Chem. Senses 30: 353–365, 2005

doi:10.1093/chemse/bji030
Advance Access publication April 8, 2005

A Metamemory Perspective on Odor Naming and Identification

Fredrik U. Jönsson, Antoinette Tchekhova, Pär Lönner and Mats J. Olsson

Department of Psychology, Uppsala University, Sweden

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

Abstract

A metacognitive perspective is utilized to elucidate why it is so difficult to name common odors and what characterizes the subjective knowledge people have about their actual odor knowledge. Odor-naming failures are often accompanied by strong feelings of knowing (FOK) or feelings of imminent retrieval of what it is that smells. The paper's two experiments investigate FOK judgements and tip of the tongue (TOT) experiences for odor and person names. The data indicate that our inability to correctly name odors are typically not due to the often proposed uniquely poor association between odors and their proper names, but rather due to failures to identify the odors, that is, failures to know 'what it is'. It was also found that (i) TOT experiences are very unusual for odor names and more so than for person names; (ii) FOK judgements about odor names are significantly less predictive of later retrieval than equivalent judgements about names of persons; (iii) FOK judgements were highly correlated with the familiarity of the cue (odor or picture of famous person), rendering some support for the idea that FOK judgements are based on the perceived familiarity of the cue triggering the FOK; and (iv) the idea that FOK judgements are based on the amount of available information about the sought-for memory (accessibility theory) was also supported.

Key words: feeling of knowing, odor identification, odor naming, tip of the tongue experience, tip of the nose experience

Introduction

Why is it so difficult to name even the most common odors? The average rate of correct naming of a set of common odors (e.g. coffee, vanilla and tar) rarely exceeds 50% (de Wijk *et al.*, 1995). Is it due to the often proposed poor activation of the name of the odorous object (Engen, 1987, Herz and Engen, 1996)? Or is it because the odorous object is not successfully identified? When we fail to name an odor, is the feeling of knowing (FOK) the odor or the recognition of the odor as one previously encountered indicative of actual knowledge about the odor (and in comparison to other modalities)? Chobor (1992, p. 357) argued that strong feelings of imminent retrieval of an odor name are very common, both in everyday life and in the clinical setting. Lawless and Engen (1977) named this phenomenon the 'tip of the nose' (TON) phenomenon, as an olfactory parallel to the more known tip of the tongue (TOT) experience (Brown, 1991; Schwartz, 2002). The above are some of the main questions of interest in the current study and we use a new metacognitive approach to investigate them. In sum, the aims of the current study were to investigate where in the naming process odor naming fails and to further investigate metamemory judgements following odor naming failures. In addition, the olfactory modality is compared with the visual modality, here pictures of famous persons. The aims of the study will now be presented in more detail.

Naming of objects, such as pictures, has been proposed to consist of three separate stages: (i) object identification; (ii) name activation; and (iii) response generation (McCauley *et al.*, 1980; Johnson *et al.*, 1996). Similar stages in the naming process have also been proposed for person naming (Burton and Bruce, 1992). Holley (2002) argued that, in nature, the function of odors is to reveal the presence of objects and substances in the environment, and for the layperson, odor naming is essentially odor-source naming (i.e. object naming). Therefore these naming stages should be applicable also for odor naming. When something is identified (e.g. an apple is identified via its smell) the participant knows what it is regardless of naming ability. After it has been identified, the object's name is (possibly) activated and, if it is the correct one, the participant generates the response 'apple'. We here propose that strong feelings of knowing an unnamed odor or TON experiences, as well as odor naming failures in general, are about the first of these stages (object identification), not the second (name activation). To clarify, *it is proposed that people fail to name odors due to failures to identify the odorous objects rather than failures to activate the name of an already identified object.* This is in contrast to the typical TOT experience that presumably arises in the 'name activation' phase. The proposition is also in contrast with the proposition that there is a weak connection between the odor and its name

(Herz and Engen, 1996). We test this hypthesis in Experiments 1 and 2, in which we use a procedure whereby the participants can distinguish between when they only lack the name of an identified object (or person) and when they have not even identified it.

Further, we compared the odor-elicited responses with those elicited by pictures of famous persons. Burke *et al.* (2004) suggested that the link between persons and their proper names are particularly susceptible to weak connections, especially among the elderly. This argument is based on the fact that, although you might have much semantic information about a person other than the name (e.g. occupation; Yarmey, 1973), two persons with the same name do not necessarily share the same semantic features. Odor naming is also quite fallible (e.g. de Wijk and Cain, 1995) and odors have also been suggested to have weak links with their names (Engen, 1987, 1991; Herz and Engen, 1996). In addition, odor naming has, at least in the olfactory literature, been highlighted as similar in many respects to person naming. Cain and Gent (1986, p.175; see also Murphy *et al.*, 1991) claimed that 'odor identification has considerable similarity to the identification of faces' and 'fits the pattern of face identification extremely well'. They suggested that both odors and faces encourage holistic encoding. Murphy *et al.* (1991) argued that in everyday life the layperson encodes an odor for 'what it is' (e.g. chocolate, strawberry, orange), rather than by an analyses of its features, and similarly for faces. Another similarity has to do with the difficulty in naming. Cain and Gent (1986) argued that both odors and persons are often highly familiar, but with a concomitant inability to retrieve the proper name. So, whereas odors seem particularly apt in eliciting strong feelings of knowing, or TON experiences (Jönsson and Olsson, 2003), faces on the other hand seem particularly apt in eliciting TOT experiences (Yarmey, 1973). However, persons may also be familiar (or recognized), but without any activation of semantic information about the person (Hanley and Turner, 2000; Hanley *et al.*, 1998). For these reasons pictures of famous persons were chosen as comparison modality to the odors.

It was hypothesized that identification failures should be more common for odors than for pictures of famous persons. The first experiment focused on TOT and FOK judgements, whereas the second experiment solely investigated FOK judgements. A FOK judgement is typically defined as a feeling that an unretrieved memory (e.g. a word) is to be retrieved or recognized. FOK judgements are more general than TOT experiences because they can range from no FOK to strong FOK, whereas TOT experiences are by definition strong feelings of subsequent retrieval.

Odor naming versus identification

Most studies of our ability to identify odors have used odor naming performance as an indicator of odor identification. Therefore, the term odor naming has in these studies been synonymous to odor identification. However, an odor can be identified by other means than its proper name. In the current study, we stress the difference between the terms odor *naming* and odor *identification*. Odor naming here reflects the ability to assign the correct label to an odor, whereas odor identification refers to the ability to identify the odor by any means (i.e. to know what it is, independent of naming ability). In the current study the participants were in the case of a naming failure asked to separate between two states of knowledge: (i) when they lack the name and have not identified the stimulus presented; and (ii) when they have identified the stimulus and only lack the name.

Support for the hypothesis that odor naming failures typically are due to identification failures has several important consequences for the understanding of (i) the ability of odors to trigger TOT experiences, as well as (ii) odor naming, (iii) the predictive validity of metamemory judgements about odor names and (iv) metamemory theory; in the current context, the understanding of the basis of metamemory judgements for odor names. These points will now be further elaborated upon.

TOT experiences

Examples of typical procedures to elicit TOT experiences are to present participants with word definitions to cue target words (Brown and McNeill, 1966) or pictures of famous persons to cue their names (Yarmey, 1973; Burke *et al.*, 2004). A TOT experience is commonly defined as a strong and imminent feeling of being on the verge of retrieving a temporarily inaccessible word. A distinctive feature of TOT experiences is the often found presence of structural–phonological information of relevance to the unretrieved target word (Brown, 1991). This may be the first letter (e.g. the missing word begins with an O, as in orange), other letters, similar sounding words (e.g. 'Lockhart' for the actress Calista Flockhart) or other information disclosing partial access to the missing word. Koriat *et al.* (2003) divided the partial clues people have about the missing word into two categories, namely, the just mentioned structural–phonological clues and semantic clues (e.g. 'it is a fruit'). As Koriat and colleagues pointed out, TOT research has mostly focused on the first category. They demonstrated that when a sought-for word is missing, people sometimes also have access to semantic information about the word. Significant for strong feelings of imminent recall in the case of odors are the absence of structural–phonological information (Lawless and Engen, 1977; Jönsson and Olsson, 2003). However, even when odor name retrieval fails, participants can sometimes categorize the odors correctly (e.g. fruit for lemon), indicating the availability of semantic information.

As previously noted, failures to name odors were hypothesized to predominantly be caused by failures to identify the odorous objects by smell. As a consequence, we hypothesized that TOT experiences, in the current study defined

as feelings of imminent retrieval of the name of an already identified odor or person, should be very infrequent for odors, and more so than for persons. In other words, odors should elicit fewer TOTs because odor naming failures are rarely due to failures in the name activation phase.

Odor naming

Our inability to name odors is well documented (Desor and Beauchamp, 1974; Cain, 1979; de Wijk and Cain, 1994a,b; de Wijk *et al.*, 1995; Cain *et al.*, 1998; Olsson and Fridén, 2001), but its cause is not fully understood. Herz and Engen (1996, p. 301) described it as the 'most contentious issue in human olfactory processing'. A common statement is that there is a poor link between the odor and its name, meaning that the verbal areas of the brain are poorly associated with the olfactory processing areas in comparison with other modalities (Engen, 1987, 1991; Herz and Engen, 1996). Other researchers have argued that our inability to link an odor to its name is not inherent but rather due to how we learn odors (de Wijk *et al.*, 1995). Learning to associate odorous objects with their proper names is not as formalized in society (e.g. in school) as is naming of visual objects. In addition, de Wijk *et al.* argued that in everyday life, odors are typically experienced in specific contexts, and what we smell is to a large extent interpreted based on the contextual information at hand. In a recent paper Sulmont-Rossé *et al.* (2005) argued that a part of the low performance found in odor naming tasks may be due to the lack of social consensus around the odor names; that is, people encounter certain odors in different contexts (or products), leading to different object associations to those odors. For example, in their experiment several participants consistently labeled artificial flower odors as 'cleaning supply' or 'bathroom freshener'. Although this was not the veridical name expected by the experimenters, they may still be considered as correct descriptions of the odorants. This is because these products are often fragranced with such odorants. Köster (2002) stressed that detection, discrimination and recognition of odors as familiar or unfamiliar is much more important in real life than odor identification or assessment of odor intensity. However, there are several occasions when accurate identification of odors is advantageous, if not essential. Examples are the accurate identification of a kitchen gas leak, spoiled food or fire, which all could have detrimental consequences if not understood (White and Kurtz, 2003).

The absence of structural–phonological information in connection with odor naming failures could be seen as support for the idea of a poor odor–name link. However, another interpretation is that when people report strong feelings of subsequent retrieval, it is not the activation of the name that is at fault; rather, they have not yet identified the odor. Actually, Cain and Potts (1996) argued that errors in odor naming are predominantly perceptual, since when their naming attempts were incorrect, participants typically did not find themselves to have problems coming up with the label. Instead, they gave labels that were close enough to suggest simple errors of discrimination. Indeed, Olsson and Cain (2000) demonstrated that the degree to which an odor needed to be substituted with another fairly dissimilar odorant (eugenol and citral were used) in order to reach a just noticeable difference in perceived odor quality was as much as 30% in liquid phase (see also Wise *et al.*, 2000). Moreover, Cain *et al.* (1998) showed how odors that were successfully named at one point (43% of the trials in their study) were not necessarily named correctly again 2 days later (10% of these were misnamed). Several studies have also found high correlations between odor naming and discrimination performance (Eskenazi *et al.*, 1983, 1986; de Wijk and Cain, 1994a,b; Cain *et al.*, 1998). Altogether, these observations indicate that discrimination is also at fault in odor naming tests.

Regarding our study, if found that odor naming typically fails in the name activation stage, it would support the poor link hypothesis. On the other hand, if odor naming typically fails already in the identification stage, it would favor the interpretation that it is to a larger extent due to odor identification failures.

Predictive validity

Another aim of the study was to compare the predictive validity of the metamemory judgements. The predictive validity is the degree of relationship between the participants' metamemory judgements and their actual retrieval or recognition of the missing name following those judgements. Cain *et al.* (1998) compared FOK judgements about odor names with FOKs for answers to general information questions, and found that the odor-elicited FOKs were less predictive of subsequent recognition in an eight-alternative forced-choice recognition test. The current study did further cross-modal comparisons and it was hypothesized that the predictive validity of the olfactory metamemory judgements would be less accurate than those of the person names.

Metamemory theory

The two main metamemory theories that have been proposed to explain the underlying basis of FOK judgements are the cue familiarity theory (Reder, 1987; Metcalfe *et al.*, 1993) and the accessibility theory (Koriat, 1993, 1995). The cue familiarity theory assumes that people base their metamemory judgements on the perceived familiarity of the cue, e.g. the familiarity of an odor that is used to cue its proper name. The accessibility perspective assumes instead that metamemory judgements are based on the sheer amount of partial information present in memory about the unrecalled target memory (independent of its correctness). Without putting the cue familiarity theory to test, Jönsson and Olsson (2003) found some support for the accessibility perspective when investigating TON experiences. Recently,

Koriat and Levy-Sadot (2001) demonstrated that the familiarity of the cue triggering the memory and the accessibility of information about that memory do interact. That is, the theories are not mutually exclusive. They showed that a certain degree of familiarity is necessary to activate search and retrieval of accessible information.

In the second experiment of the current study a familiarization procedure of half the odors and pictures was used to investigate the role of cue familiarity as an underlying basis for the FOK judgements. It was expected that the familiarization procedure should increase the strength of the FOKs as compared with the stimuli that had not been familiarized. The role of access to partial information was also investigated. It was hypothesized that the more identifications reported, the stronger the FOK judgements.

Experiment 1

The aims of this experiment were (i) to investigate where in the naming process odor naming fails; (ii) to investigate the ability of common odors to elicit TOT experiences; and (iii) to compare the olfactory modality with another modality, namely pictures of famous persons. Following naming failures of odors and famous persons, the participants judged whether they had or had not identified the stimulus presented. If the stimulus was identified, they judged whether they had a TOT experience or not. If it was not identified, they instead made a FOK judgement about how sure they were that they would be able to identify the odor (by the veridical label or by a description).

Method

Participants

Forty participants (20 women) with a mean age ± SD of 24.10 ± 2.59 years (range = 18–31 years) were recruited from 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

Thirty common odorants and 30 pictures of famous people were used as test stimuli (see Appendix Table A1). Some odorants (e.g. apple and lemon) were changed regularly to keep them fresh and the odor quality stable. The odorants were all commonly encountered stimuli (i.e. not artificial odorants) except for the three essences violet, peppermint, and geranium. Odorants were presented in 160 ml tinted glass jars with screw lids. Cotton pads prevented visual inspection of the stimulus material in the jars. The pictures all consisted of famous people from different professions (e.g. politician, singer or actress) and were presented to the participants on a 17″ computer screen. Because the current experiment entailed a comparison between stimuli from two different modalities, it was important to ensure that they were equally difficult to name. The odors and pictures were therefore pre-standardized and matched for naming difficulty in a pilot experiment. This was done in the following way: first, 73 pictures of famous persons were chosen, shown to a group of 41 participants (mean age = 24.90 ± 6.00; range = 19–48), and tested for naming difficulty. Then 32 of the pictures were chosen to match the naming difficulty (mean and variance), of 32 odorants, of which we already had naming data from previous studies. These two stimulus sets were then used in a pilot experiment with a similar methodology to Experiment 1 (31 participants with a functional sense of smell; mean age = 27.45 ± 6.62; range = 20–49). The overall naming performance in that experiment was 0.25 ± 0.14 for the pictures and 0.34 ± 0.09 for the odorants [$t(30) = 3.44$; $P = 0.002$]. Because the odorants were easier to name, the two easiest odorants and the two most difficult pictures were removed and two equally difficult stimulus sets of 30 was created [odors, 0.30 ± 0.09; pictures, 0.26 ± 0.14; $t(30) = 0.1.42$; $P = 0.17$]. In both pilot experiments each trial consisted of a maximum of 45 s and the trials were always ended when a participant retrieved the sought-for name.

Procedure

All participants were tested individually. After being welcomed, they were seated in front of a table with a 17″ computer screen on it. The experimenter sat at the same table and could always see the participants. The odor jars and the other computer equipment (apart from the screen) were placed in front of the experimenter and were hidden from the participants by a screen. First the participants filled in a questionnaire with some background data, such as age, sex and if they had a functional sense of smell. Then they read the instructions.

The experiment consisted of two conditions: a picture and an odor condition. The pictures of famous persons were presented to the participants on the computer screen, whereas the odor jars were handed to them by the experimenter. During the odor trials the computer screen was blank. The presentation order of the stimuli was fully randomized for each participant, with the only exception being that every second trial was an odor and every other a picture trial. The experiment started with two practice pictures and two practice odors, and the experimenter ensured that the participants had understood the instructions. Each odor and picture could be sampled repeatedly during the allotted time (i.e. the odor could be smelled as much as needed and the picture was visible all the time). A maximum of 90 s was given for each trial. The whole experiment took on average 76 ± 14 min.

For each trial the participants had a questionnaire in front of them that consisted of three separate parts, section A, B and C. For each trial they only filled in one of these sections. The main procedure is graphically demonstrated in Figure 1. The task was to smell the odor or look at the picture and try to name it or him/her. Regarding the pictures, the name targeted was only the surname. If the participants thought that they

immediately could name the stimulus (henceforth called *immediate naming*), they filled in the name in section A of the questionnaire and the experiment continued with the next trial. If they could not immediately name the stimulus, they instead filled in either section B or C of the questionnaire. They filled in section B if they thought they had identified the stimulus and section C if they had not yet identified it. The latter two sections will now be explained in further detail.

Section B: TOT experiences. They filled in this section if they thought they had identified the stimulus but could not retrieve the name. They first judged whether they had a TOT experience (by marking if it was strong or very strong) or not. Then they continued to search for the name. The trial always ended either when they retrieved the name or after 90 s. The distinction between knowing (section B) and not knowing (section C) the identity was extensively explained and exemplified in the instructions. Here is a translated excerpt of the instructions for section B, including the TOT instructions:

> If you know who a person is you can normally place him/her in the right context, for example, that a person is married to a friend of you or that the person is an actor in a movie you have seen. If you know which odor it is you can maybe visualize the object or go to the store and buy the product. That is, you know precisely what or who it is, you just have not yet retrieved the name. It is not enough that the person or the odor is very familiar; you also have to know who he/she is or what it is.
>
> If you have the name on the tip of the tongue you feel that it is possible to retrieve the name and that you will do so soon. The name is on its way to emerge from memory. Sometimes you can feel frustrated and maybe emotional over that the name you seek escapes you. If you have such a feeling of that you are on the verge of retrieving the name from memory, you should indicate if the experience is strong or very strong. If you know who it is/what it is without having the name on the tip of the tongue you should instead mark this.

Section C: FOK judgements. If the participants neither could name nor felt that they had identified the odor or person, they first filled in whether they had (i) no or a weak FOK, (ii) a strong FOK or (ii) a very strong FOK, indicating that they would retrieve who or what it was (either via a description or via the proper name). An excerpt of the instructions for this section follows below:

> Sometimes an odor or a person might feel everything from not familiar to very familiar, but you do not know exactly who it is/what odor it is ... You can not quite place the person in his/her right context or you can not imagine the object from which the odor is emanating. It can be that you are unsure of whether an odor comes from, for example, asphalt or tar.

As can be seen in Figure 1, the participants discontinued filling in the questionnaire as soon as they retrieved the sought-for name or if they had not retrieved it within 90 s from the start of the trial. In both sections B and C, and while searching for the proper name of the stimulus, the participants were also asked to fill in a possible profession (e.g. actress) of the person or, if it was an odor, to categorize it (e.g. fruit).

Results and discussion

An alpha level of 0.05 was used for all statistical tests. Missing values were in the analyses handled by using casewise deletion. When applicable, Cohen's *d* was used to denote the effect size for the difference between two means. As can be seen in Table 2, the proportions of non-TOTs, strong and very strong TOT experiences (section B) were very low for the odors, and the number of strong to very strong

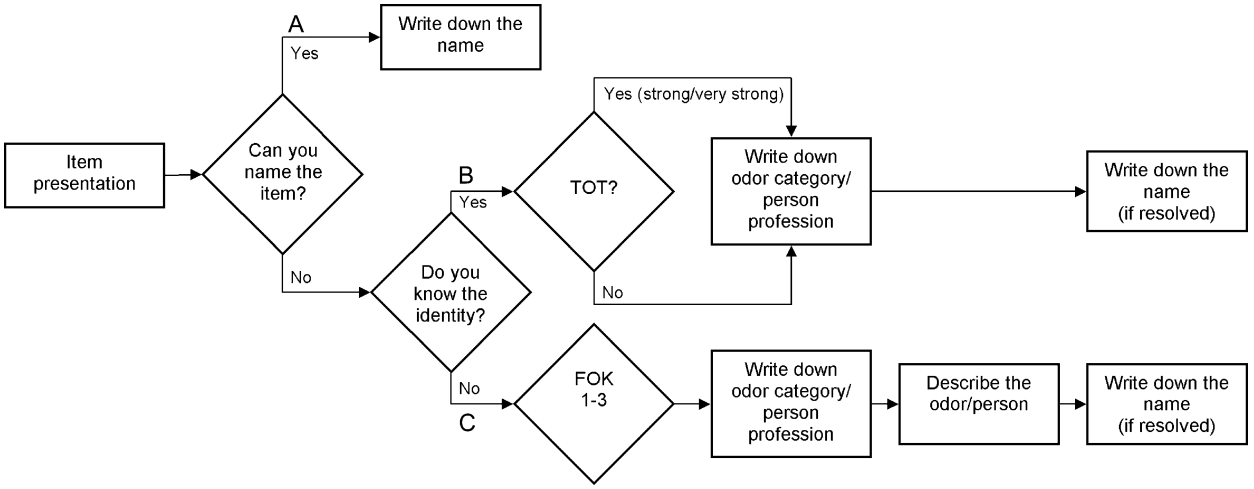

Figure 1 The procedure of Experiment 1 is shown. The letters A (immediate naming), B (name unretrieved, but stimulus is identified) and C (unidentified stimulus and name unretrieved) denotes the three separate parts of the questionnaire. For each trial only one part was filled in. All trials always ended after 90 s or sooner if the name was retrieved.

FOK judgements (section C) was very low for the pictures. For that reason it was generally not viable to do cross-modal comparisons separately for the two sections as a function of metamemory strength.

Overall naming

The measure of overall difficulty of the two stimulus sets was the proportion correctly named stimuli, independent of whether they were immediately named or preceded by a metamemory judgement and resolved later. This is henceforth referred to as *correct overall naming*. A paired *t*-test showed that, overall, the odors and pictures of famous persons were equally difficult to name [$t(39) = 0.37$; $d = 0.07$; $P = 0.71$]. This means that the pre-standardization of the odors and pictures was successful. The results are presented in Table 1 (line 1), which also presents the results from Experiment 2.

Immediate naming

The participants were instructed that if they immediately could name the stimulus they should do so; otherwise they should instead continue to search for the name and fill in section B or C of the questionnaire (Figure 1).The odors were less accurately named than the persons. This is shown by the significant difference in proportion correct immediate naming in Table 1 (line 2), [$t(7) = 9.76$; $d = 2.21$; $P < 0.0001$]. Note the difference between this measure of correct immediate naming (the proportion correct responses calculated for the immediate naming attempts only) and the above correct overall naming (calculated on all trials).

We also calculated the proportion out of all trials that led to immediate naming attempts (line 3 in Table 1). Out of the 30 presentations for each modality, they attempted to name slightly more odors than persons, but the difference only reached the level of a statistical tendency [$t(39) = 1.94$; $d = 0.44$; $P = 0.06$].

Subjective identifications

For all unnamed stimuli we calculated the proportion of these that were identified (referred to as subjective identifications) (line 4 in Table 1). For each modality, the number of identifications (section B responses in the questionnaire) was divided by the total number of naming failures (section B + C responses). These analyses therefore excluded the immediate naming trials. The unnamed pictures were identified significantly more often than the unnamed odors [$t(39) = 11.93$; $d = 2.46$; $P < 0.001$]. The number of subjective identifications were indeed very low for the odors ($n = 67$ out of totally 850 responses). This is in line with the hypothesis that odor naming failures are typically due to failures to identify the odors.

TOT experiences and FOK judgements

It was hypothesized that odor-elicited TOT experiences should be unusual and more unusual following failures to name odors than persons. For each participant and modality the number of strong to very strong TOT experiences was divided with the total amount of naming failures. This was the dependent variable. A paired *t*-test showed that the pictures elicited significantly more TOT experiences (0.24 ± 0.14) than the odors [0.07 ± 0.07; $t(39) = 7.42$; $d = 1.54$; $P < 0.0001$]. The proportion of strong to very strong FOK judgements was calculated in a similar fashion. Strong to very strong FOK judgements were more common for odors (0.55 ± 0.17) than for pictures [0.06 ± 0.07; $t(39) = 17.86$; $d = 3.77$; $P < 0.0001$]. In Table 2, a more detailed presentation of how the metamemory judgements were distributed is shown.

In a review of the TOT phenomenon, Brown (1991) noted that the typical TOT incidence across studies is between 0.08 and 0.18 (mean: 0.13). However, in those studies the TOT incidence was calculated as the proportion TOTs of *all* trials, not as a proportion of the unnamed trials only. We therefore re-calculated the TOT incidences for the two modalities accordingly, confirming that TOTs for odor names are very unusual [odors, 0.05 ± 0.05; pictures, 0.18 ± 0.10; $t(39) = 7.57$; $P < 0.0001$]. This latter measure is the one to be used when comparing with other TOT studies. To conclude, whereas unnamed odors elicit few TOT experiences they do elicit many strong to very strong FOKs. These strong FOKs

Table 1 (1) Proportion correct overall naming, (2) proportion correct immediate naming, (3) proportion immediate naming attempts and (4) proportion subjective identifications for Experiments 1 and 2

	Experiment 1		Experiment 2	
	Odor	Picture	Odor	Picture
1. Correct overall naming	0.30 (0.12)	0.29 (0.17)	0.29 (0.12)	0.27 (0.20)
2. Correct immediate naming	0.68 (0.16)**	0.96 (0.08)	0.59 (0.20)**	0.87 (0.18)
3. Immediate naming attempts	0.28 (0.14)*	0.22 (0.16)	0.44 (0.16)**	0.28 (0.19)
4. Subjective identifications	0.08 (0.09)**	0.43 (0.18)	0.25 (0.20)**	0.46 (0.18)

Correct overall naming and immediate naming attempts are calculated as proportions of all trials (i.e. including the FOK and TOT trials). Correct immediate naming at line 2 refers only to the proportion of the immediate naming attempts that were correctly named. Subjective identifications at line 4 refer to the proportion of the unnamed stimuli that were reported to be identified. The latter proportion thus excludes the immediate naming attempts.
*$P = 0.06$; **$P < 0.0001$

Table 2 How the metamemory judgements were distributed

Modality	Metamemory judgements					
	Non-TOT	Strong TOT	Very strong TOT	weak FOK	strong FOK	very strong FOK
Odor	0.02 (0.04)	0.03 (0.04)	0.04 (0.05)	0.37 (0.16)	0.30 (0.12)	0.24 (0.14)
Picture	0.19 (0.14)	0.09 (0.10)	0.15 (0.13)	0.51 (0.17)	0.05 (0.06)	0.01 (0.02)

Values are mean proportions (SD) across participants and are calculated on the basis of all trials that led to naming failures (e.g. on average, 24% of the odor naming failures consisted of very strong FOK judgements).

for odors are typically associated with object identification failures rather than name activation failures. The present data also inform us about the often cited TON experience. Based on the distribution of the metamemory judgements in Table 2, it can be concluded that the TON experience is actually a mix of mainly FOK judgements about the retrieval of odor identity and a few TOT experiences.

Other semantic information

Following the naming failures, the participants were asked to give a plausible profession (e.g. politician or singer) for the persons and a plausible category for the odors (e.g. fruit or spice). The proportion correct such responses were analyzed. The participants significantly more often knew the persons' professions (0.40 ± 0.16) than they could categorize the odors [0.29 ± 0.15; $t(37) = 7.34$; $d = 1.41$; $P < 0.001$]. It should be noted that the criterion for scoring an odor category as correct was liberal, making this difference conservative. The results are in line with the participants' subjective identity judgements, that they had more often identified the persons than the odors.

Experiment 2

The experiment aimed to perform a cross-modal comparison of the predictive validity of FOK judgements. Only one such cross-modal study exists in the literature (Cain *et al.*, 1998). Another aim was to replicate the finding of Experiment 1, indicating that when participants make strong predictions of recalling odor names they are typically further back in the naming process than is the case for person names.

Yet a third aim targeted metamemory theory—more precisely, the cue familiarity and accessibility theories. Familiarity has in studies of other modalities been found to be related to FOK judgements (Reder, 1987; Metcalfe *et al.*, 1993). Jönsson and Olsson (2003) demonstrated that accessibility does seem to play a role because participants' odor categorizations were better, the higher their predictions of subsequent odor name retrieval. However, they hypothesized that due to the overall low level of knowledge the participants had about the odors presented to them, the familiarity might play an important role as the underlying basis for FOK judgements about odor names. We here used a familiarization procedure to evaluate the role of cue familiarity and in comparison with the reference modality. It was expected that the familiarization procedure should increase the strength of the FOKs as compared with the non-familiarized stimuli. Also, the accessibility perspective was further evaluated. It was hypothesized that the more subjective identifications reported by the participants, the stronger the FOK judgements.

Method

This design was similar to that of Experiment 1, but with some important differences. Here, only FOK judgements about the missing names were gathered, the time for each trial was decreased from 90 to 60 s and a new questionnaire was used. Instead of judging whether they had identified the stimulus before the metamemory judgements, this was now done after those judgements.

Participants

One person reported a poor sense of smell and was excluded from the analyses. The remaining 38 participants (20 women) had a mean age of 25.82 ± 6.02 years (range = 19–48 years) and were all recruited from Uppsala University. They participated for course credits or were given a movie ticket voucher worth ~75 SEK.

Stimuli

The experimental setting and the stimuli were identical to those in Experiment 1 (see Appendix Table A1), with the only exception that the odorants peppermint, vinegar and Tabasco were exchanged for caraway, tar and cocoa. The odorants were all non-artificial odorants except for the essences violet and geranium.

Procedure

The participants were tested individually. The experiment consisted of three separate phases: (i) familiarity judgements of the stimuli (familiarization); (ii) a short break with filler activities; and (iii) the test phase. To facilitate comprehension, the procedure of the test phase is graphically presented in Figure 2. After being welcomed and seated in front of a desk with a computer screen on it, the participants were given written instructions about the first phase of the experiment (familiarization). The task was to smell each odor or watch each picture during 3 s, and rate the familiarity of the stimulus immediately afterwards. This was done on a scale

ranging from 1 (not at all familiar) to 10 (extremely familiar). A total of fifteen odors and fifteen pictures were presented. The odors and pictures were individually randomized for each person from the set of 30, meaning that each person was presented a set of 15 unique odors and pictures. The presentation order was also randomized, with the only exception that odors and pictures were always presented interchangeably. After all the stimuli had been presented, the same procedure was repeated a second time with the same stimuli but in a new randomized order of presentation. During the second phase the participants filled in a questionnaire with some background data, such as age and sex, as well as the Affective Impact of Odors (AIO) questionnaire (Wrzesniewski *et al.*, 1999). The latter was used as a filler activity between the sessions and the results are not treated here.

The final phase (test) started with the participants reading written instructions. As noted, the main procedure of this phase is outlined in Figure 2. The total sets of 30 pictures and 30 odors were presented one at the time, 15 of them from each modality being repeated from the first phase (old) and 15 being new. The task was to smell each odor or watch each picture and try to name it/him/her (immediate naming). A maximum of 15 s was allotted to this task. If a naming attempt was made, the experiment continued with the next trial. If the participants could not name the stimulus during the allotted time, they instead made a FOK judgement about how sure they were of subsequent retrieval of the correct name (within ~60 s, with the time spent for the immediate naming attempt subtracted). The FOK judgement was made on a scale from (1) no/weak, (2) average strength, (3) strong to (4) absolutely sure. Immediately after the FOK judgement they judged whether they had or had not identified the unnamed stimulus. This judgement was identical to that of Experiment 1. If they thought that they retrieved the correct name for the stimulus, they wrote it down, interrupted all further activities and continued with the next trial. The presentation order for the stimuli was randomized for each individual, but odors and pictures were presented interchangeably. Each odor and picture could be sampled repeatedly during the allotted time (i.e. the odor could be smelled as much as needed and the picture was visible all the time).

The experiment started with two practice odors and two practice pictures, not included in the analyses. Spoken instructions complemented the written ones to ensure that the participants understood. The whole experiment took on average 80 ± 10 min.

Results and discussion

An alpha level of 0.05 was used for all statistical tests. Missing values were in the analyses handled by using casewise deletion. When applicable, Cohen's *d* was used to denote the effect size for the difference between two means.

First phase

Familiarity judgements. In the first phase, the participants judged the familiarity of half the stimulus material twice. A modality (odor/picture) × presentation (1st/2nd) repeated measures analysis of variance (ANOVA) showed that the odors (6.39 ± 1.29) and pictures (6.30 ± 1.62) were equally familiar [$F(1,37) = 0.12$; $MS_e = 2.62$; $P = 0.73$]. A significant main effect of presentation showed that the second presentation (6.72 ± 1.53) was associated with higher familiarity ratings than the first [5.96 ± 1.09; $F(1,37) = 20.29$; $MS_e = 1.09$; $P < 0.0001$]. No significant interaction was observed.

Test phase

Overall naming. A modality × type (old/new; where old represents the stimuli that were also presented in the first phase) repeated measures ANOVA was performed to test if the odors and pictures were equally difficult to name. The average proportion of correctly named stimuli, independent of them being named immediately or preceded by a FOK judgement, was non-significantly different for the two modalities [see Table 1; $F(1,37) = 0.52$; $MS_e = 0.04$; $P = 0.48$]. This means that the odor and picture sets were well matched for difficulty of naming. A main effect of type was observed, with old stimuli (0.31 ± 0.15) being better named than new

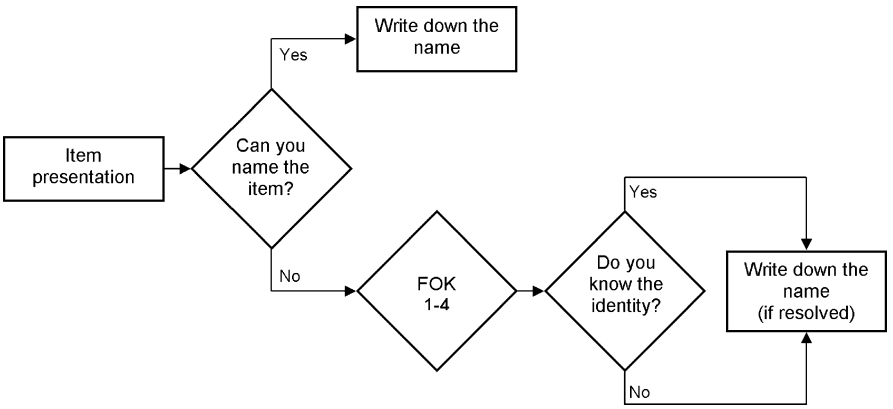

Figure 2 The main procedure of the test phase of Experiment 2 is shown.

[0.25 ± 0.13; $F(1,37) = 8.33$; $MS_e = 0.02$; $P = 0.01$]. This suggests that it was easier to name the stimuli after pre-exposure. However, there was also a significant modality × type interaction [$F(1,37) = 7.10$; $MS_e = 0.01$; $P = 0.01$]. Whereas pairwise comparisons showed no difference between old (0.30 ± 0.15) and new (0.29 ± 0.14) odors [$t(37) = 0.34$; $d = 0.07$; $P = 0.74$], the old pictures (0.32 ± 0.24) were significantly better named than new [0.22 ± 0.19; $t(37) = 4.01$; $d = 0.46$; $P = 0.0003$]. To conclude, the familiarization procedure resulted in better naming performance of the persons, but not the odors.

Immediate naming. The participants named the stimuli directly if they thought they could. Two modality × type repeated measures ANOVAs of (i) the mean proportion of correct immediate naming and (ii) the mean proportion of immediate naming attempts showed the following (Table 1).

The analysis of proportion correct immediate naming, calculated as a proportion of the immediate naming attempts only, showed that the participants were much poorer in naming the odors than the pictures [see Table 1; $F(1,29) = 51.81$; $MS_e = 0.04$; $P < 0.0001$]. There were no other main or interaction effects. This replicates the results of Experiment 1.

The proportion immediate naming attempts were significantly higher for the odors than the persons [$F(1,37) = 20.37$; $MS_e = 0.04$; $P < 0.0001$]. A main effect of Type showed that familiarized stimuli more often led to direct naming responses than those not familiarized [old, 0.39 ± 0.16; new, 0.33 ± 0.14; $F(1,37) = 11.20$; $MS_e = 0.01$; $P = 0.002$]. There was a significant modality × type interaction [$F(1,37) = 6.18$; $MS_e = 0.01$; $P = 0.02$]. Post-hoc comparisons of the means showed no difference between the proportion immediate naming attempts for the new (0.43 ± 0.16) and old (0.45 ± 0.19) odors [$t(37) = 0.83$; $d = 0.11$; $P = 0.41$], but previous exposure had a significant effect on the pictures. Newly presented pictures (0.24 ± 0.19) were significantly less frequently named than old (0.33 ± 0.23) pictures [$t(37) = 3.87$; $d = 0.43$; $P = 0.0004$]. This is consistent with the finding that the overall naming performance increased only for the pictures. Also, both new and old odors were more often immediately named than both old and new pictures (all Ps < 0.05).

To conclude, the participants more often thought they could immediately name the odors than the persons, as shown by the higher amount of such naming attempts, but did so less accurately.

Cue familiarity. To investigate the role of familiarity as an underlying basis for odor-elicited FOKs (in comparison to the reference modality) a modality × type repeated measures ANOVA was performed with mean FOK strength as the dependent variable. It was expected that the familiarization procedure would increase the FOKs for the familiarized (old) stimuli as compared with the new. The participants made significantly higher FOK ratings for the odors (2.11 ± 0.54) than for the pictures [1.67 ± 0.49; $F(1,37) = 18.36$; $MS_e = 0.39$; $P = 0.0001$]. However, the familiarization procedure failed to have any effect on FOKs for either modality. No main effect of type, nor any interaction was present in the analysis (both Fs < 1).

We also analyzed proportion correct naming following the FOKs as a function of modality × type. This analysis showed a statistical tendency towards odor-elicited FOKs (0.08 ± 0.08) being better resolved than those elicited by the pictures [0.05 ± 0.06; $F(1,37) = 3.27$; $MS_e = 0.01$; $P = 0.08$]. Although no main effect of type was present, there was a significant modality × type interaction [$F(1,37) = 5.08$; $MS_e = 0.09$; $P = 0.03$]. Simple comparisons of the means showed that whereas the previously presented odors and pictures did not differ [odors, 0.07 ± 0.10; pictures, 0.07 ± 0.12; $t(37) = 0.12$; $d = 0.00$; $P = 0.91$], the newly presented ones did [odors, 0.09 ± 0.13; pictures, 0.03 ± 0.05; $t(37) = 2.57$; $d = 0.61$; $P = 0.01$]. As a matter of fact, the newly presented pictures were associated with a significantly lower naming performance as compared with all the three other means (all Ps < 0.05).

In sum, the FOKs were for both modalities unaffected by the familiarization procedure, which means that the cue familiarity theory was not supported. It is not clear whether this should be taken as an argument against cue familiarity being important or if it rather mirrors an unsuccessful familiarization of the stimuli. The initial high familiarity of the stimuli (common odors, famous persons) could have counteracted further familiarization. With regards to correct naming performance following the FOKs, old pictures were better named than new ones, but odors were unaffected by the familiarization procedure.

Correlation between familiarity and FOK. Because the familiarization procedure failed to affect the FOKs, a second analysis was made targeting the relationship between familiarity and FOKs. First, for each trial in the first phase, the mean of the two familiarity judgements was calculated and then correlated with the FOK strength in the test phase. As a consequence, only the repeated stimuli associated with a FOK judgement could be included in the analysis. A gamma correlation was calculated for each participant and modality. The mean gammas showed that for both modalities, there was a high familiarity–FOK correlation, but it was significantly higher for pictures (0.74 ± 0.24) than for odors [0.51 ± 0.36; $t(31) = 3.10$; $d = 0.75$; $P = 0.004$]. It is unclear why the correlation was higher for the persons. One possible reason could be that the measurement error in the familiarity ratings was not identical between the two modalities. Because the participants made two consecutive familiarity ratings (on half the stimuli) in the first phase, we could calculate the test–retest reliability of those ratings. For each participant we calculated a Pearson correlation between presentations 1 and 2, and a paired t-test showed that the mean Pearson correlation was significantly lower for the odors (0.66 ± 0.18) than the pictures [0.79 ± 0.22; $t(34) = 2.79$; $P = 0.01$]. The fact that the familiarity ratings were more reliable for the pictures could therefore, at least partially, explain why the correlation between familiarity and FOK was higher for the pictures.

Subjective identifications. Directly after making a FOK judgement the participants judged whether they had or had not identified the stimulus in question. Due to fragmentary data, FOK strengths 1 and 2 were merged to FOK_{12} and FOK judgement 3 and 4 to FOK_{34}. Then the proportion subjective identifications for the two FOK categories were calculated for each participant. A modality × FOK strength repeated measures analysis of variance showed that the participants had identified the pictures significantly more often than the odors [see Table 1; $F(1,27) = 28.71$; $MS_e = 0.04$ $P < 0.0001$]. A main effect of FOK strength showed that the participants identified the stimuli more often when the FOK judgements were strong (0.64 ± 0.29) than when they were weak [0.06 ± 0.12; $F(1,27) = 178.42$; $MS_e = 0.05$; $P < 0.0001$]. This means that the stronger the FOK, the more information was at hand about the stimuli, which is in line with the idea that FOKs are based on the amount of accessible information about the unnamed stimulus (Koriat, 1993, 1995). Note that if the participants reported that they had identified a stimulus, this does not necessarily imply that the identification was correct. However, the accessibility theory states that FOKs are based on the amount of accessible information independent of its accuracy.

A significant modality × FOK strength interaction [$F(1,27) = 16.10$; $MS_e = 0.04$; $P < 0.0001$] and simple comparisons of the means showed that for the stronger FOK judgements, pictures (0.82 ± 0.26) were associated with identification responses significantly more often than odors [0.47 ± 0.35; $t(27) = 4.90$; $d = 1.14$; $P < 0.001$]. Although this was also true for the weaker FOKs [pictures, 0.09 ± 0.14; odors, 0.03 ± 0.07; $t(37) = 2.56$; $d = 0.54$; $P = 0.01$], the difference between the modalities was much smaller. In addition, for odors, strong FOKs were associated with more subjective identifications than weak [$t(32) = 7.10$; $d = 1.63$; $P < 0.001$], which was also the case for the pictures [$t(30) = 16.99$; $d = 3.37$; $P < 0.001$].

To conclude, (i) the pictures of famous persons triggered more subjective identification responses than the odors, thus replicating the finding of Experiment 1; (ii) this difference was larger for the strong FOKs; and (iii) for both modalities, strong FOKs were associated with more subjective identifications than weak FOKs.

Predictive validity. The Pearson correlation between mean FOK and proportion correct naming across the individuals was analyzed. The *z*-test for the difference between two correlation coefficients showed that the predictive validity of the FOKs was significantly higher for the pictures [$r(38) = 0.60$; $P < 0.0001$] than for the odors [$r(38) = 0.10$; $P = 0.55$; $z = 2.48$; $P = 0.007$], with a non-significant correlation for the odor-elicited FOKs. Nelson (1984) argued that the best measure of FOK accuracy is the non-parametric Goodman–Kruskal gamma (G) correlation between FOK and subsequent memory performance. Also with this method the mean gamma was lower for the odors (0.48 ± 0.62; $n = 22$) than the pictures (0.93 ± 0.14; $n = 19$). Because of the generally low proportion correct naming following the FOKs, several participants did not come up with a single correct name. For these persons it was not possible to calculate a correlation as indicated by the low number of observations. Even though the number of paired participants with calculable correlations in both conditions were low, the paired *t*-test reached the level of a statistical tendency [$t(7) = 2.08$; $P = 0.08$). The aggregated gammas (i.e. across the whole data set) showed the same pattern [odors, $G(662) = 0.54$, $P < 0.001$; pictures, $G(835) = 0.89$, $P < 0.001$]. These gammas are similar, but slightly higher than what was found in Jönsson and Olsson (2003). In sum, the results on predictive validity corroborate previous findings. Although FOK judgements about odor names are predictive of later retrieval (this study) or recognition (Jönsson and Olsson, 2003), they are less predictive than FOKs for other modalities (Cain *et al.*, 1998).

General discussion

Cain *et al.* (1998, p. 321) argued that the 'ability of subjects to rate the veridicality of their answer . . . has unexplored dimensions of possible strategic importance to odor identification'. The current study utilized the insight that people's metamemory judgements can broaden our knowledge of the odor naming process. Below follows a short summary of the main results, followed by a discussion.

1. People are more likely to attempt to immediately name odors than famous persons but are on these occasions better in naming the persons.
2. Odor naming failures are typically due to failures to identify the odor rather than name activation failures. This finding is important for the understanding of why it is so difficult to name even common odors.
3. TOT experiences, that is, feelings of imminent retrieval of an appropriate name, are very uncommon for odor names, but more frequent for person names.
4. FOK judgements about odor names are less predictive than equivalent judgements about person names.
5. Neither of the two modalities showed an increase in the strength of the FOK judgements as an effect of the familiarization procedure. However, correlational analyses showed that FOK and familiarity were highly correlated for both modalities.
6. The stronger the FOK judgements, the more often the participants reported that they had identified the stimuli, giving further support for the accessibility perspective.

Immediate naming

In Experiment 1, correct immediate naming was 96% for the persons and in Experiment 2 it was 87% (Table 1). The correct immediate naming was significantly lower for the odors (68% in Experiment 1 and 59% in Experiment 2). Note that the instruction was identical for the odor and picture conditions, that is, the participants should only use the immediate

naming option if they thought they could name the stimulus in question correctly. If, as instructed, they applied the same criterion in the two conditions for responding in the immediate naming section, it would indicate overconfidence in the veracity of the emitted odor names, and a larger bias than for person names. This interpretation is supported by the high overconfidence in odor naming found in two other studies (Jönsson and Olsson, 2003; Jönsson *et al.*, 2005).

The immediate naming results are inconsistent with the idea of poor activation of the odor name for odors. On the contrary, the frequent use of the option to immediately name the odors indicates that odor names are often available (i.e. activated). The problem is instead that they are often incorrect. This is consistent with the idea put forward here that it is not odor naming that fails *per se*, but the identification of the odorous object.

Subjective identifications

For the odors, there was a notable difference in the amount of subjective identifications in the first (8%) and second (25%) experiments. This is probably due to the different experimental methods. In Experiment 1, the identity judgements were made before the metamemory judgements, whereas in Experiment 2 they were instead made after these judgements. In addition, half the stimuli of the second experiment had already been presented (twice) in the familiarization phase. This gave the participants more time to search for the identity, which may explain at least part of the difference. In spite of this, both experiments indicate that the inability to name an odor is typically due to an inability to identify the odor, that is, naming of odors typically fails already in the object identification phase. This was corroborated (Experiment 1) by the degree of semantic information available about the odors and persons. The statement that there is a poor link between an odor and its veridical label (Engen, 1987, 1991; de Wijk *et al.*, 1995; Herz and Engen, 1996) would instead predict that naming typically fails in the name activation phase. The current results speak against such a proposition. If a person encounters an ambiguous odor he or she is typically preoccupied in understanding what it is that smells, rather than knowing this and feeling that the missing odor name is more or less on the verge of being retrieved.

TOT experiences

The first experiment showed that very few odors elicited TOT experiences, especially in comparison with the pictures of famous persons. This confirms the hypothesis that TOT experiences for odors are unusual. However, an alternative explanation of the large cross-modal discrepancy can be that, rather than odor-elicited TOTs being exceptionally few, it is the TOTs for person names that are numerous (see e.g. Burke *et al.*, 2004). Most probably it is a combination of both. Brown (1991) reviewed TOT research up to 1991 and found that the typical TOT incidence across studies is $\sim$13 ± 5%. In Experiment 1, odors (5%) elicited fewer TOTs than typically found in studies of other modalities, whereas for persons (18%) the number was instead slightly higher. A poor odor–name link would instead predict a higher TOT incidence than for other modalities.

However, the fact that two studies of the TON experience have been unable to find any evidence for the availability of structural–phonological information (Jönsson and Olsson, 2003; Lawless and Engen, 1977) supports the idea of a poor odor–name association but, as noted above, such a poor association does not seem to be the main reason for the poor naming performance in odor naming tasks. Because odors elicit few TOT experiences, future studies should increase the amount of stimuli and/or participants to further elucidate the apparent lack of structural–phonological information.

Predictive validity

The second experiment corroborated and extended previous findings that metamemory accuracy is lower for odors than for other modalities (Cain *et al.*, 1998; Jönsson and Olsson, 2003). A central question is why odor-elicited FOKs are less predictive. The present study sheds some light on this issue. The participants' subjective identifications as well as their poor ability to categorize the unnamed odors (see also Jönsson and Olsson, 2003) indicated that their knowledge about the unnamed odors was low. Indeed, even when they made strong to very strong FOK judgements they were typically still trying to identify the odorous object. This means that the knowledge of the unnamed odors, compared with persons, was generally lower. This, in combination with our documented poor ability to discriminate between odors, suggests that the odor-elicited FOK judgements are based on less precise or discriminating information than FOKs elicited by other modalities, which may explain the differences in predictive validity. In other words, people's metamemory judgements are better for other modalities because they generally have more information about the unnamed stimulus and/or the missing target memory. Thus they are able to better discriminate between instances when they will and when they will not be able to retrieve or recognize the missing person name than is the case for odors.

Metamemory theory

The accessibility perspective proposes that people's metamemory judgements are based on the sheer amount of available information about the sought-for target, irrespective of the accuracy of that information. Experiment 2 further supported the perspective because the FOK judgements were higher when the stimuli were identified than when they were not (see also Jönsson and Olsson, 2003). However, the level of knowledge about the unnamed odors is generally low, and lower than for the unnamed persons. This can be seen in the odor categorization performance, the relatively low amount of subjective identifications and the apparent lack of structural–phonological information about missing

odor names. One possibility is that, if the amount of accessible information is not as available or informative for odor-elicited FOKs as it is for other modalities, other cues may have an increased role as a basis for the FOKs for odors. One such cue could be the familiarity of the odor. However, Experiment 2 failed to show any effect of familiarization on FOK. Although this could be taken as evidence against familiarity as a basis for FOKs, the fact that the two variables were highly correlated (for both modalities) speaks against such a conclusion. Also, familiarity has repeatedly been shown to be related to FOK judgements in other contexts (Metcalfe, 2000; Metcalfe *et al.*, 1993; Reder, 1987). Possibly, the two stimulus sets were too familiar from the beginning, counteracting further familiarization.

Leibert and Nelson (1998) used word pairs (cues and targets) to investigate the role of cue familiarity for FOKs. They also failed to show an effect on the FOKs when familiarizing the cue only, because mean FOK strength was identical to a condition where the cue had not been repeated. However, when both the cue and target words were repeated together they found an increase in the FOKs, but this also increased mean recall. They concluded that their results were inconsistent with the hypothesis that FOKs depend solely on cue familiarity. To conclude, more research is needed to evaluate the role of cue (i.e. odor) familiarity as an underlying basis for FOKs about odor names and a different methodology might be needed to clarify the issue. A possible better approach would be to use more unfamiliar stimuli.

Table A1 The stimuli used in Experiment 1

Odorants	Pictures
Aniseed	Britta Lejon
Apple	Calista Flockhart
paprika	Camilla Thulin
Black pepper	Carina Lidbom
Clove	Charles Aznavour
Curry	Christopher Walken
Dill	Colin Powell
Garlic	Condoleezza Rice
Geranium[a]	Demi Moore
Ginger	Efva Attling
Juniper	Elijah Wood
Ketchup	Ernst Kirchsteiger
Lemon	Ingela Thalén
'Messmör'[c]	James Gandolfini
Motor oil	Jeff Bridges
Nutmeg	Jerzy Einhorn
Onion	Johnny Rotten
Orange	Julianne Moore
Oregano	Keanu Reeves
Peanut butter	Leif Pagrotsky
Peppermint[b]	Liza Marklund
Plastic padding	Marita Ulvskog
Shoe cream	Monica Lewinsky
Soft soap	Paul Simon
Strawberry jam	Robert Mitchum
Tabasco	Sissela Kyle
Tea	Spike Lee
Vanilla	Thomas Östros
Vinegar	Ulrika Messing
Violet[a]	Viggo Mortensen

[a]An essence from Stockholm Aeter och Essencefabrik AB.
[b]An essence from Apoteksbolaget.
[c]A unique Swedish dairy product.

Acknowledgements

The authors would like to thank Henrik Olsson for helpful comments on an earlier version of the manuscript. This research was supported by the Bank of Sweden Tercentenary Foundation (1998-0270).

References

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

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

Burke, D.M., Locantore, J.K., Austin, A.A. and **Chae, B.** (2004) *Cherry pit primes Brad Pitt*. Psychol. Sci., 15, 164–170.

Burton, A.M. and **Bruce, V.** (1992) *I recognize your face but I can't remember your name: a simple explanation?* Br. J. Psychol., 83, 45–60.

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

Cain, W.S. and **Gent, J.F.** (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 Publishing Company, New York, pp. 170–186.

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.

Chobor, K.L. (1992) *A neurolinguistic perspective of the study of olfaction*. In Serby, M.J. and Chobor, K.L. (eds), Science of Olfaction. Springer Verlag, New York, pp. 355–376.

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

de Wijk, R.A. and **Cain, W.S.** (1994b) *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 Associates Inc. Mahwah, NJ, pp. 21–37.

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

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

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.B.** (1986) *Exploration of olfactory aptitude*. Bull. Psychon. Soc., 24, 203–206.

Hanley, J.R., Smith, T.S. and **Hadfield, J.** (1998) *I recognize you but I can't place you: an investigation of familiar-only experiences during tests of voice and face recognition.* Q. J. Exp. Psychol., 51A, 179–195.

Hanley, J.R. and **Turner, J.M.** (2000) *Why are familiar-only experiences more frequent for voices than faces?* Q. J. Exp. Psychol., 53A, 1105–1116.

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

Holley, A. (2002) *Cognitive aspects of olfaction in perfumer practice*. In Rouby, C., Schaal, B., Dubois, D., Gervais, R. and Holley A. (eds), Olfaction, Taste and Cognition. Cambridge University Press, pp. 16–26.

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

Jönsson, F.U. and **Olsson, M.J.** (2003) *Olfactory metacognition*. Chem. Senses, 28, 651–658.

Jönsson, F.U., Olsson, H., and **Olsson, M.J.** (2005). *Odor emotionality affects the confidence in odor naming.* Chem. Senses., 30, 29–35.

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. 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.

Koriat, A., Levy-Sadot, R., Edry, E. and **de Marcas, S.** (2003) *What do we know about what we cannot remember? Accessing the semantic attributes of words that cannot be recalled*. J. Exp. Psychol. Learn. Mem. Cogn., 29, 1095–1105.

Köster, E.P. (2002) *The specific characteristics of the sense of smell.* In Rouby, C., Schaal, B., Dubois, D., Gervais, R., and Holley, A. (eds), Olfaction, Taste, and Cognition. Cambridge University Press, New York, pp. 27–43.

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

Leibert, T.W. and **Nelson, D.L.** (1998) *The roles of cue and target familiarity in making feeling of knowing judgements.* Am. J. Psychol., 111, 63–75.

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.

Murphy, C., Cain, W.S., Gilmore, M.M. and **Skinner, B.R.** (1991) *Sensory and semantic factors in recognition memory for odors and graphic stimuli: elderly versus young persons*. Am. J. Psychol., 104, 161–191.

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

Olsson, M.J. and **Cain, W.S.** (2000) *Psychometrics of odor quality discrimination: method for threshold determination*. Chem. Senses, 25, 493–499.

Olsson, M.J. and **Fridén, M.** (2001) *Evidence of odor priming: edibility judgements are primed differently between the hemispheres*. Chem. Senses, 26, 117–123.

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

Schwartz, B.L. (2002) Tip-of-the-tongue States: Phenomenology, Mechanism, and Lexical Retrieval. Lawrence Erlbaum Associates, Mahwah, NJ.

Sulmont-Rossé, C., Issanchou, S. and **Köster, E.P.** (2005) *Odor naming methodology: correct identification with multiple-choice versus repeatable identification in a free task*. Chem. Senses, 30, 23–27.

White, T.L. and **Kurtz, D.B.** (2003) *The relationship between metacognitive awareness of olfactory ability and age in people reporting chemosensory disturbances*. Am. J. Psychol., 116, 99–110.

Wise, P.M., Olsson, M.J. and **Cain, W.S.** (2000) *Quantification of odor quality.* Chem. Senses, 25, 429–443.

Wrzesniewski, A., McCauley, C. and **Rozin, P.** (1999) *Odor and affect: individual differences in the impact of odor on liking for places, things and people*. Chem. Senses, 24, 713–721.

Yarmey, D.A. (1973) *I recognize your face but I can't remember your name: further evidence on the tip-of-the-tongue phenomenon*. Mem Cognit., 1, 287–290.

Accepted February 25, 2005