identification in the clinic. In a comparison between identification of the real-world items in the test administered at the Connecticut Chemosensory Clinical Research Center and quality discrimination among various chemicals, Eskenazi *et al.* (1983) found coefficients (r) of –0.5 for normals and –0.8 for patients with temporal lobe damage.

Clinical tests of odor identification provide alternatives to the subject and therefore actually tap a type of recognition. An association between true identification and discrimination also holds interest for how it may implicate discrimination in the process of identification. The fourth study explored that association within the range of performance available among young persons with normal smell.

By an intervention of corrective feedback, the study permitted inspection of how discriminability related to identification in conditions of unstable access, i.e. prior to feedback, and more stable access, i.e. after feedback.

Odor-specific versus general practice

What, besides repeated bouts of recall and reference to a list of attributes, might enhance odor identification? It seems to take surprisingly little corrective feedback to raise performance for everyday materials to high levels (Cain, 1979; Schemper *et al.*, 1981), though studies have usually not pitted the performance of subjects who have received feedback, with its benefit of stabilized access, against those who do not. The fifth study fills this gap while it examines possible transfer of training. Any transfer from items in a 'feedback' set to others would open possibilities for how to train persons to identify. Training with some odors could then reap benefits for identification of others. The phenomenon known traditionally as transfer of training has seen little attention in recent years, perhaps because its effects, even when significant, have generally proved small (Howard, 1995). In the much more active area of generation of expertise, however, transfer of skills has a prominent place based upon much more varied outcomes (Holyoak, 1991).

Study 1: fluctuations in performance

In this study, subjects sought to identify various items over four sessions. The issue was consistency: would subjects identify an odor incorrectly one day but correctly another, and vice versa? Conceivably, subjects could labor under an illusion that any inability to identify everyday odors is temporary. If the impressions derive from illusion, then errors once made would in fact virtually always recur, as would correct responses. If inconsistency proves true, then its reason merits attention.

In recall of nonolfactory material, successful recall adds some stabilization to the process of retrieval, such that performance improves progressively. Bahrick and Hall (1991) viewed repeated bouts of recall as a mechanism for preventive maintenance of retention of naturally acquired material over the life span. Hence, with each bout of recall one performs some preventive maintenance against failure to access the material in the next bout.

Method

Subjects

Thirteen young adults (four males and nine females) participated in four sessions separated, one from the next, by an average of 2 days. Few subjects had any experience in odor experiments. The subjects received $5 per session.

Procedure

The task entailed verbal identification, without feedback, of 26 odors per session, presented in irregular sequence from 6 oz opaque plastic jars. Gauze covered the items to preclude visual identification. Fifteen items appeared in random sequence scattered among 11 others that changed completely from one session to the next. No two subjects had the same 11 variable items in the same session. The 15 fixed items comprised almond extract, apple, chocolate, cigar butt, cough syrup, crayons (Crayons), ketchup, disinfectant (Lysol), mustard, dehydrated onion, orange, pencil shavings, soy sauce, tea leaves and turpentine. The remaining 11 in each session came from a pool of 44 (e.g. baby powder, beer, cinnamon, garlic powder). (Prior to a session, the experimenter smelled the items and replaced them as necessary.) Variation in composition was intended to disguise interest in day-to-day performance for the fixed set of 15. As the subjects sought to identify the odors with specific labels, they rated confidence in the correctness of their responses on a five-point scale, where 1 represented very low confidence and 5 very high confidence. If entirely at a loss for a name, subjects could so indicate by the response 'no idea' and, by definition, the rated confidence equaled 1.

Results

For the fixed set of 15 items, the subjects began with 42% correct identification in the first session and increased to

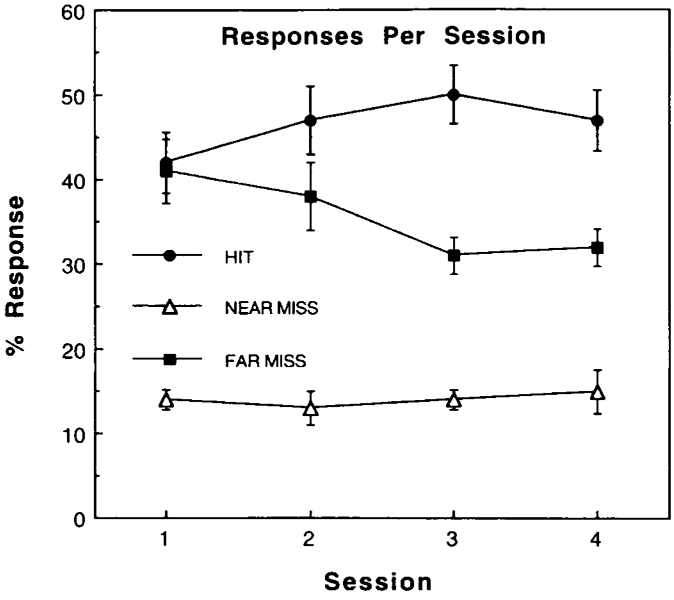


Figure 2 Percentage (± SE) of correct names (hits) and incorrect names (near and far misses) across four sessions.

47% by the fourth [$F(3,36)=2.80$, $P = 0.055$] (Figure 2). They named about one-third (31%) of items correctly in all sessions, another third (33%) correctly in some though not in others and the final third (37%) incorrectly in all sessions. Performance on the odors that appeared just once per subject, i.e. fillers, went from 40% correct in the first session to 36, 36 and 34% in the second through fourth sessions respectively.

The dendrogram in Figure 3 shows the sequence of correct responses (hits) and incorrect responses (near and far misses) for the fixed odors across sessions. Examples of near misses (15% in session 1, with essentially no change thereafter) included spearmint for peppermint, lighter fluid for turpentine and grapefruit for orange. Far misses, which comprised both specifically wrong answers (e.g. mothballs for cigar butts, pepper for garlic powder, honey cough syrup for potato chips) and generic answers (e.g. air freshener for tea leaves, spice for almond extract, shampoo for rubbing alcohol), accounted for 42% of the responses in the first session, but declined eventually to 32% (Figure 2).

Responses given in the first session predicted subsequent responses reasonably well, if imperfectly. A hit in the first session led to an 88% chance of a hit on that item in the next session. A miss led to an 81% chance of a miss on that item in the next session. Stated obversely, 12% of those items hit on the first day were missed on the second and 19% of those missed on the first day were hit on the second. This trend continued throughout and largely accounted for the net improvement over the sessions. The conditional probabilities that subjects would emit a given type of correct or incorrect response after one, two or three previous such

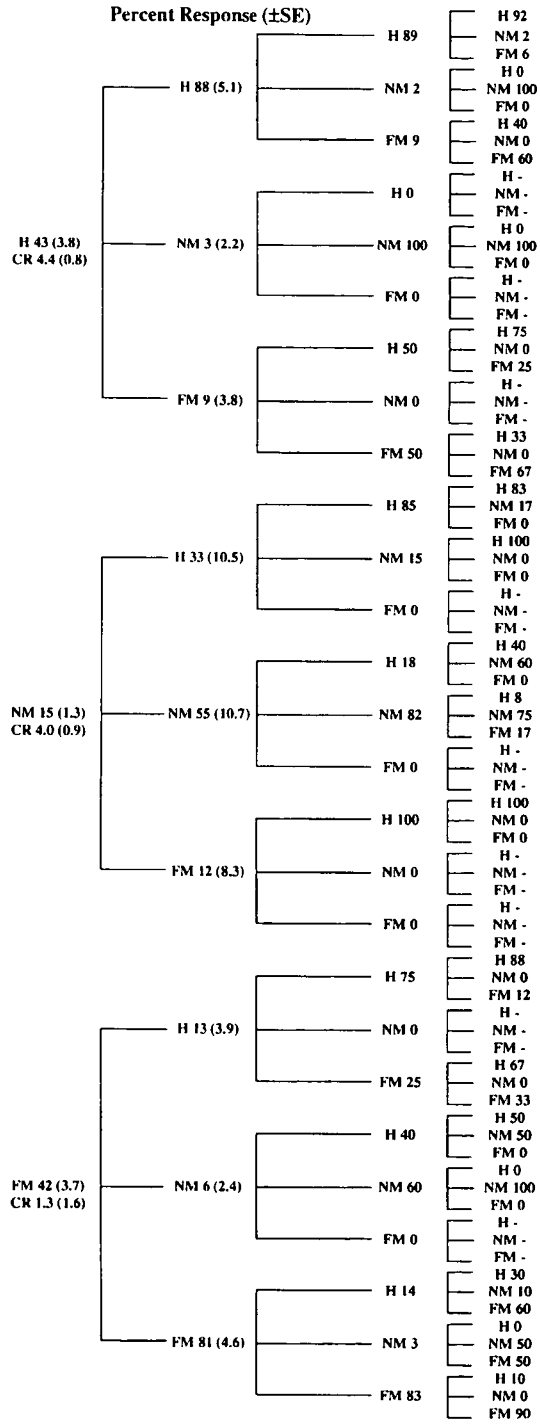


Figure 3 Dendrogram showing the percentage (± SE) of hits, near misses and far misses session by session. To illustrate with the topmost branch: the hit rate (i.e. correct identification) in session 1 equaled 43%, and of those items hit in session 1, the hit rate in session 2 equaled 88%. Of those items hit in session 2, the hit rate in session 3 equaled 89%. Of those items hit in session 3, the hit rate in session 4 equaled 92%. Regarding the lowermost branch of each bifurcation, the far misses in session 1 equaled 42%. Of those missed in this way in session 1, the far miss rate in session 2 equaled 81%. And so on. The pattern shown throughout the dendrogram is one of relative, though imperfect, stability. Confidence ratings of the judgements (scale of 0–5) appear for session 1.

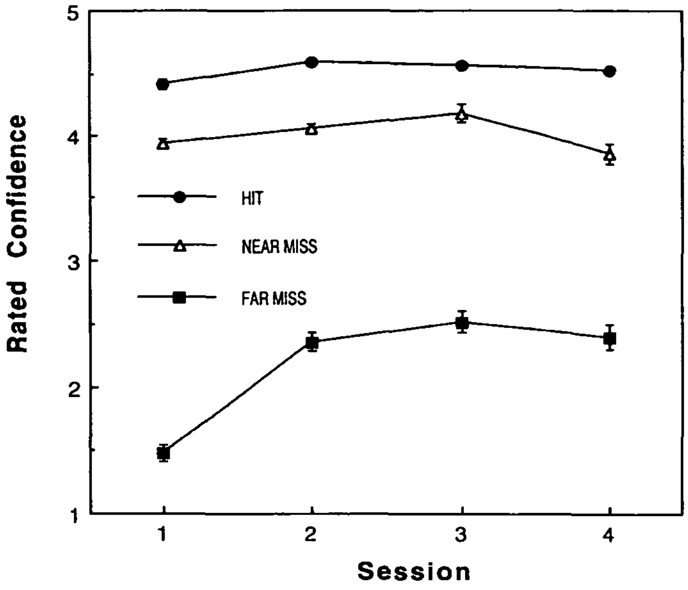


Figure 4 Average confidence ratings (± SE) for hits and misses. The error bars are smaller than the symbols in some cases.

responses was ~80–90%, with a tendency to increase as the string of similar responses lengthened.

Figure 3 also shows average confidence ratings associated with the hits and misses in the identification task (for clarity only session 1 is shown, but see Figure 4). Three features stand out: (i) subjects showed high confidence for hits and low confidence for far misses; (ii) confidence for near misses fell closer to that for hits than to that for far misses, i.e. subjects apparently thought they had smelled the item named; and (iii) variability of the confidence ratings for far misses were about twice as high as that for hits. That is, confidence ratings for far misses included some high values, whereas those for hits almost never included low values. After session 1, confidence ratings associated with each level of performance stayed about the same (Figure 4).

Discussion

The impression that an odor may prove unidentifiable today but identifiable tomorrow, and vice versa, received ample support. In daily life, the failures of today presumably have more subjective weight than successes. People have little reason to ponder whether items they identified today may prove unidentifiable tomorrow. The present data indicate that success follows success somewhat more frequently than failure follows failure, which accounts for the modest net gain in performance. Others have shown that recall of the same body of material will vary from one instance of recall to another (Underwood, 1983). A gain in performance via repeated instances shows the stabilizing influence of mere retrieval (Bahrick and Hall, 1991). Such findings support the point of view that what we call forgetting may actually represent failures of retrieval [see Cain and Potts (1996) for a discussion of this regarding odor memory].

What mattered considerably here was that improvement in identification took place without feedback, an instance of what can be called hypermnesia, and which Bahrick and Hall (1991) referred to as 'preventive access maintenance'. A second attempt to recall a name after a failure carries various potential, if somewhat minor, advantages, including a second sampling of the stimulus, which could make more 'features' available for synthesis into a configuration, and memory of the previous response, with retention of it if it seemed right and elimination of it if it seemed wrong. The second point requires some appreciation of the correctness of responses. The finding that subjects could express a level of confidence roughly commensurate with accuracy provides a key element.

The basis for the typical positive relationship between confidence and accuracy seems to lie in feelings of familiarity (Clark, 1997), and this may play a defining role in the persistence of correct responses. Subjects presumably tend to remember from one opportunity (session) to the next a label that gave a feeling of high confidence which would help among other things to define the set of responses for the next session. When seen strictly as a task of information transmission, odor identification should be limited by discriminability among stimuli, and specification of and discriminability among responses (see Desor and Beauchamp, 1974, Garner, 1962). [Of the various clinical tests of olfaction, only Wright's (1987) defined itself in terms of information transmission.] An information analysis does not concern itself with where the responses come from, whether from memory or from the experimenter. When the task does not include a predefined set of responses, suboptimal performance could just as well stem from the existence of too many possibilities, as well as too few. The memory of a response that leads to a judgement of high confidence could therefore reduce the set of potential responses for the next session.

When seen as a task of memory as opposed to a task of information transmission, the question of where the response of identification comes from, irrespective of stimulus or response discriminability, tends to define the issue. We will return to the issue of memory in the General discussion.

Study 2: profiling as a means to prompt identification

Sessions of this study followed those of Study 1 by about an hour on the same days. Odorants that served as fillers (distracters) in Study 1 became the focus of attention regarding whether profiling via the technique of the American Society of Testing and Materials (Dravnieks, 1985) might precipitate identification. Rabin (1988) used the task of profiling as a way to cause subjects to process odors deliberately and perhaps more deeply than merely smelling them (Craik and Lockhart, 1972). Profiling caused some increase in how well subjects discriminated between the profiled odors, though learning their names (labeling) caused a bigger increment in discrimination (see also de Wijk and Cain, 1994). This suggested that profiling actually increased the semantic information available to the subject, though it did not show that profiling could precipitate identification.

If subjects improve their performance from one opportunity to the next because they sample and synthesize progressively more features, then profiling would seem a suitable task to accelerate progress for it might index the features. Viewed in terms of information analysis, if subjects could improve their performance by a reduction of possible responses from the theoretical infinite (and hence poorly discriminable) number available in simple naming to a much smaller (and hence less poorly discriminable) though still very large number, profiling might also accelerate progress. Hence, profiling seemed to possess two latent assets: multiple sampling of features in an analytical context and reduction of the set of responses.

Method

Subjects

The 13 subjects of Study 1 also served here. The subjects received $5 per session for four sessions.

Procedure

After a subject had sought to identify all 26 odors in each session of Study 1, the experimenter selected six items from among the 11 fillers in that session according to the rule that two should have received hits, two near misses and two far misses. If the responses failed to allow ideal choices of two of each type of response, particularly near misses and far misses, the experimenter sought the best approximation.

With the six odors chosen, the experimenter selected one randomly from that set and asked the subject to describe it via the list of 146 descriptors arranged in irregular sequence (Dravnieks, 1982). Next to each descriptor, the numbers 0–5 gave the subject the chance to rate its applicability to the test odor. The experimenter encouraged the subject to think about each descriptor and to make a conscious decision about its applicability. There were no limits on how many of the 146 the subject could use. The experimenter did not explicitly ask the subject to use descriptors as a means to identify the odor, but did instruct the person ahead of time to write the name of the odor on the rating sheet when the profiling was finished. The experimenter then presented the next odor, until the subject had profiled and indicated the identity of all six in that session. The same procedure occurred in four sessions of the study.

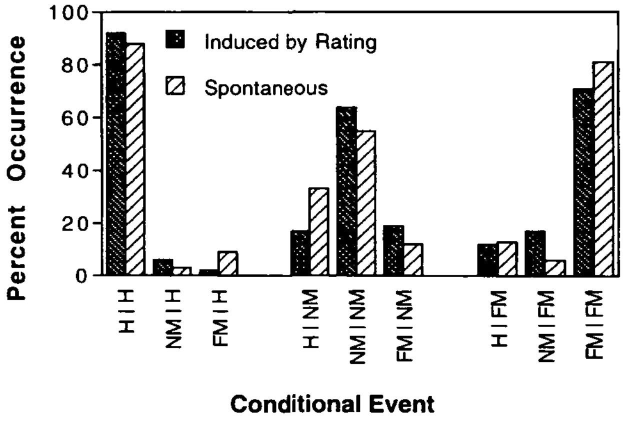


Figure 5 Dark striped bars show the conditional occurrences of identifying odors better or worse after profiling them. The conditional event NM/H, for example, refers to giving a near miss after profiling when having given a hit before profiling. Light striped bars refer, for point of reference, to the responses from the first to the second sessions for the 15 target odorants of Study 1.

Results

The average number of descriptors per item was eight, with essentially no progressive change over sessions (range across sessions: 7.8–8.0). Different subjects, however, characterized themselves by how many descriptors they used, ranging from as few as 1.4 per odorant to 14.8, for subjects from the least to the most descriptivist respectively. By their use of multiple descriptors, even if just a few, the subjects seemed to treat the task in the analytical manner intended. For the odorants used, which varied uniquely from subject to subject, one-quarter had the adjectival forms of their veridical names on the profiling sheets. Such a small proportion illustrates how a list of even 146 descriptors fails to capture many of the odor names of everyday products.

When subjects identified odorants correctly (hit) before profiling, they rarely changed their responses afterwards (Figure 5). This outcome held no surprise, for there would seem little reason for the list of descriptors to throw subjects off the correct trail. Insofar as any deterioration occurred, however, it apparently had nothing to do with the descriptors since essentially the same deterioration occurred from the first to the second session of Study 1.

When subjects had given near misses before profiling, they also tended to do so afterwards, and they showed no greater tendency to improve their responses than to degrade them. When the subjects had given far misses before profiling, they again tended to do so afterwards, and any improvement came very close to that seen spontaneously. Net improvement in quality of label, i.e. H|NM, H|FM and NM|FM, equaled 46 and 52% in the induced and spontaneous cases, respectively, and net deterioration in quality, i.e. NM|H, FM|H and FM|NM, equaled 20 and 23% respectively. Hence, general profiling not specifically geared to the odorants under study apparently afforded no advantage to precipitate identification.

Discussion

As indicated, profiling could potentially have facilitated identification in various ways. It could have guided a person to conscious perception of aspects that might otherwise have gone unnoticed. If imperfect performance in odor identification derived principally from failure to access relevant odor names in memory, profiling might have facilitated [stabilized] retrieval and allowed for the performance seen when subjects received feedback with veridical names (Cain, 1979; Cain and Krause, 1979). Even if the exact name of a stimulus failed to appear on the profiling sheet, the list could have functioned similarly to a police Identikit used to create likenesses and possible identification via 'features' of a person's face. [Cain and Gent (1986) noted similarities between odor perception and face perception regarding discretionary holistic-analytic processing (see Carey and Diamond, 1977).] That is, profiling could lead to the accumulation of features that, at a less conscious level, could contribute to an identifiable configuration. It seemed to do none of these things.

Subjects had not, however, been told to use the sheets to precipitate identification. If this weakens the test that profiling might precipitate identification, it would seem to do so specifically regarding the possible use of the list to retrieve labels and thereby to use the exercise as a huge multiple-choice task. From the perspective of information transmission (Garner, 1962), the optimal number of choices should equal a number consistent with the discriminability of the stimuli. In principle, too many choices could prove as unproductive as too few. This would at most mitigate only slightly the general point that the depth and type of processing induced spontaneously by profiling seems to have little influence on identification.

Study 3: feeling of knowing

In Study 1, subjects demonstrated via their confidence judgements that, within limits, they knew whether they had given a correct answer. When subjects gave veridical labels, they generally did so with almost perfect confidence. When they gave near misses, they gave lower but still rather high ratings. The outcome suggested that near misses commonly reflected errors of either perceptual discrimination or of response discrimination. When they gave far misses, however, they seemed in general to know that the answers were incorrect. These would seem more likely to include errors of retrieval.

Study 3 directly addressed 'feeling of knowing' regarding items that subjects fail to identify. In a common investigation into this aspect of metacognition, subjects go through various items (e.g. trivia questions) and accumulate a specified number of unknowns, and then indicate whether they might recognize the answer from among alternatives. Insofar as that procedure focuses upon failures to know an answer outright, it deals with a foreshortened version of the continuum of knowing what one knows. That is, it omits the hits and, by the protocol used here, many of the near misses and those far misses in which the subjects had confidence. Furthermore, it deals with anticipated rather than past or present performance. In this respect, it differs from the kind of information transmitted by the confidence rating of Study 1.

Metcalfe (1986) found that people could anticipate recognition of the answers to trivia questions but not their recognition of the answers to 'insight' problems or brain teasers that involved reorganization of information. Both trivia problems and odors seem sometimes to induce a tip-of-the-tongue state, whereby the answer seems almost, but not quite, available [Lawless and Engen (1977) drew a distinction between the tip-of-the-tongue state versus the tip-of-the-nose state]. We might therefore expect a reasonable predictive performance in both tasks. Trivia problems, though, offer a much richer context upon which to base a prediction. For example, subjects who know a great deal of geography may trust that they will recognize answers on that topic even if the answers hardly seem on the tip of the tongue (see Koriat, 1995).

Method

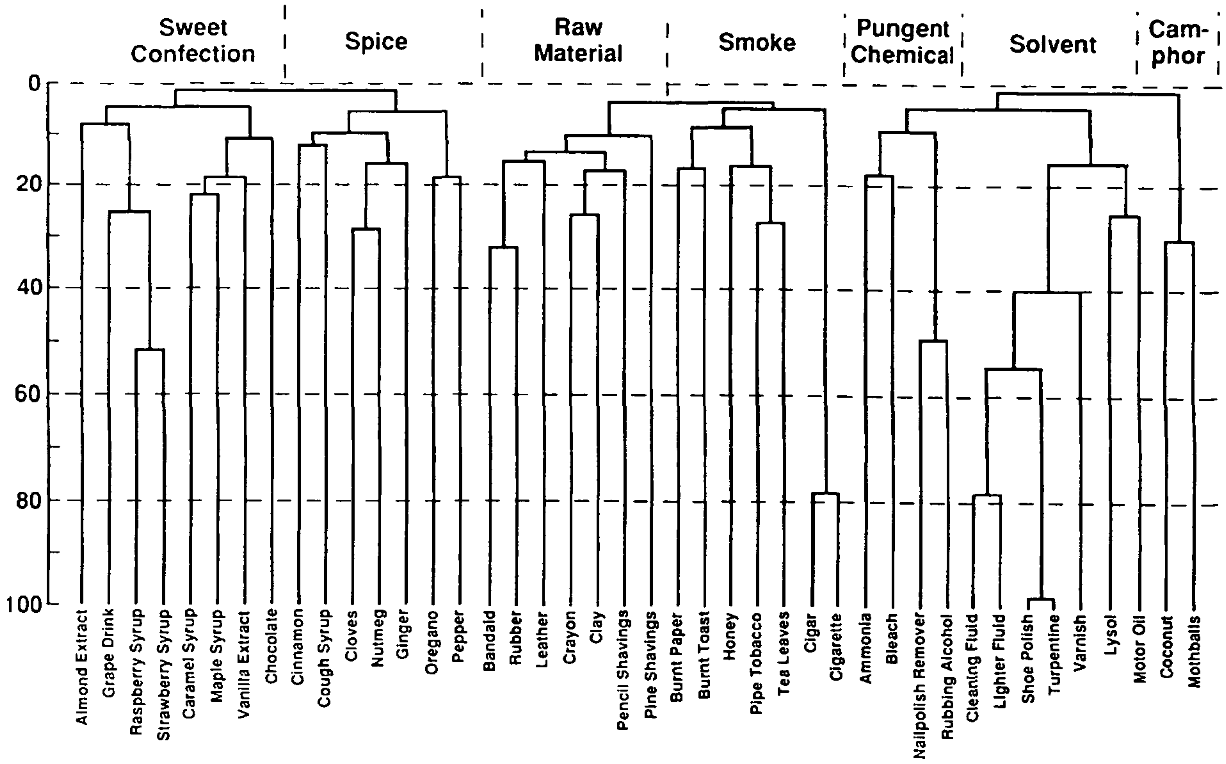
Subjects

Seventy-eight young adults (39 males and 39 females) participated without pay. None had previously served in odor experiments.

Stimuli

The stimuli included 49 common odors and 40 trivia questions. Examples of the trivia questions are: (a) Who was Helen Keller's teacher? (b) What renowned sour mash whiskey is made in Lynchburg, Tennessee? (c) What general first commanded United Nations' Forces in the Korean War? and (d)What is the number one killer in industrialized countries?

For the recognition part of the experiment, where subjects had to choose a correct answer from among choices, we offered eight choices that seemed, as a set, to make the item reasonably difficult. For the trivia question (a), above, the choices were: Annie Hall, herself, Florence Nightingale, Anne Franklin, Sybil Luddington, Anne Baxter, Harriet Tubman and Anne Sullivan. For (b), the choices were: Old Grandad, Seagrams, moonshine, Southern Comfort, Wild Turkey, Canadian Club, Crown Royal and Jack Daniels. For the odor choices, we used data previously collected in this laboratory on errors of odor identification. In order to make these choices reasonably difficult, we selected items that generally fell within the same cluster or adjacent clusters in a hierarchical cluster analysis of the data (Hartigan, 1977). The solution was generated from an 80-by-80 confusion



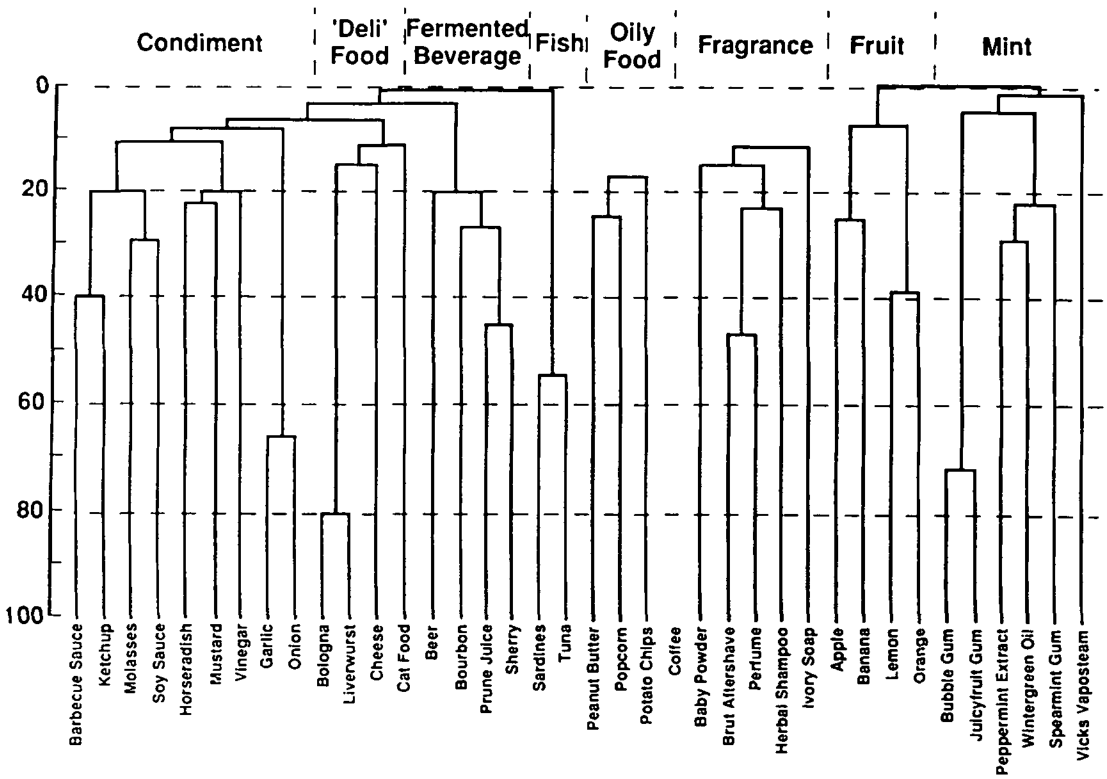


Figure 6 Hierarchical cluster analysis of 80 odorants via Hartigan's (1977) method of average linkage. The hierarchical tree is scaled from 0 to 100, with higher values at the point of clustering corresponding to greater similarity.

matrix of 82 subjects each of whom had contributed 3–5 responses per item. Some subjects had smelled 40 odorants and some 80. We show the solution in Figure 6 since it may prove beneficial to other experimenters who may wish to create multiple-choice responses for odor recognition. To illustrate, choices for pencil shavings, for example, included rubber, leather, Band-Aid, burnt paper, clay, pencil shavings, Lysol [disinfectant] and burnt toast.

Procedure

The first part of a session entailed screening odors (trivia questions) for seven that a subject could not name (answer). Subjects administered the task to themselves. Arbitrarily numbered jars of odorants sat on a large tray placed in front of the subject, as did a stack of cards, numbered on the back and placed face down. Each card contained a question.

Half the subjects sampled the odors first and half the questions first. A subject would smell an odor (read a question) chosen arbitrarily from the set and would indicate on a response sheet whether he/she knew its identity (could answer the question). If unable to name (answer) it, the subject put it aside for subsequent use. This process continued until the subject had accumulated seven such items. Then the subject ranked the items from least to most likely to be recognized. Thereafter, the subject scrambled the items, sampled each again and rated them on a visual analog scale marked at intervals from 0 (Definitely will not get the answer.) to 10 (Definitely will get the answer.).

After making the ratings, the subject sought to recognize the correct answer from among eight names printed on cards. The subject would re-smell (re-read) the odor (trivia question) before making a choice, would record the choice and would then proceed to the next item of the seven.

Results

The trivia questions and the odors seemed well matched for difficulty. For example, it took on average 16.4 trials to obtain seven unanswerable trivia questions and 16.2 trials to obtain seven unidentifiable odors. In the recognition task, subjects obtained correct answers for 3.4 (SD = 1.4) out of seven trivia questions and for 3.6 (SD = 1.7) out of seven odors. Hence, use of recognition improved performance considerably over recall. The average familiarity ratings for the trivia ranged from 2.6 to 7.8 for the items ranked first through seventh, respectively, in recognition ability and 2.9 to 7.3 for the comparable odors. Recognition performance varied from 15/40 to 26/40 for the first- through the seventh-ranked trivia items and from 19 to 24 for the first- through the seventh-ranked odors. The correlation between rated familiarity and recognition equaled 0.91 for trivia ($P < 0.01$) and 0.61 (n.s.) for odors. A test of the slope of the linear regressions supplemented the correlation coefficient as a measure of the relationship between familiarity and recognition. The slope of the relationship for trivia, 1.82, reached significance [$t(5) = 3.55$, $P < 0.025$], but that for odors, 0.76, failed to [$t(5) = 1.79$].

Nelson (1984) concluded that the Goodman–Kruskal gamma test provides the best measure of feeling of knowing. The index, like other measures of correlation, varies between +1.0 and –1.0. Its value depends upon the number of concordances and discordances in a set. When applied to the rank order of predicted recognition versus concordances for individual subjects, gamma averaged 0.07 [$t(76) = 1.17$, n.s., when tested for significance of difference from zero] for odor and 0.19 [$t(76) = 3.17$, $P < 0.01$] for trivia. Gammas based on aggregate data equaled 0.60 for odors and 0.89 for trivia, almost the same as the Pearson coefficients. Hence, irrespective of the means of examination of the data, the outcome suggested significant predictive validity of feelings of knowing for trivia but not for odor.

Discussion

The data indicated that people will often recognize odors they feel they could, but that the magnitude of their feeling may not correspond with the reality of their performance. Although feelings of knowing often do correspond to performance, investigators have shown for non-olfactory material that familiarity can vary somewhat independently of ability to identify. Whittlesea *et al.* (1990), for example, increased feelings of familiarity by manipulation of what they termed perceptual fluency. In the simplest case, a primed stimulus may give a feeling of high familiarity even if semantic information about it may not become any more accessible via the priming.

As in the case of profiling, this study gives just a single view of the phenomenon of interest. The performance at recognition could depend on the ease or difficulty of the choices given. Subjects did not know ahead of time how challenging the choices would be, whether for odors or for trivia. The attempt to make the tasks about equally difficult succeeded with respect to average difficulty, and in this regard we can feel some confidence in the conclusion that subjects can anticipate their recognition of answers for trivia better than their answers for odors. Although people may feel in the tip-of-the-nose state, they apparently do little better when feeling that way than when feeling less certain, if we assume that the tip-of-the-nose state anticipates easy recognition.

Study 4: identification versus discrimination

The vaunted ability of the sense of smell to discriminate differences in quality has received surprisingly little empirical attention (see Cain, 1988). The nose of some individuals can undoubtedly discriminate subtle nuances between chemicals, but the ordinary person may show considerably poorer ability. Eskenazi *et al.* (1983, 1986), Rabin (1988) and Martinez *et al.* (1993) found that discrimination among members of small sets of diverse odors failed surprisingly often. The possibility of such failures permits examination of the question of whether discrimination and identification may rise and fall together among people without professed problems with smell. In the present experiment, subjects sought to discriminate among a small set of unfamiliar odors and sought to identify a large set of familiar odors.

Method

Subjects

Eighteen young adults (10 females, eight males) participated in two sessions. The subjects received $5 per session.

Procedure

The first session involved two rounds of odor identification of common substances like those used above and the second involved quality discrimination of 10 generally unfamiliar chemicals. In the first session, subjects received corrective feedback during one round of presentations of 42 substances and no feedback during a second round of presentations of a subset of 32 substances. The 10 substances discarded had previously proven themselves rather easy to identify and their absence from the set served both to make the second round of identification more difficult and to circumvent identification via a process of elimination.

In the second session, subjects sought to decide which member in each of 90 triads of two identical odors and one different odor comprised the odd one. A pilot study had established concentrations (by volume in diethyl phthalate) of approximately equal perceived intensity as follows: *n*-propyl alcohol, 0.5%; *n*-butyl alcohol, 0.5%; *n*-octyl alcohol, 0.15%; *n*-decyl alcohol, 0.3%; linalyl acetate, 0.5%; *n*-amyl butyrate, 0.15%; acetophenone, 0.037%; 3-heptanone, 0.075%; methyl *n*-propionate, 0.125%; and cinnamyl *n*-butyrate, 0.5%. Subjects smelled the triads in irregular order from 8 oz plastic squeeze-bottles.

Results

The subjects identified 38 ± 7% (mean ± SD) of the items on the first round and improved to 59 ± 13% on the second. The large increase in performance presumably came from the corrective feedback since, as Studies 1 and 2 showed and the next one will show more definitively, subjects make only modest progress without such feedback. If we restrict attention only to those 32 items presented on both the first and second rounds of identification, the results look much the same: 31 ± 9% on round 1 and 59 ± 13% on round 2.

Discrimination equaled 40 ± 17% after correction for chance performance of 33%. Hence, subjects made plenty of errors. The correlation between the performance of individuals at discrimination and their performance in the first and second rounds of identification, respectively, equaled −0.09 (n.s.) and 0.61 ($P < 0.01$) (Figure 7). (The absence of a significant correlation in round 1 held true irrespective of whether the correlation was computed on all 42 items or on the 32 items, where it equaled 0.12.) That is, a significant association between how a person performed on discrimination and how that person performed on identification emerged after the first round of identification.

Interestingly, the correlation between initial performance at identification and the increase in performance from round 1 to 2 equaled a significant −0.52 ($P < 0.05$). That is, subjects who performed more poorly on identification in round 1 tended to show more improvement with feedback than those who performed better initially. Such an outcome, apparently not an artifact of a ceiling since performance still fell well below perfection, underscores the importance of feedback in the uncovering of the correlation.

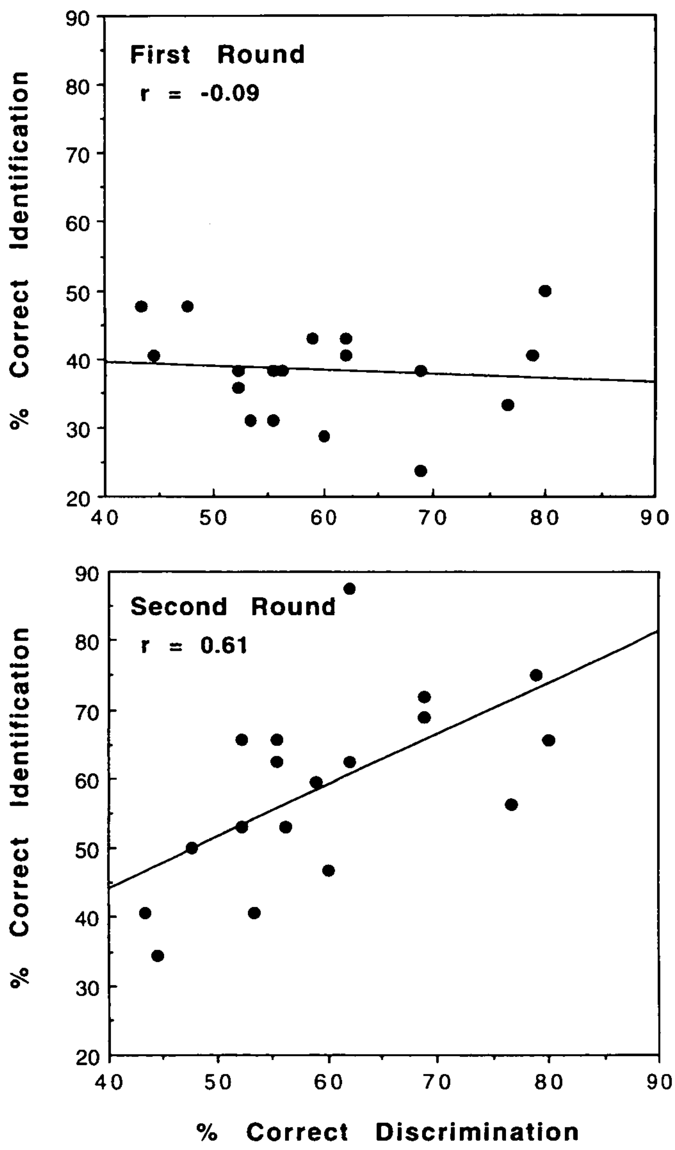


Figure 7 Relations between identification and discrimination for the first and second rounds of identification.

Discussion

As indicated in the Introduction, the correlation between performance in discrimination and identification seen previously had occurred for 'cued identification'. The present results indicate the importance of defining permissible responses, either via multiple choice or by feedback. These presumably serve to stabilize access to the available semantic information. Viewed another way, however, they turn the task of identification into one of recognition. Without definition of the permissible responses, unstable access to semantic information presumably largely governs performance at identification. With such definition, differences among normal young people emerge. de Wijk *et al.* (1994) also found stable individual differences in quality discrimination and concluded that such differences existed as a talent separate from threshold sensitivity in normal subjects. Previous studies of neurologically damaged subjects had suggested the same (e.g. Eskenazi *et al.*, 1986; Martinez *et al.*, 1993). Clearly, then, the factors that control access to semantic information deserve attention in the quest to understand the discriminative basis of odor identification.

Study 5: odor-specific versus general training

In the various studies that have given subjects feedback with labels, the subjects have received the feedback on all items. This leaves open the question of whether subjects learn about odors only one at a time or whether feedback might sharpen general ability. This question has received no attention, yet it would seem to merit it if only to exclude the possibility that some feedback would generalize to a much larger set.

Correlations such as that between discrimination and identification can occur only because of reliable differences among people. The previous study implied that individual differences of sensory (i.e. discriminative) relevance in identification might only manifest themselves after feedback, but it did not specify what feedback actually accomplished. Does it sharpen discrimination, as Rabin (1988) found (see also de Wijk and Cain, 1994)? If so, then feedback might transfer beyond odors in the feedback set and, if that also proved true, transfer of improvement would have important implications for how to train people about odors.

Method

Subjects

Thirty-two young adults (16 females, 16 males) participated in five sessions each. They received $5 per session.

Procedure

In their first session, all subjects sought to identify 40 common odors. Thereafter, half the subjects joined a training group and half a control group. For both, the answers given in the first session allowed a division of the 40 items into target and nontarget sets of 20 each. The experimenter sought to make these sets comparable for each individual. If a subject identified 20 of the 40, then 10 were assigned to each set. If the subject also achieved 10 near misses, then five were assigned to each set, and so on. For the control group, the designation of odors into target and nontarget sets had no subsequent functional meaning.

For the training group, the three subsequent sessions entailed presentation of all 40 items, with corrective feedback for the 20 items in the target set and none for the 20 items in the nontarget set. For the control group, the three subsequent sessions entailed participation in an odor task of intensity perception, i.e. they judged the perceived intensity of chemical mixtures, with no relevance to odor identification. In a fifth session, all subjects sought to identify the 40 odorants again. The questions of principal interest concerned whether subjects would improve on the training set and whether any gain on the training set would generalize to the 20 accompanying odors for which no training was given.

Results

Figure 8 illustrates that subjects who received training on half the odors increased their performance on those markedly from the first session (44%) to the fifth (81%). An ANOVA revealed a significant effect of session [$F(1,15) = 245$, $P < 0.0001$], of training [$F(1,15) = 66$, $P < 0.0001$], and of the interaction training by session [$F(1,15) = 90$, $P < 0.0001$]. The group that received neither training nor interpolated exposure to the items in sessions 2 through 4

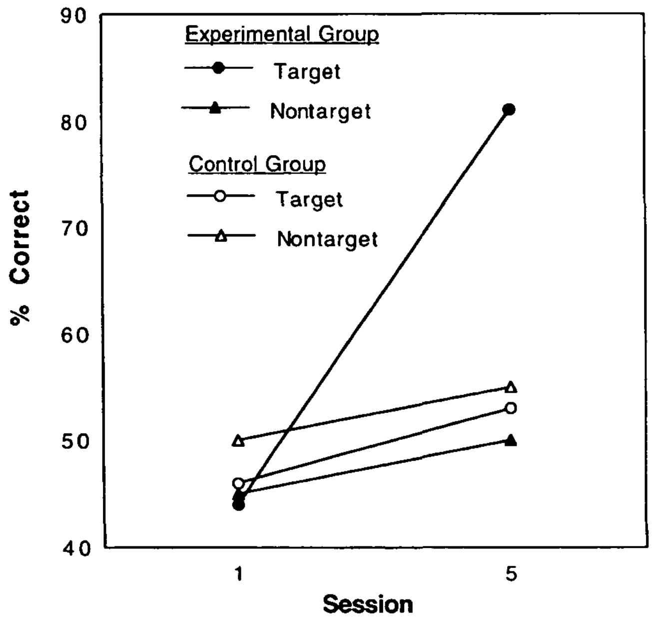


Figure 8 Identification of target and nontarget odors in test subjects before and after three sessions of feedback and in control subjects who received no feedback.

also showed a significant effect of session (48% in session 1 to 54% in session 5) [$F(1,15) = 6.3$, $P < 0.03$]. Hence, as in Study 1, some spontaneous improvement occurred from one presentation of an identification set to another, despite the absence of interpolated experience with the set. Nevertheless, for this group of subjects, performance did not differ between the arbitrarily designated target and nontarget sets [$F(1,15) = 2.3$, n.s.], nor did the interaction of odor-set by session approach significance [$F(1,15) = 0.4$, n.s.]. The parallel course of the increase in performance from sessions 1 through 5 among those 20-odor sets for which no feedback was given underscores the absence of any generalization of training from the target to the nontarget set in the subjects who received feedback.

Discussion

This study demonstrated, perhaps for the first time against appropriate controls, how a relatively small amount of training (feedback) can improve identification markedly. The study thereby offered a contrast of what might happen spontaneously and what might happen with help. Some spontaneous improvement occurred, but no transfer.

General discussion

To recapitulate: (i) inability to identify an odor may indeed be temporary; from day to day, unidentifiable odors became identifiable and vice versa, though the shifts gravitated in the direction of improvement rather than deterioration; once a person identified an odor, that odor tended to remain identified. (ii) Profiling, which seemed a priori to have certain virtues to prompt identification, failed to exhibit them; perhaps a change in orientation toward use of profiling specifically to retrieve labels might have increased performance, but this would not alter the point that more analytic or deeper processing of odors brings no incremental advantage over simple smelling. (iii) Subjects felt high confidence in the accuracy of their hits and near misses and considerably less in the accuracy of their far misses; they could, as expected, recognize odors better than they could identify them, but their relative level of confidence regarding how well they would recognize unidentified odors proved poorly prognostic. (iv) Among people with normal smell, differences in ability to identify odors were associated with differences in ability to discriminate among qualities, but only after specification of the response set for identification. (v) The most obvious way to enhance odor identification, namely feedback with veridical names, offered no transfer beyond items in a training set, but subjects in Study 5, as those in Study 1, improved their performance a little even without training.

These various outcomes served to confirm some and deny other expectations regarding identification. The finding that access to semantic memory for odors varied but improved from day to day confirmed expectations, whereas the finding that an unidentified odor of high familiarity would prove no more recognizable than one of lower familiarity denied them. The ability of subjects to rate the veridicality of their answers rather well confirmed general findings outside of olfaction, but has unexplored dimensions of possible strategic importance to odor identification. Tulving (1983) decried the absence of information on this topic generally: 'the fact that usually confidence judgements correlate positively with objective measures of veridicality may have discouraged experimenters and theorists from raising questions about the nature of the relation between subjective and objective veridicality, and conditions and variables related to this relation' (p. 188). In so far as ratings of confidence reflect semantic information, we might use them more effectively than the quality of the labels themselves to probe the amount of information accessible even when subjects fail to emit veridical labels. Retrospective confidence judgements given at the point of attempted identification and prospective confidence judgements given for anticipated recognition (as in Study 3) by the same subjects could offer a more definitive composite picture than the one obtained here. The question concerns how much semantic information subjects can actually access, even if they cannot access a veridical label.

If one says that human beings exhibit poor ability to identify odors, one needs to place it in context. As indicated at the outset, subjects may actually convey considerable semantic information even in far misses. [This term—for which we must take blame (Cain, 1979)—has unfortunate consequences for on the surface it may imply just 'junk' answers.] If subjects give the answer fruit to lemon, they have by our definition made a far miss, though coming this close probably deserves more 'credit', which numerical ratings of confidence could offer, if shown to correlate well with degree of veridicality. To illustrate partial transmission of information, we can look at the errors from past or present studies. For example, Engen's (1987) table 1 lists identification responses given to Johnson's baby powder: correct response (21/48), no response (6), powder (3), soap (3), air freshener, baby oil, baby wipes, Band-Aid, bubble gum, dentist's office, flower, hand lotion, man's perfume, rose, suntan lotion, tissue, toilet paper, vanilla and wax. Most of the wrong answers contained some semantic information relevant to the stimulus. Whether a particular error derived from faulty perception or faulty access to stored information remains uncertain, though some of each seem to appear in the list. (A dichotomy between perception and access undoubtedly has arbitrary characteristics, though it probably retains a certain heuristic merit.)

Results from the present study proved not unlike Engen's, as illustrated in Figure 9. (The profiles in Figure 9 offer more-or-less representative outcomes for the 15 items in Study 1.) Depending upon the stimulus, subjects transmitted considerable information, with errors sometimes seemingly based on perception and sometimes seemingly based on

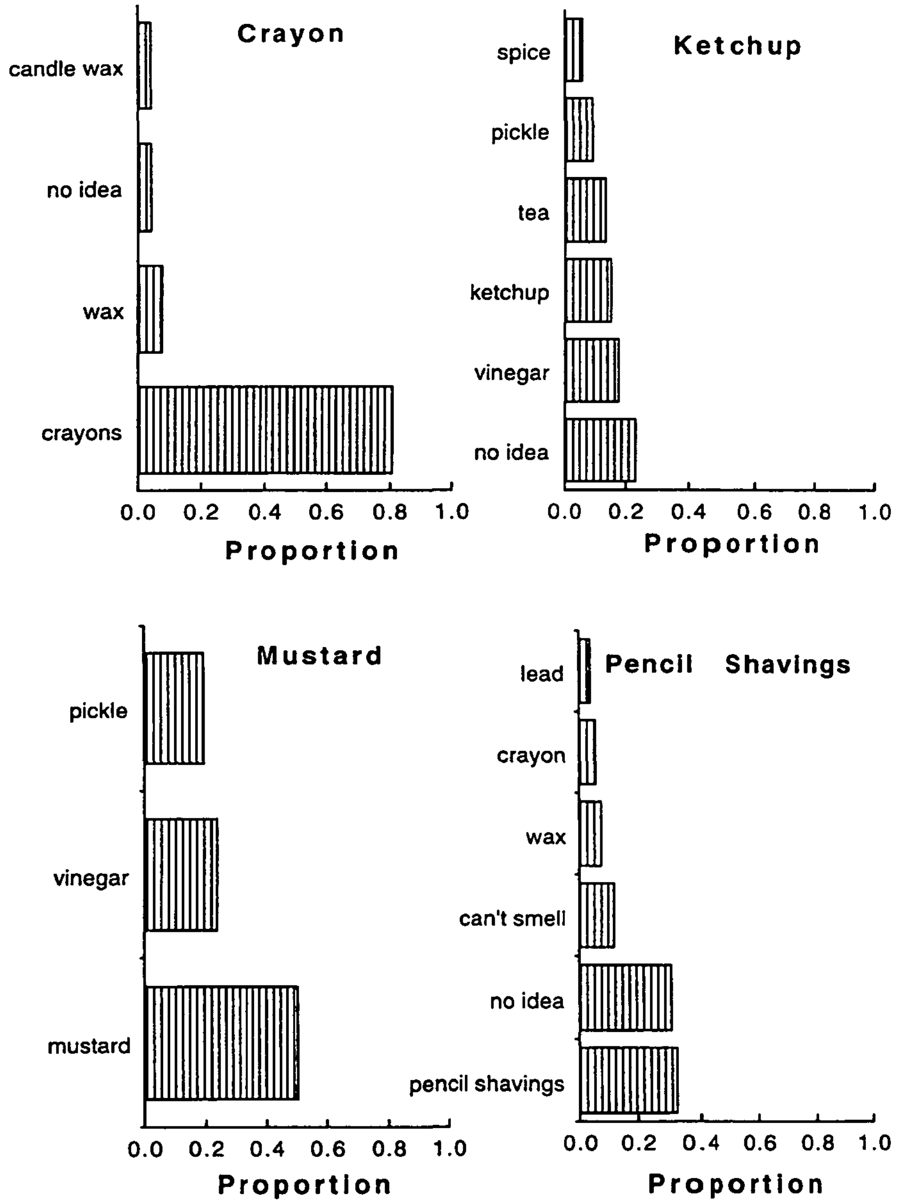


Figure 9 Responses given for four of the items in Study 1. The histograms include responses given on at least 2% of trials. Responses that did not exceed that criterion were as follows. For crayons: clay and chalk. For ketchup: radish, cigarette, raisin, chutney, rose leaves, lemon and mustard. For mustard: relish, curry and dill. For pencil shavings: soap, curry and plastic.

semantic similarity. Here again, there is no easy answer in many instances to whether an error has a semantic or perceptual basis. In the case of mustard, for example, the constituent vinegar forms part of the actual product and when given as an answer presumably represents a perceptual error. Mustard does have a vinegary aroma, but it has more than just that. A similar analysis could hold for an answer of pickle for the product mustard-relish contains both pickles and mustard, which could indicate that the answer has a semantic basis. Pencil shavings would seem to have no particular perceptual similarity to crayons, but does seem to have a semantic similarity via the concept of writing instrument. The stimulus crayon, on the other hand, will probably never draw the response pencil shavings for its semantic network may not extend that far. Ketchup drew a rich variety of responses. Conceivably the complexity of its ingredients or its overlap with the ingredients of other products of its type (sauces and condiments) causes it to be identified poorly in a decontextualized state. Like mustard, ketchup rarely occurs as an olfactory stimulus alone. Perhaps the difficulty of its decontextualized identification reflects how well it blends with the foods it serves to enhance in flavor.

Given that subjects tend to transmit partial information about identity even when they fail to give exactly the right answer, they show a provocative inconsistency over time.

When they make far misses, subjects find it difficult to apply the names consistently, i.e. they may call ketchup mustard one time and spice another (Cain, 1979; Rabin and Cain, 1984; Cain and Potts, 1996)—a reflection of instability of access. This held true in the present Study 1, just as in the past. When a subject gave far misses for the same item in successive sessions, the probability that the subject would give exactly the same far miss label on both occasions equaled only 0.33. When a subject gave near misses for the same item in successive sessions, the probability that the subject would give exactly the same near miss label equaled 0.88. (When a subject gave veridical names on successive sessions, by definition the subject gave exactly the same name.) If we took the name as indicative of the perceptual event, then a stimulus encoded with an incorrect label on any given occasion might well be perceived differently on different occasions. Cain and Potts (1996), in their study of how subjects fell for bait they had themselves created in an episodic memory task, concluded that some such perceptual fluctuation does occur. Within this context, however, it would seem productive to allow subjects to make second or third guesses in order to discover whether they simultaneously entertain more than one hypothesis regarding identity, and hence whether constraints of the experimental protocol simply preclude such information from becoming known.

The path to enhancement of semantic memory for odors seems at the same time complex and simple. Complexity applies in so far as profiling and transfer of training offer so little incremental performance as to make the process of semantic access seem opaque. To the complexity, we can add sluggishness of learning new names for unfamiliar odors. Simplicity applies in so far as the simple route of making labels available works so well. Indeed, trial-by-trial feedback with veridical names leads to such rapid learning that the process seems almost like insight compared with the painfully slow process of paired-associate learning of names to unfamiliar odors (Davis, 1975, 1977; Cain *et al.*, 1995). How long this facilitation of performance, i.e. stabilization, lasts might give some insight into why it occurs. Would, for example, names given at feedback persist beyond the subjects' ability to recall them in free recall?

In so far as some semantic information other than a veridical label would provide incremental performance, category cues might. Heller *et al.* (1996) showed that category cues, such as furniture, fruit and vehicles, aided sighted (but blindfolded) and blind persons in tactual identification of raised-line drawings of objects, such as chair, apple and truck. Without such cuing, subjects could identify about one-quarter to one-third of items, but with cuing could identify about two to three times as many. Multiple-choice testing also led to very high performance levels. In the difference between uncued identification and multiple-choice testing, performance in tactual picture identification behaved like odor identification. In so far as the responses that subjects give in uncued identification fall outside whatever 'natural' category one would give, then it will afford some opportunity for improvement. This would presumably prove true for odors and may have worth. Nevertheless, the outcome will inevitably depend on the categories chosen. Such testing would presumably demonstrate how the addition of semantic information short of the actual label could increase the information transmitted, but it would not allow an index of the amount of semantic information added. For ketchup, for example, the cue condiment might reduce uncertainty considerably since subjects often gave names of items other than condiments, though it would not reduce uncertainty for mustard quite so much since errors for that tended to fall into the category condiment already. The category cue food would reduce uncertainty for ketchup or mustard hardly at all since the responses already fell into that category. The category pureed food would fall intermediate between condiment and food for ketchup. And so on.

For the ordinary person, i.e. one not trained to perceive odors analytically, odors offer little or no redundancy of information. Perceptual 'point of view', however, may alter that situation. Experts involved in the development or assessment of a chemosensory product will have little trouble perceiving it more or less analytically (see Cain and Algom, 1997). Sensory analysis of food and consumer products generally entails attribute descriptions (Stone and Sidel, 1985; Meilgaard *et al.*, 1987). It actually takes little time to develop the analytical approach. Within an hour or two, a panel of untrained persons working with a panel leader can develop a set of descriptors for a product, such as a coffee or candy, and then can use these various descriptors consistently to compare products.

Once one has found particular notes in a sample, it becomes possible to 'look' for these to help identify that sample. The ordinary person has undoubtedly never had to do this and accordingly derives no benefit from the analytical task of the sort put forth in the present Study 2. Persons trained to look at odors analytically could possibly benefit from attribute rating in a way naive persons would not. Attributes once discovered could become part of a code. Naive subjects might possibly stumble upon such an encoding strategy by chance, but this would seem unlikely. In any case, development of an analytical perspective may provide more of a key to transfer of training than mere association of labels.

With so little known about semantic memory for odors, no single investigation will reveal all its aspects. (For perspective, we can note that the literature on human memory, mostly for material presented via vision, grows at a rate of >100 articles per month.) Certain promising leads have emerged from the present investigation:

1. The relationship between veridicality and confidence ratings could stand more use since such ratings may

provide a promising way to quantify information accessible from semantic memory. The use of confidence ratings retrospectively and prospectively by the same subjects may offer a clearer view of how much semantic information the subjects can access. Within this context, the simultaneous study of semantic memory for nonolfactory, perceptual information would offer a useful point of comparison.

2. The inconsistency of incorrect answers seems to merit study. A simple place to begin would be to allow multiple guesses for each item with accompanying confidence ratings. Perhaps subjects hold multiple hypotheses about identity, but experimental protocol allows expression of only one.
3. What, we may ask, is the minimum amount of semantic information needed to precipitate correct identification? Category cuing, as discussed above for both its assets and liabilities, could be used to address the matter. Since feedback with veridical labels increases performance dramatically, it would be useful to follow how long the advantage endowed by such manipulations lasts.
4. Category cuing and feedback with veridical labels can be seen to lie on a continuum of degree of information supplied, though in some respects they can have categorically different effects. Feedback with veridical labels reduces the task of identification to one of recognition, with two attendant benefits: (i) it reduces the repertoire of permissible responses to a specific set, which should aid retrieval; and (ii) it reduces the discriminative burden on the subject, who, after feedback, may not need to discriminate, say, one fruit from another since only one name of a fruit appeared among the labels given in feedback. Category cuing may reduce the repertoire of permissible responses, but may not decrease discriminative burden and instead may increase it since the cuing will merely point up the need to discriminate among the presumably similar-smelling items, such as fruits, within a category.
5. Depth of processing of olfactory information seems to have little relevance to odor identification in untrained persons, but perhaps the matter merits further exploration in trained persons. The depth involved in training for quantitative descriptive analysis goes well beyond any used in experiments. Perhaps a subject so trained has better access to semantic memory and would show incremental benefits from the addition of small amounts of semantic information.
6. Although subjects appear to make both semantic and perceptual errors in their identifications, the matter could be demonstrated more clearly with suitably chosen stimuli (e.g. those with little chemical commonality) and mixtures that might form either perceptually or semantically ambiguous stimuli (see Rabin and Cain, 1989).
7. Some researchers (e.g. Reber, 1993) have pointed out that the studies of memory and of learning have exhibited remarkably little intersection, a matter that researchers in odor science might keep in mind. As we ask how well subjects can access semantic memory for odors, we might also ask how the information gets there in the first place. It appears that untrained subjects learn associations to new odors only very slowly. Indeed, Cain *et al.* (1995) have speculated that learning about everyday odors may proceed so slowly as to continue into middle age, where performance peaks but where the gains from experience begin to erode from age-associated loss of sensitivity or discriminative ability (see de Wijk and Cain, 1994). An inability to access semantic memory for odors may largely indicate that people do not overlearn odors the way they do visual objects. Hence, as with a list of foreign vocabulary words for the young student, performance may be patchy until overlearning occurs. In any case, the identification of odors learned to different specified criteria and then presented for testing much later might show that, although a reasonable amount of semantic information exists in semantic storage without training, much more could be stored from training, with an attendant increase in accessibility (see Jehl *et al.*, 1997; Larsson, 1997).

Acknowledgements

This paper was presented at the E.P. Köster Symposium on Olfaction, Utrecht University, on 22 October 1996. Supported by research grant 5 RO1 DC 00284-12 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health. We thank Michael D. Rabin for the cluster analysis. We also thank Anthony Cecala, Edward Kim and Maryann Savage for their experimental contributions. Prof. Thomas Morton of the University of California, Riverside, offered the idea that multiple guesses may reveal more about semantic memory than we saw with one answer.

References

Bahrick, H.P. and **Hall, L.K.** (1991) *Preventive and corrective maintenance of access to knowledge*. Appl. Cogn. Psychol., 5, 1–18.

Cain, W.S. (1979) *To know with the nose: keys to odor identification*. Science, 203, 467–470.

Cain, W.S. (1980) *Chemosensation and cognition*. In van der Starre, H. (ed.), Olfaction and Taste VII. IRL Press, London, pp. 347–357.

Cain, W.S. (1982a) *Odor identification by males and females: predictions versus performance*. Chem. Senses, 7, 129–142.

Cain, W.S. (1982b) *Sumner's 'On testing the sense of smell' revisited*. Yale J. Biol. Med., 55, 515–519.

Cain, W.S. (1988) *Olfaction*. In Atkinson, R.C., Herrnstein, R.J., Lindzey, G. and Luce, R.D. (eds), Stevens' Handbook of Experimental Psychology, Vol. I: Perception and Motivation, rev. edn. Wiley, New York, pp. 409–459.

Cain, W.S. and **Algom, D.** (1997) *Perceptual and mental mixtures in odor and in taste: are there similarities and differences between experiments or between modalities? Reply to Schifferstein (1997)*. J. Exp. Psychol.: Hum. Percept. Perform., 23, 1588–1593.

Cain, W.S. and **Gent, J.** (1986) *Use of odor identification in clinical testing of olfaction*. In Meiselman, H.L. and Rivlin, R.S. (eds), Clinical Measurement of Taste and Smell. Macmillan, New York, pp. 107–116.

Cain, W.S. and **Krause, J.** (1979) *Olfactory testing: rules for odor identification*. Neurol. Res., 1, 1–9.

Cain, W.S. and **Potts, B.C.** (1996) *Switch and bait: probing the discriminative basis of odor identification via recognition memory*. Chem. Senses, 21, 35–44.

Cain, W.S. and **Rabin, M.D.** (1989) *Comparability of two tests of olfactory functioning*. Chem. Senses, 14, 479–485.

Cain, W.S. and **Turk, A.** (l985) *Smell of danger: an analysis of LP-gas odorization*. Am. Ind. Hyg. Assoc. J., 46, 115–126.

Cain, W.S., Gent, J., Catalanotto, F.A. and **Goodspeed, R.B.** (1983) *Clinical evaluation of olfaction*. Am. J. Otolaryngol., 14, 252–256.

Cain, W.S., Gent, J.F., Goodspeed, R.B. and **Leonard, G.** (1988) *Evaluation of olfactory dysfunction in the Connecticut Chemosensory Clinical Research Center*. Laryngoscope, 98, 83–88.

Cain, W.S., Stevens, J.C., Nickou, C.M., Giles, A., Johnston, I. and **Garcia-Medina, M.R.** (1995) *Life-span development of odor identification, learning and olfactory sensitivity*. Perception, 24, 1457–1472.

Carey, S. and **Diamond, R.** (1977) *From piecemeal to configuration representation of faces*. Science, 195, 312–314.

Clark, S.E. (1997) *A familiarity-based account of confidence–accuracy inversions in recognition memory*. J. Exp. Psychol.: Learn. Mem. Cognit., 23, 232–238.

Craik, F.I.M. and **Lockhart, R.S.** (1972) *Levels of processing: a framework for memory research*. J. Verbal Learn. Verbal Behav., 11, 671–684.

Crowder, R.G. and **Schab, F.R.** (1995) *Introduction*. In Schab, F.R. and Crowder, R.G. (eds), Memory for odors. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 1–7.

Davis, R.G. (1975) *Acquisition of verbal associations to olfactory stimuli of varying familiarity and to abstract visual stimuli*. J. Exp. Psychol.: Hum. Learn. Mem., 104, 143–142.

Davis, R.G. (1977) *Acquisition and retention of verbal associations to olfactory and abstract visual stimuli of varying similarity*. J. Exp. Psychol.: Hum. Learn. Mem., 3, 37–51.

Desor, J.A. and **Beauchamp, G.K.** (1974) *The human capacity to transmit olfactory information*. Percept. Psychophys., 16, 551–556.

de Wijk, R.A. and **Cain, W.S.** (1994) *Odor quality: discrimination versus free and cued identification*. Percept. Psychophys., 56, 12–18.

de Wijk, R.A., Nordin, M., Cain, W.S., Nordin, S. and **Murphy, C.** (1994) *Relative independence of odor quality discrimination and odor detection in aging*. Chem. Senses, 19, 461.

Doty, R.L., Shaman, P. and **Dann, M.** (1984) *Development of the University of Pennsylvania Smell Identification Test: a standardized microencapsulated test of olfactory function*. Physiol. Behav., 32, 489–502.

Dravnieks, A. (1982) *Odor quality: semantically generated multi-dimensional profiles are stable*. Science, 218, 799–801.

Dravnieks, A. (1985) Atlas of Odor Character Profiles. American Society for Testing and Materials, Philadelphia, PA.

Engen, T. (1987) *Remembering odors and their names*. Am. Scient., 75, 497–503.

Eskenazi, B., Cain, W.S., Novelly, R.A. and **Friend, K.B.** (1983) *Olfactory functioning in temporal lobectomy patients*. Neuropsychologia, 21, 365–374.

Eskenazi, B., Cain, W.S. and **Friend, K.** (1986) *Exploration of olfactory aptitude*. Bull. Psychon. Soc., 24, 203–206.

Garner, W.R. (1962) Uncertainty and Structure as Psychological Concepts. Wiley, New York.

Hartigan, J. (1977) Cluster Analysis of Variables. University of California Press, Berkeley, CA.

Heller, M.A., Calcaterra, J.A., Burson, L.L. and **Tyler, L.A.** (1996) *Tactual picture identification by blind and sighted people: effects of providing categorical information*. Percept. Psychophys., 58, 310–323.

Herz, R.S. and **Eich, E.** (1995) *Commentary and envoi*. In Schab, F.R. and Crowder, R.G. (eds), Memory for Odors. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 159–175.

Herz, R.S. and **Engen, T.** (1996) *Odor memory: review and analysis*. Psychon. Bull. Rev., 3, 300–313.

Holyoak, K.J. (1991) *Symbolic connectionism: toward third-generation theories of expertise*. In Ericsson, K.A. and Smith, J. (eds) Toward a General Theory of Expertise. Cambridge, Cambridge University Press, pp. 301–335.

Howard, R.W. (1995) Learning and Memory: Major Ideas, Principles, Issues and Applications. Praeger, Westport, CT.

Jehl, C., Royet, J.-P. and **Holley, A.** (1997) *Role of verbal encoding in short- and long-term odor recognition*. Percept. Psychophys., 59, 100–110.

Koriat, A. (1995) *Dissociating knowing and the feeling of knowing: further evidence for the accessibility model*. J. Exp. Psychol.: Gen., 124, 311–333.

Larsson, M. (1997) *Semantic factors in episodic recognition of common odors in early and late adulthood*. Chem. Senses, 22, 623–633.

Lawless, H. and **Engen, T.** (1977) *Associations to odors: interference, mnemonics and verbal labeling*. J. Exp. Psychol.: Hum. Learn. Mem., 3, 52–59.

Lawless, H. and **Zwillenberg, D.** (1983) *Clinical testing of taste and olfaction*. Trans. Penn. Acad. Ophthalmol. Otolaryngol., Fall, 190–196.

Martinez, B., Cain, W.S., de Wijk, R., Spencer, D., Novelly, R. and **Sass, K.** (1993) *Olfactory functioning before and after temporal lobe resection for intractable epilepsy*. Neuropsychology, 7, 351–363.

Meilgaard, M., Civille, G.V. and **Carr, B.T.** (1987) Sensory Evaluation Techniques. CRC Press, Boca Raton, FL.

Metcalfe, J. (1986) *Feeling of knowing in memory and problem solving*. J. Exp. Psychol.: Learn. Mem. Cognit., 12, 288–294.

Nelson, T.O. (1984) *A comparison of current measures of the accuracy of feeling-of-knowing predictions*. Psychol. Bull., 95, 109–133.

Nelson, T.O., Gerler, D. and **Narens, L.** (1984) *Accuracy of feeling-of-knowing judgements for predicting perceptual identification and relearning*. J. Exp. Psychol.: Gen., 113, 282–300.

Rabin, M.D. (1988) *Experience facilitates olfactory quality discrimination*. Percept. Psychophys., 44, 532–540.

Rabin, M.D. and **Cain, W.S.** (1984) *Odor recognition: familiarity, identifiability and encoding consistency*. J. Exp. Psychol.: Learn. Mem. Cognit., 10, 316–325.

Rabin, M.D. and **Cain, W.S.** (1989) *Attention and learning in the perception of odor mixtures*. In Laing, D.G., Cain, W.S., McBride, R.L.

and Ache, B.W. (eds), Perception of Complex Smells and Tastes. Academic Press, Sydney, pp. 173–188.

Reber, A.S. (1993) Implicit Learning and Tacit Knowledge. Oxford University Press, New York.

Schab, F.R. and **Crowder, R.G.** (1995) *Odor recognition memory*. In Schab, F.R. and Crowder, R.G. (eds), Memory for odors. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 9–20.

Schemper, T., Voss, S. and **Cain, W.S.** (1981) *Odor identification in young and elderly persons: sensory and cognitive limitations*. J. Gerontol., 36, 446–452.

Schneider, W. (1985) *Developmental trends in the metamemory–memory behavior relationship: an integrative review*. In Forrest-Pressley, D.L., MacKinnon, G.E. and Waller, T.G. (eds), Metacognition, Cognition and Human Performance, Vol. 1. Theoretical Perspectives. Academic Press, Orlando, FL, pp. 57–109.

Stone, H. and **Sidel, J.L.** (1985) Sensory Evaluation Practices. Academic Press, Orlando, FL.

Sumner, D. (1962) *On testing the sense of smell*. Lancet, ii, 895–897.

Tulving, E. (1983) Elements of Episodic Memory. Oxford University Press. Oxford.

Underwood, B.J. (1983) Attributes of Memory. Scott, Foresman, Glenview, IL.

Whittlesea, B.W.A., Jacoby, L.L. and **Girard, K.** (1990) *Illusions of immediate memory: evidence of an attributional basis for feelings of familiarity and perceptual quality*. J. Mem. Lang., 29, 716–732.

Wright, H.N. (1987) *Characterization of olfactory dysfunction*. Arch. Otolaryngol. Head Neck Surg., 113, 163–168.

Accepted on January 23, 1998