Chem. Senses 28: 651–658, 2003 DOI: 10.1093/chemse/bjg058

Olfactory Metacognition

Fredrik U. Jönsson and Mats J. Olsson

Department of Psychology, Uppsala University, Sweden

Correspondence to be sent to: Fredrik Jönsson, Department of Psychology, Uppsala University, Box 1225, S-75142 Uppsala, Sweden. e-mail: fredrik.jonsson@psyk.uu.se

Abstract

The current paper focuses on the subjective knowledge people have about their ability to name odors. Previous investigations of such metacognitive aspects of olfactory cognition are very scarce and have yielded results that need further scrutiny. In two experiments, we investigated three metamemory judgments about odor identity. As opposed to previous findings, participants' feeling of knowing judgments about odor identity predicted later recognition. Participants were also accurate but highly overconfident in their retrospective confidence in odor identification. A strong and imminent feeling of being able to name an odor, a so-called 'tip of the nose' experience, was found to predict later recall, but was otherwise poorly related to any partial activation of the odor name or other information associated with the odor. This makes it different from the commonly investigated 'tip of the tongue' phenomenon. The current study shows that olfactory metamemory is related to actual knowledge, a finding that is in line with what has been observed for other modalities.

Key words: confidence, feeling of knowing, metamemory, odor identification, tip of the nose, tip of the tongue

Introduction

It is quite well established that people are poor in naming odors. The average rate of correct naming of a set of common odors (e.g. coffee, cinnamon, and tar) rarely exceeds 50% (Desor and Beauchamp, 1974; Cain, 1979; de Wijk, 1994; de Wijk and Cain, 1994; de Wijk *et al.*, 1995; Cain *et al.*, 1998), and the identification rate of any single item rarely reaches 100%. However, if a person instead is presented with a picture of a common visual object, e.g. a pencil, every healthy subject will be able to correctly label it as a 'pencil'. The fact that this is not the case with common odors might mirror the lack of importance of precise odor identification to humans. Instead of being swiftly identified, odors may be divided into broader semantic categories or be associated with places and events (Engen, 1991; de Wijk *et al.*, 1995). Cain (1982) demonstrated that despite our relative inability to correctly name odors based on olfactory cues, we are typically under the impression that we can. Whereas many studies have investigated odor naming *per se*, the subjective knowledge people have about their odor identifications, their identification attempts and how correct these actually are have not received much attention in the literature. The little that has been done has also yielded some surprising results, which are reviewed below (Lawless and Engen, 1977; Cain *et al.*, 1998). The current study focused on three types of metamemory phenomena about odor identity: *feeling of knowing* (FOK), the *tip of the nose* (TON) phenomenon (Lawless and Engen, 1977) and *retrospective confidence*. They will now be introduced in more detail.

Feeling of knowing

Sometimes when we try to retrieve information from memory we fail. An interesting phenomenon accompanying such retrieval failures is the phenomenological experience that although not retrieved we anyway know this information. This is in the literature referred to as a FOK (Hart, 1965). The FOK can be of varying degree, from no FOK at all to a very strong FOK. A relevant question is whether the FOK for an odor's name, despite not being able to retrieve it from memory, mirrors actual storage of it in memory? The correctness of these experiences can be measured in the laboratory and the typical procedure is to calculate a Goodman–Kruskal gamma correlation between the FOK judgments and a recognition test following the judgments (Nelson, 1984). Most FOK studies have focused on verbally presented general information questions (e.g. 'What is the capital of Australia?'). There is an extensive body of FOK data showing that people are, in general, moderately accurate in monitoring their knowledge (Metcalfe, 2000). However, less attention has been directed towards FOK judgments for non-verbal stimuli (Peynircioglu *et al.*, 1998;

Perfect and Hollins, 1999). Cain and colleagues (Cain *et al.*, 1998) were the first to investigate FOK judgments about odor identity. They let young adults try to identify a set of common odors and answer general information questions. The participants made FOK judgments on the first seven odors not identified and the first seven questions not answered. Each such trial was followed by an eight-alternative, forced-choice recognition test. Cain *et al.* found no significant predictive validity for FOK judgments about odor identity, but did for general information questions. The conclusion is, therefore, that when people feel that they know the name of an odor, their FOK judgments are actually at chance. In experiment 1 we investigated the predictive validity of such FOK judgments using a more powerful method than Cain *et al.*

The tip of the tongue phenomenon

A phenomenon closely related to the FOK is the strong feeling that a currently inaccessible word is stored in memory and will be retrieved. This is often referred to as the 'tip of the tongue' (TOT) phenomenon. [The TOT experience is a prediction of recall imminence, while the FOK is usually a prediction of recognition performance (Koriat, 1998). FOK judgments can be made on any non-recalled item and are requested by the experimenter, while TOT experiences occur involuntarily. In addition, participants with a TOT experience are confident that they can eventually recall the target information, with or without additional cues. An FOK judgment, on the other hand, is usually an assessment of the recognition likelihood for the missing word.] Brown and McNeill (1966) defined the TOT phenomenon as 'a state in which one cannot quite recall a familiar word but can recall words of similar form and meaning' (p. 325). Lawless and Engen (1977) investigated strong feelings of imminent recall in the case of odor identification. They named it the tip of the nose phenomenon (TON) as a parallel to the TOT experience. They let participants make familiarity ratings and odor identification attempts for a set of 48 odors. If the participants were unable to identify an odor but made a high familiarity rating and felt that they had the odor name on the tip of the tongue (i.e. a TON experience), they filled out a questionnaire asking for partial information about the sought-for odor label and other associated information. When people have a TOT experience, they often have partial information about the name, such as how many syllables it contains, information about letters in the word, syllabic stress, and words with a similar meaning and sound (Brown and McNeill, 1966; Brown, 1991). In Lawless and Engen's study, the participants had virtually no such partial information about the sought-for name of the odor. In experiment 2 we replicated Lawless and Engen's experiment, with some important additions.

Retrospective confidence judgments

Anyone having tried to name odors knows how difficult it is to identify them without visual and other contextual cues. Cain (1982) showed that people tend to be quite confident about their ability to identify odors, but anyone having performed laboratory tests of odor naming knows that participants are often surprised by the difficulty of identifying even very familiar odors. Taking into consideration the difficulty of naming even common odors, an interesting approach is to find out to what extent people know whether they have identified an odor correctly or not. This type of judgment, a retrospective confidence judgment, is different from the FOK and TON experiences, in that it is not a prediction of future retrieval or recognition of information from memory, but instead a judgment made about the correctness of already retrieved information. In our context, participants tell how confident they are in the correctness of an odor label just retrieved. Cain *et al.* (1998), collected confidence ratings on a category scale from 1 (very low confidence) to 5 (very high confidence), asking the participants to rate the correctness of their odor-naming attempts. The results indicated that people can differentiate between correct and incorrect identifications, because the mean confidence was higher for the correct responses than for the incorrect ones. However, Cain *et al.* did not report the exact degree of the relationship between confidence and odor-naming performance (as measured by a correlation between confidence and the correctness of the answers). The study also does not give any information regarding whether the participants were over- or underconfident. It is important to note that although a person may be able to discriminate between correct and incorrect responses, i.e. to generally give a higher confidence rating for correct responses than for incorrect ones, he or she can still be over- or underconfident. An example of overconfidence is when a participant reports 100% confidence in his or her responses while averaging only 80% correct. An example of underconfidence would be if a participant averages 60% correct on trials where he or she is 80% confident. There are thus two different ways of looking at the confidence–accuracy relationship, which complement each other. These issues are investigated in experiment 2.

Experiment 1

Experiment 1 aimed to investigate the predictive validity of FOK judgments about odor identity in a study with a more powerful method than that of Cain *et al.* (1998).

Methods

Participants

Thirty-nine participants (11 men and 28 women; mean age = 24 years, range = 20–31 years) were recruited from the Department of Psychology at Uppsala University. They participated for course credits or were given a movie ticket

Table 1 The odors in experiments 1 and 2

Aniseed	Curry	Oregano	Vanilla	Hair-styling gel*
Apple	Dill	Paint	Vicks[a]	Licorice*
Apricot	Garlic	Peanut butter	Window-cleaning fluid	Meat buljong*
Bell pepper	Gasoline	Peppermint[b]	Vinegar	'Messmör'*[c]
Black pepper	Geranium	Pipe tobacco	Violet[d]	Onion*
Bleach	Ginger	Sage	Yeast	Plastic Padding*
Blue cheese	Juicy Fruit (chewing gum)	Sandalwood	Acetone*	Potato chips*
Cacao	Juniper	Shoe cream	All-purpose cleaner*	Raisin*
Caraway	Ketchup	Soft soap	Baby powder*	Shampoo*
Cardamom	Laundry detergent	Spruce	Banana*	Snuff*[e]
Cigarette-butt	Lemon	Thick felt tip	Cheese doodles*	Soap*
Cinnamon	Motor oil	Tabasco	Dishwashing liquid*	Soy*
Clove	Nutmeg	Tar	Furniture polish*	Strawberry jam*
Coffee	Orange	Tea	Glue*	Toilet refreshener*

Odors marked with an asterisk (*) were used as stimuli in experiment 2 only. All other odors were used in both experiments.
[a]A product inhaled through the nose to relieve cold.
[b]An essence from Apoteksbolaget.
[c]A Swedish dairy product.
[d]An essence from Stockholm Aeter & Essencefabrik AB.
[e]A Swedish tobacco product.

voucher (worth ~75 SEK). All participants reported a functional sense of smell.

Stimuli

The test stimuli consisted of 48 common odorants, e.g. coffee, garlic, apple and tea. All odorants are listed in Table 1. Some odorants, e.g. apple, were changed regularly to keep them fresh and prevent the odor quality from changing. The stimuli were all real-world items except for the essences violet and peppermint. Odorants were presented in 160 ml tinted glass jars with screw lids. Cotton pads prevented visual inspection of the stimulus material in the jars.

Procedure

Participants were tested individually. First, they were welcomed and given a questionnaire concerning some background data, such as their age and sex. Second, they smelled and tried to name two practice odors and the 48 test odors. They were instructed to smell each odor only once and with eyes closed. If they thought they could name the odor they did so and the next odor was distributed to them. If they were unable to name an odor, they instead made a FOK judgment, i.e. they rated how sure they felt that they would recognize the correct odor name if it was presented to them directly after the rating. This was done on a scale ranging from 50% (chance level) to 100% (totally sure), with 10% intervals. The participants were thoroughly acquainted with the meaning and use of the scale. They were instructed to be as realistic in their judgments as possible. After the FOK rating, a one-alternative identification test followed where they were shown either the correct or an incorrect odor name printed in black on white paper. The participants were informed that 50% of the odor names presented would be correct and 50% incorrect, which was also the case. The presentation order of the odors was uniquely randomized for each participant, as was the presentation of the false and correct cues following the FOK judgments. Upon completion of the test, participants were thanked and debriefed. The whole experiment took an average of 25 min to complete.

Results and discussion

The participants tried to name on average 28% (n = 443) of the odors presented to them. About 16% (n = 252) of all odors were correctly named (i.e. 57% of the trials where an attempt was made to name the odor). If near misses (e.g. lemon or fruit for orange) were counted as correct answers, the proportion correct increased to 19% (n = 310). The low naming accuracy confirms previous findings of the difficulty in naming even familiar odors. The remaining trials were classified as retrieval failures and led to FOK judgments.

On average, there were 29.74 (SD = 8.13) FOK judgments per individual. A Goodman–Kruskal gamma correlation (Nelson, 1984) between FOK judgments and recognition performance was calculated for every participant. The gamma correlation (G) is a non-parametric measure that relates the ranking of the FOKs to the correctness of response on the memory test. The mean gamma correlation was moderate and statistically reliable [$G(39) = 0.34$; SD =

0.37; 95% confidence interval (CI) = 0.22–0.45]. An aggregated analysis of all the participants' data yielded a similar gamma correlation [G(1159) = 0.37; $P < 0.0001$]. This is close to what is typically found in FOK experiments of other kinds of materials, with correlations usually between 0.45 and 0.55 (Metcalfe, 2000). Another way of looking at metamemory accuracy is to calculate an overall proportion of correct responses for each FOK category, and to correlate these proportions with the FOK categories. Using this method, Cain *et al.* (1998) did not find a statistically reliable Pearson correlation (r). The correlation for the present data, however, was higher and statistically reliable [$r(6) = 0.99$; $P < 0.001$]. Overall, these results suggest that FOK judgments about odor identity are associated with subsequent recognition of correct names.

Experiment 2

In the second experiment, the overall aim was to replicate and further investigate the TON phenomenon (Lawless and Engen, 1977). First, we assessed the extent to which the TON experience was resolved with the correct odor name. Second, we investigated whether the strength of the TON feeling would predict the resolution probability, with more resolved TON experiences the stronger the experience. If the amount of correctly resolved TON experiences is positively related to the strength of the experience, it means that the participants have some valid cue or cues as a basis for the experience. A third question concerned the type of odor information that is related to the TON experience, i.e. partial information about the odor name or other information associated with the odorous object. In this experiment we also investigated the accuracy of participants' retrospective confidence in odor identity.

Method

Participants

Forty participants (18 men and 22 women; mean age = 24 years, range = 19–31 years) were recruited from the Department of Psychology at Uppsala University. They participated for course credits or were given a movie ticket voucher (worth ~75 SEK). All participants reported a functional sense of smell.

Stimuli

The stimuli consisted of 70 common odorants, e.g. coffee, garlic, orange and tea, including those in experiment 1. All odorants are listed in Table 1. Some odorants, e.g. apple and orange, were changed regularly to keep them fresh and prevent the odor quality from changing. As in experiment 1, the stimuli were real-world items, except for the essences violet and peppermint. Odorants were presented in 160 ml tinted glass jars with screw lids. Cotton pads prevented visual inspection of the stimulus material in the jars.

Procedure

All participants were tested individually. Upon arrival, the participant was welcomed and given a questionnaire concerning some background data. He or she was then given an oral overview of the experimental procedure, followed by more extensive written instructions. At test, odors were presented one at a time. The participant was requested to smell each odor only once and one of three outcomes of the smelling was possible; either the person (1) made an attempt to name the source of the odor by its proper name, (2) had a TON experience, or if neither of these two, (3) he or she made a FOK judgment.

If an attempt to name the odor was made, the participant also rated how confident he or she was that the name was correct on a scale from 0% to 100%, with 20% intervals. The participant was informed that it would be a contradiction to first name the odor and afterwards make a confidence judgment of 0% (which means that he or she is absolutely sure that the name is incorrect).

If the participant could not name the odor and had a TON experience, 90 s were allotted to retrieve the sought-for odor name. [In the instructions, the participants were given the following definition/explanation of the TON experience: 'Sometimes when you smell something, you don't always come up with what it is, even though you have a very strong feeling that you know it; i.e. you have it on the tip of the tongue. If you have such a strong feeling that you are on the verge of coming up with what odor it is and believe that you will recall it if just given time, then you should do the following…' (authors' translation from Swedish).] During this period a questionnaire was also filled out and the participant could smell the odor again as many times he or she wanted. If and when the participant retrieved the name during the allotted time it was to be written down on the questionnaire, all further activities were stopped and the experiment continued with the next odor. For reasons of comparability, the questionnaire was adapted from Lawless and Engen (1977) and consisted of two parts. One part asked about partial information about the sought-for word. The other part asked for other associated information about the odor. The partial information of the odor name concerned (1) whether the participants could name any or some of the letters in the sought-for odor name; (2) how many syllables it had; (3) on what syllable the accent was; (4) whether they could come up with words with a similar meaning; or (5) words that sounded similar. The associated information asked for was (6) if the participants could name a similar odor; (7) if they could name the category to which the odor might belong; (8) an object from which the odor might come; (9) a place from which the odor might come; (10) if they could form a visual image of the object; or (11) a place from which the odor might have come. On the questionnaire, the participants were also asked to rate the strength of their TON experience on a two-point scale ('strong' or 'very strong'). If the participants could neither come up with a

label for the odor, nor had a TON experience, they were instructed to make a FOK judgment (as in experiment 1).

Before the experiment started, the participants were thoroughly instructed about the idea behind the judgment scales for confidence (from 0% to 100%) and FOK (from 50% to 100%). They were also instructed to note the difference between them and to be as realistic as they could. When they made a confidence rating or an FOK judgment, they reported their answers to the experimenter, who wrote them down. All TON responses were gathered on the questionnaire. To prevent confusion between the two scales, the participants had both scales visible in front of them during the whole experiment. The experiment began with two practice odors. The order of the stimuli was uniquely randomized for every participant. After finishing the experiment, participants were thanked and debriefed. The whole experiment took an average of 65 min to complete.

Results and discussion

The participants tried to name and made confidence judgments on about 37% (n = 1034) of the odor trials. About 19% (n = 531) of all odors were correctly named (i.e. 51% of the trials where an attempt was made to name the odor). This is similar to what was found in experiment 1. If near misses (e.g. lemon for orange) were counted as correct answers, the proportion correct increased only slightly to 22% (n = 625). About 17% (n = 484) of all trials were judged to be TON experiences and the remaining 46% (n = 1272) constituted FOK judgments.

Retrospective confidence judgments

A gamma correlation between confidence and naming accuracy was calculated for each participant. On average, there were 26.83 (SD = 10.39) confidence judgments per individual, and the mean gamma correlation across participants was $G(40) = 0.64$ (SD = 0.30; 95% CI = 0.54–0.73). When aggregating all individuals' judgments per confidence category (see Figure 1), the Pearson correlation was high and statistically reliable [$r(6) = 0.93$; $P < 0.05$].

The participants showed a strong overconfidence in their ability to identify odors in the upper confidence categories, whereas they were closer to perfect calibration in the lower categories (20% and 40%, respectively; Figure 1). The mean of all participants' individual O/U scores was high and significantly different from zero (O/U = 0.26; SD = 0.16; 95% CI = 0.21–0.31). [The O/U score is a directional measure of the appropriateness of assigned confidence levels. The O/U score equals the average deviation between the probability assigned to a subcategory and the proportion correct within the subcategory. The formula is

$$O/U = \frac{1}{N}\sum_{t=1}^{T} n_t(r_t - c_t)$$

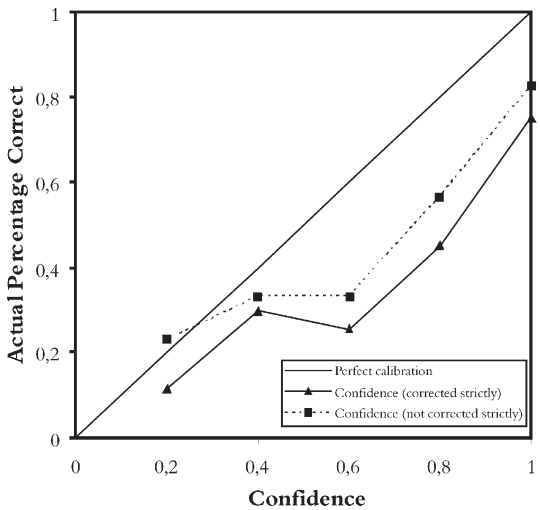

Figure 1 Confidence judgments plotted against actual proportion of correct answers per confidence category. The diagonal represents perfect calibration. The solid line shows the calibration curve for responses that were scored with a strict criterion for what constituted a correct answer. The dashed line shows a calibration curve that also includes near misses counted as correct responses.

where N equals the total number of judgments, c_t equals the proportion correct in subcategory t, n_t equals the number of judgments in subcategory t, and r_t equals the probability assigned to subcategory t (Yates, 1990).] To conclude, confidence and correctness in odor identification was highly correlated, but participants were markedly overconfident in their judgments.

Feeling of knowing

On average, there were 31.80 (SD = 10.49) FOK judgments per individual. Because the participants made both TON and FOK judgments in this experiment, the stronger and supposedly more predictive FOKs appeared as TON experiences. This diminished the observed predictive validity of the FOKs as seen in the aggregated gamma correlation [$G(1272) = 0.14$; $P = 0.003$; cf. experiment 1].

Tip of the nose phenomenon

All 70 odorants evoked at least one TON experience and the number of TON experiences per person varied from 1 to 28, with a mean of 12.1 (SD = 7.4). The typical TOT incidence for verbal stimuli is ~13 ± 5% (Brown, 1991), which is similar to what we found in this experiment: 17.3% (n = 484) of all odor trials evoked a TON experience. Of these and within the allotted 90 s, 23.8% (n = 115) were named correctly, 41.3% (n = 200) were still unresolved, whereas as much as 34.9% (n = 169) were 'resolved' with an incorrect label.

One of the aims of the current study was to investigate whether the strength of the TON experience and actual knowledge would be correlated. The resolution frequency for 'strong' versus 'very strong' TONs was analyzed as a function of the number of correctly resolved strong (or very

Table 2 Mean proportion of correct responses and mean total proportion of responses on the TON questionnaire as a function of TON strength

Question	Proportion of correct responses		
	Overall	Strong TON	Very strong TON
6. Can you name a similar odor?	0.08	0.07	0.08
7. Can you name a general category for the odor?	0.17	0.12	0.22[a]
8. Can you name an object from which the odor might have come?	0.13	0.13	0.14
9. Can you name a place from which the odor might have come?	0.20	0.19	0.21

[a]A statistically reliable difference ($P < 0.05$) between strong and very strong TONs as measured by proportion of correct responses.

strong) TONs, divided by the total number of strong (or very strong) TONs. A paired sampled *t*-test showed that very strong TONs were more frequently resolved (36.7%) than TONs that were judged to be just strong [19.0%; $t(32) = 2.97$; $P < 0.01$]. This indicates that the participants have some reliable cue or cues on which they base their feeling that they will soon be able to retrieve the sought-for odor name.

Questions 6–9 in the questionnaire targeted information associated to the odor. Table 2 shows the proportion of correct responses to these questions. In response to question 6, the participants tried to name a similar odor but did so very poorly. Questions 7–9 all concerned relevant information about the odor source. Overall, the proportion of correct responses on these questions was relatively low. The scoring of question 9 was very liberal due to the difficulty in checking the correctness of the participants' responses; hence the resolution rate in Table 2 is probably an overestimation.

Because the strength of the TON experience was positively related to later retrieval of the odor name, we analyzed whether the strength would also be related to partial or associated information collected in the TON questionnaire (see Table 2). Also in these analyses, we investigated proportion of correct responses. On questions 6, 8 and 9, there were no statistically reliable differences between strong and very strong TON experiences. For question 7, which asked for a general category for the TON-evoking odor, there was a significantly higher proportion of correct responses [$t(32) = 2.30$; $P = 0.03$] following very strong as compared to strong TON experiences.

The number of responses for the questions (1–5) asking for partial information about the odor name was very low, with most of the answers being incorrect. There was only one (incorrect) attempt to report a word that sounded similar to the sought-for name. On the question asking for words with similar meaning there were only eight responses, one of them correct. On the question concerning on what syllable the accent was, six out of ten responses were correct. On the question asking for how many syllables the sought-for odor label had, 11 out of 32 responses were correct, whereas on the question asking if the participants could name any or some of the letters in the word, only 9 out of 22 responses were correct. Statistical comparisons between strong and very strong TON experiences were therefore not possible.

General discussion

As noted, Cain *et al.* (1998) did not find a predictive validity of FOK judgments about odor identity but did so for general information questions. However, with another method, the current study did find that participants' subjective experience of knowing a yet unidentified odor is related to later recognition of that odor's name. Hence, the conclusion drawn from these studies must be that there is no qualitative, although possibly a quantitative, difference between olfaction and other senses with respect to the predictive validity of FOK judgments.

In the second experiment participants' retrospective confidence in their odor identifications was investigated. The results are consistent with the data of Cain *et al.* (1998) and demonstrate that participants know quite well when they have identified an odor correctly and when they have not (as shown by the gamma correlation). However, they did have an overly high belief in their identifications (see Figure 1) and especially in the higher confidence categories. It remains unclear why people are so overconfident in their odor identifications, and this is certainly an interesting topic for future research. One suggestion is that we sometimes misrepresent odors perceptually. It has been argued (Cain and Potts, 1996) that participants sometimes 'misapprehend' the source of an odor (e.g. lemon is perceived as orange), but their ability to realize this is compromised. If this *confusion hypothesis* is true, people's confidence in their accuracy may be distorted, i.e. people will be overly confident in their answers. In other words, a person who is repeatedly highly confident in incorrectly identified odors due to misapprehension of them will *on average* be more confident than correct in his or her odor naming attempts. Another less interesting possibility is that the overall overconfidence is, at least in part, a methodological artifact (Juslin *et al.*, 2000).

In laboratory studies of the TOT phenomenon, the probability that a target will be retrieved in a minute or two varies around 50% (Brown, 1991). Here we found that only ~24% of the TON experiences were resolved correctly within the assigned 90 s. However, this should be seen in light of Lawless and Engen's (Lawless and Engen, 1977) resolution rate of 38%, which is somewhat closer to what is typically found in TOT studies, as well as to what we found for very strong TONs (37%). [Lawless and Engen (1977) did not report any resolution rate, but did report that 14 out of a total of 37 TON experiences were resolved correctly (38%).

In addition, they did not report whether or not they had any time limit for the participants to resolve their TON experiences, meaning that our resolution rates may not be fully comparable, due to methodological differences.] More replication is needed, but so far the conclusion is that TON experiences are not as predictive of later retrieval as TOT experiences.

Although participants sometimes solve their TON experiences and are also able to differentiate between stronger and weaker experiences, when in such a state, they do not seem to experience any partial activation of the sought-for name as is often the case with TOTs. The latter replicates previous findings (Lawless and Engen, 1977) and seems to be an important difference between TON and TOT experiences. In addition, when we requested information about the odor source other than the name of the odor (especially questions 7–9, in Table 2), the level of knowledge was also low. These findings, together with the lower overall resolution probability of TON experiences, suggest that the TON experience is generally related to the identification phase rather than the naming phase. More precisely, the TON experience arises when attempting to realize that the present odor comes from, for example, cinnamon, rather than from attempting to recall that this red-brown spice that I already know is called 'cinnamon'. The naming of pictures, for example, has been proposed to consist of three separate stages, first an object identification (to know what it *is* regardless of veridical naming ability), then a name activation and finally a response generation (McCauley *et al.*, 1980; Johnson *et al.*, 1996). It might be that a person with a TON experience is often in the first (object identification) stage of the three.

An interesting question is on what people base their phenomenological experience of knowing an odor. The fact that participants resolved stronger TON experiences better than less strong ones has some bearing on metacognitive theory. This indicates that the participants based their experiences on some valid cue or cues. There are two prominent theories of metamemory that are of interest in the current context, i.e. the cue-familiarity theory (Reder, 1987; Metcalfe *et al.*, 1993) and the accessibility theory (Koriat, 1993, 1995). These theories were developed in the context of FOK judgments but have later been expanded to encompass the TOT phenomenon as well (Schwartz and Smith, 1997). The cue-familiarity hypothesis proposes that the familiarity of the eliciting cue is an important basis of metamemory judgments. If a person is cued with a face of a famous person and is unable to retrieve the name of the person, the FOK for the name would thus mainly be based on the perceived familiarity of the face. The accessibility theory instead postulates that it is the amount of related partial information available about the target memory (i.e. the sought-for name) that is important (whether correct or incorrect). Recently, an attempt to bring these two views together has been made regarding FOK judgments (Koriat and Levy Sadot, 2001). Generalized to our data, a TON experience does not seem to be based on the partial access to the odor name, because of the simple fact that there is virtually no such information. However, to the extent that TON experiences of different strengths are positively related to the amount of correct associated information reported (see Table 2), Koriat's theory would find some support. On questions 6, 8 and 9, no such relationships could be observed. However, participants' odor categorizations were better the stronger the TON experience. In line with this, the strength of the experience was also positively related to subsequent retrieval of the odor's name. The conclusion is thus that although the current data showed no strong support for an accessibility perspective concerning TON experiences, some support could be found in participants' odor categorizations.

In the current study, common everyday odors were used to probe odor identification. The participants had most likely encountered most of the them many times before in their lives and it is likely that the odors appeared familiar to most people, even in the absence of correct identification. Indeed, a further analysis of data from an earlier study (Broman *et al.*, 2001) showed that many odors were rated as very familiar although these odors were rarely identified (i.e. named). In line with the cue-familiarity hypothesis (Reder, 1987; Metcalfe *et al.*, 1993) it is possible that the strong feeling of familiarity of the odor created an impression of being able to retrieve the odor's identity. To the extent that odors of low identifiability and high familiarity are common, and if the TON experience is mainly based on cue familiarity, unresolved TON experiences would be more frequent. A relevant topic for further study would therefore be to investigate the role of the familiarity component as a basis of the TON experience. A superior role of the familiarity component, as opposed to the accessibility to pertinent information about the target name, would also be in line with the above hypothesis that TON experiences are more about trying to retrieve the identity of the odor rather than having the name on the tip of the tongue. Future studies should include cross-modal comparisons of metacognitive ability and further investigate the role of existing metacognitive theories in the context of odor identification. The latter is important not only for the understanding of olfactory metacognition, *per se*, but also for the generalizability of current metacognitive theories across different sense modalities.

Acknowledgements

The authors would like to thank Anders Johansson for experimental assistance in the second experiment and Maria Larsson for helpful comments on an earlier version of the manuscript. This research was supported by the Bank of Sweden Tercentenary Foundation (1998–0270).

References

Broman, D.A., **Olsson, M.J.** and **Nordin, S.** (2001) *Lateralization of olfactory cognitive functions: effects of rhinal side of stimulation.* Chem. Senses, **26**, 1187–1192.

Brown, A.S. (1991) *A review of the tip-of-the-tongue experience.* Psychol. Bull., **109**, 204–223.

Brown, R. and **McNeill, D.** (1966) *The 'tip of the tongue' phenomenon.* J. Verb. Learn. Verb. Behav., **5**, 325–337.

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

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

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., **de Wijk, R., Lulejian, C., Schiet, F.** and **See, L.C.** (1998) *Odor identification: perceptual and semantic dimensions.* Chem. Senses, **23**, 309–326.

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

de Wijk, R.A. and **Cain, W.S.** (1994) *Odor identification by name and by edibility: life-span development and safety.* Hum. Factors, **36**, 182–187.

de Wijk, R.A., Schab, F.R. and **Cain, W.S.** (1995) *Odor identification*. In Schab, F.R. (ed.), Memory for Odors. Lawrence Erlbaum, Mahwah, NJ, pp. 21–37.

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

Engen, T. (1991) Odor Sensation and Memory. Praeger, New York.

Hart, J.T. (1965) *Memory and the feeling-of-knowing experience.* J. Educ. Psychol., **56**, 208–216.

Johnson, C.J., **Paivio, A.** and **Clark, J.M.** (1996) *Cognitive components of picture naming.* Psychol. Bull., **120**, 113–139.

Juslin, P., **Winman, A.** and **Olsson, H.** (2000) *Naive empiricism and dogmatism in confidence research: a critical examination of the hard-easy effect.* Psychol. Rev., **107**, 384–396.

Koriat, A. (1993) *How do we know that we know? The accessibility model of the feeling of knowing.* Psychol. Rev., **100**, 609–639.

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

Koriat, A. (1998) *Illusions of knowing: The link between knowledge and metaknowledge*. In Yzerbyt, V.Y. (ed.), Metacognition: Cognitive and Social Dimensions. Sage: Thousand Oaks, CA, pp. 16–34.

Koriat, A. and **Levy Sadot, R.** (2001) *The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing.* J. Exp. Psychol. Learn. Mem. Cogn., **27**, 34–53.

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

McCauley, C., **Parmelee, C.M., Sperber, R.D.** and **Carr, T.H.** (1980) *Early extraction of meaning from pictures and its relation to conscious identification.* J. Exp. Psychol. Hum. Percept. Perform., **6**, 265–276.

Metcalfe, J. (2000). *Metamemory: Theory and Data*. In Tulving, E. and Craik, F.I.M. (eds), The Oxford Handbook of Memory. Oxford University Press, New York, pp. 197–211.

Metcalfe, J., **Schwartz, B.L.** and **Joaquim, S.G.** (1993) *The cue-familiarity heuristic in metacognition.* J. Exp. Psychol. Learn. Mem. Cogn., **19**, 851–864.

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

Perfect, T.J. and **Hollins, T.S.** (1999) *Feeling-of-knowing judgments do not predict subsequent recognition performance for eyewitness memory.* J. Exp. Psychol. Appl., **5**, 250–264.

Peynircioglu, Z.F., **Tekcan, A.I., Wagner, J.L., Baxter, T.L.** and **Shaffer, S.D.** (1998) *Name or hum that tune: feeling of knowing for music.* Mem. Cogn., **26**, 1131–1137.

Reder, L.M. (1987) *Strategy selection in question answering.* Cogn. Psychol., **19**, 90–138.

Schwartz, B.L. and **Smith, S.M.** (1997) *The retrieval of related information influences tip-of-the-tongue states*. J. Mem. Lang., **36**, 68–86.

Yates, J.F. (1990) Judgment and Decision Making. Prentice Hall, Englewood Cliffs, NJ.

Accepted August 23, 2003