UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Social Sciences 6

Olfactory Metacognition

A Metamemory Perspective on Odor Naming

FREDRIK JÖNSSON

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2005

ISSN 1104-232X
ISBN 91-554-6269-3
urn:nbn:se:uu:diva-5821

Dissertation presented at Uppsala University to be publicly examined in Sal IX, Universitetshuset, Uppsala, Friday, June 3, 2005 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in Swedish.

Abstract
Jönsson, F. 2005. Olfactory Metacognition: A Metamemory Perspective on Odor Naming. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences* 6. 57 pp. Uppsala. ISBN 91-554-6269-3.

Although many aspects of odor naming have received attention during the years, the participants' own cognitions (metamemory) about their naming attempts have not. (i) We showed that feeling of knowing (FOK) judgments accompanying odor naming failures are predictive of later recognition (Study I) or retrieval (Study III) of the missing name, but to a lesser degree than equivalent judgments about names of persons. "Tip of the nose" (TON) experiences do predict later odor name recall (Study I), but are otherwise poorly related to any partial activation of other information associated with the odor. (ii) We evaluated two theories proposed to explain the underlying basis of FOK judgments. Correlational analysis showed that FOK judgments about odor names are related to the perceived familiarity of the cue triggering the FOK (cue familiarity theory; Study III). FOK judgments are based on the amount of available information about the sought-for memory (accessibility theory; Study I and III). (iii) We demonstrated that the participants are overconfident in their odor naming attempts (Study I and II). This may to some degree be due to the arousing properties of the odors (Study II), suggesting that emotional variables should be taken into account when researching metamemory. (iv) Our inability to correctly name odors are typically not due to an uniquely poor association between odors and their proper names, but rather due to failures to identify the odors (Study III), that is, failures to retrieve "what it is". It was also found that TOT experiences are unusual for odor names and more so than for person names. (v) We discuss potential differences between olfactory metamemory and metamemory for other modalities. The TON experience differs from the tip of the tongue (TOT) experience and the predictive validity is lower for metamemory judgments about odor names compared to other modalities.

Keywords: odor identification, odor naming, metamemory, confidence, tip of the tongue experience, tip of the nose experience, feeling of knowing

Fredrik Jönsson, Department of Psychology, Box 1225, Uppsala University, SE-75142 Uppsala, Sweden

ISSN 1652-9030
ISBN 91-554-6269-3
urn:nbn:se:uu:diva-5821 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-5821)

List of Papers

I Jönsson, F. U., & Olsson, M. J. (2003). Olfactory metacognition. *Chemical Senses*, 28, 651-658.

II Jönsson, F. U., Olsson, H., & Olsson, M. J. (2005). Odor emotionality affects the confidence in odor naming. *Chemical Senses*, 30, 29-35.

III Jönsson, F. U., Tchekhova, A., Lönner, P., & Olsson, M. J. (in press). A metamemory perspective on odor naming and identification. *Chemical Senses.*

Errata

Jönsson and Olsson (2003), p. 652, right column, 1st paragraph, line 36-38:
The sentence "An example of underconfidence would be if a participant averages 60% correct on trials where he or she is 80% confident" should be changed to "An example of underconfidence would be if a participant averages 80% correct on trials where he or she is 60% confident".

Jönsson and Olsson (2003), p. 655, Figure 1:
The word "confidence" in the description of the curves should be exchanged for "calibration".

Jönsson and Olsson (2003), p. 656, Table 2:
In the table text it is written "Mean proportion of correct responses and mean total proportion of responses on the TON questionnaire as a function of TON strength". The words "and mean total proportion of responses" should be deleted from the sentence.

Contents

Introduction

The importance of olfaction to humans is evidenced by, for example, the perfume industry with a multibillion dollar turnover worldwide. People use perfumes to smell better, turn more attractive and to hide natural body odors. Also, the "taste" experience is primarily dependent on olfaction (Ikeda, 2002). This explains why food can be experienced as flavorless when olfaction is dysfunctional (e.g., when having a cold) and why we are much poorer in identifying foods with the sense of taste alone compared to when the sense of smell is functioning (Mozell et al., 1969). Recognition and identification of relevant odors are important and a prime example is wine evaluation (Parr et al., 2002), being able to sense if food is edible or not, identifying a gas leak or not, and so forth. Also, mothers with a functional sense of smell can with a 90% success rate identify their newborns by olfactory cues after only 10-60 minutes exposure (Kaitz et al., 1987; men were not tested in that study). Identification of an odor even changes how it is perceived. Peoples' ratings of odor familiarity, intensity and pleasantness is higher when an odor is identified (either spontaneously or with the help of a provided odor name), than when it is not (Distel and Hudson, 2001). Moskowitz, (1979; in Distel & Hudson, 2001) found that liking, sweetness, and intensity changed when the brand names of five cologne fragrances were presented compared to when they were not.

Despite the importance of olfaction, we are quite poor in naming odors. The ability, or better inability, to properly name odors without the help of visual or other contextual cues has vexed researchers of olfaction for decades. Herz and Engen (1996, p. 301) described it as the "most contentious issue in human olfactory processing". If a participant is shown a common object or product like, for example, an orange all healthy participants would be able to say that "this is an orange". However, if he or she instead smelled the odor of orange (without showing the odorous object or any other cue) naming performance drops considerably, although the odor is most probably considered quite familiar by most people. The present thesis focuses on our ability to identify and name odors. More precisely, the work is focused on a specific aspect of odor identification, namely the subjective knowledge people have about the odors they encounter in everyday life and its relation to actual knowledge.

Odor naming

Correct naming performance of a set of common everyday odors rarely average above 50% (Cain, 1979; Cain et al., 1998; Desor & Beauchamp, 1974; de Wijk & Cain, 1994a; de Wijk & Cain, 1994b; de Wijk et al., 1995; Distel & Hudson, 2001; Lawless & Engen, 1977; Olsson & Fridén, 2001), and as a consequence much less for more unfamiliar and uncommon odors. The identification rate of any single item rarely reaches 100% across a group of participants. The cause for this is not fully understood, but several competing or complementary explanations have been proposed. It has been suggested to be due to (i) odor identification is not important to humans, (ii) there is a poor link between an odor and its name, (iii) odor and language information processing share some of the same cortical resources leading to interference, and (iv) there is a lack of social consensus around odor names.

Köster (2002) argued that our poor odor identification performance has to do with its minor importance for humans and that odor identification is not the primary function of the nose. He stressed that, in real life, rather than being able to identify odors and judge their intensity, it is much more important to be able to detect their presence, discriminate between them and recognize them as either familiar or unfamiliar. 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 & Kurtz, 2003).

A common explanation for why it is so difficult to name odors is that there is a *poor link* between the odor and its name. This is in line with Köster's argument that odor identification is not as important as other olfactory abilities. It has been suggested that this link is inherently weak, that is, the verbal areas of the brain are poorly associated with the olfactory processing areas and more so than for other modalities (Engen, 1991; Engen, 1987; Herz & Engen, 1996). Lorig (1999) instead argued that odor information processing and language processing *share* some of the same cortical resources. Due to this, simultaneous processing of olfactory and language information leads to interference, which in turn can explain why it is so difficult to name odors. Other researchers have argued that our inability to link an odor to its name is not inherent but rather has to do with how we learn odors. De Wijk et al. (1995) argued that learning to associate odorous objects with their proper names is not formalized in society (e.g., in school) as is naming of visual objects, which in turn would lead to less well learned odor-name associations. De Wijk et al. proposed that our inability to name odors is due to a combination of a poor odor-name association (either inherent or due to less learning) and a poor odor discrimination ability. De Wijk et al. further 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. Sulmont-Rossé et al. (2005) proposed yet another explanation. They argued that a part of the low performance found in odor naming tasks may be due to the lack of social consensus around odor names, that is, people encounter certain odors in different contexts or in different products, which means that which odor is associated with which objects differs between persons and cultures. For example, in their experiment several participants consistently labeled artificial flower odors as "cleaning supply" or "bathroom freshener". This was not the veridical name expected by the experimenters, but they may still be considered as correct descriptions of the odorants. This is because these products are often fragranced with such odorants.

Odor discrimination

Wise and colleagues (2000) pointed out that almost all tests of olfactory performance are limited to some degree by the property of discrimination (see also Cain, 1979; Schab & Cain, 1991). De Wijk et al. (1995) noted that "all subsequent higher order processing, including recognition memory and identification, can only be as accurate as the resolution of the sensory system (p. 24)". Similarly, Eskenazi et al. (1986) stated that "the ability to discriminate must underlie the ability to identify (p. 204)" an odor.

Several empirical studies have tied odor discrimination ability to the ability to identify or name an odor correctly. Olsson and Cain (2000) used eugenol (cloves) and citral (lemon) to demonstrate that the degree to which an odor needed to be substituted with another fairly dissimilar odorant 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). However, odor discrimination can improve with repeated exposure (Jehl et al., 1995; Rabin, 1988). In the first experiment of Cain et al. (1998) participants tried to name the same set of odors in four sessions (separated by about 2 days). The overall naming accuracy increased over time, but the naming performance varied across the sessions and an odor that was correctly named one day was sometimes incorrectly named another. The hit rate varied between 42% and 47%, across the sessions. However, only about 33% of the odors were correctly named in all sessions, indicating errors of discrimination.

Several studies have made correlational analyses of the relationship between odor discrimination and odor naming ability. In another of the Cain et al. (1998) experiments, they found that participants' odor discrimination ability for one odor set was highly correlated with their naming ability for another set (see also Rabin, 1988). Eskenazi et al. (1983) presented several odors to a group of healthy participants and temporal lobectomy patients. Their task was to judge which of three odors were dissimilar, and their performance on this task was significantly correlated with their odor naming ability. De Wijk and Cain (1994a) showed that odor naming performance

follows an inverted U-shaped course over age. It is best in young adults and worst in children and elderly. Odor discrimination ability showed the same pattern and was correlated with naming ability. De Wijk and Cain (1994b) found significant correlations between odor discrimination and naming performance for both free and cued naming (see also Eskenazi et al., 1986) as well as with consistency in applying the same label to a given odor with different intensity. In the above studies the correlations (r) varied between .48 and .80.

The odor naming attempts people make in themselves often reveal errors of discrimination (e.g., Engen, 1987), like when the name lemon or lime is applied to the smell of orange. Researchers of olfaction have long been aware of this and often find it meaningful to separate between two measures of naming performance, a strict measure of accuracy (the odor labels has to be fully correct) and a measure also including so called "near misses". An example of the latter would be to name lemon as orange or garlic as onion. Apart from this, a response can also be too generic, for example when the label edible or fruit is reported for lemon. Such responses may, at least in some cases, reveal some level of knowledge about the odor.

The difference between odor identification and odor naming

Most research about odor identification has consisted of odor naming tasks (de Wijk et al., 1995). The typical procedure is to let a group of participants smell different odors, one by one, without visual or other contextual cues and try to name the specific object from which the odor emanates (e.g., orange, beer or black pepper for the corresponding odorous objects). The odor naming can be done without any cues, so called free identification, or the participants can be given a list with alternatives to chose from, so called cued identification. The underlying assumption is that if a person is able to name something, then he or she normally also knows what it is, that is, have identified the item in question (see e.g., Johnson et al, 1996 for similar thoughts in picture naming). However, an object can be identified by other means than by the proper name and an object can be identified even though its name is sometimes not retrieved or even learned. An example would be if you saw a picture of a sextant, were unable to name it, but could give a proper definition like "instrument used in celestial navigation to determine latitude and longitude", or any other description revealing that the object is identified. An olfactory example could be if the odor of orange was smelled and the participant is able to imagine how the fruit looks, the tree on which it grows and/or knows where to find the fruit and so forth. So, although odor identification has been used analogously to odor naming it may sometimes be useful to separate between the two. This is what we do in study III of this thesis in an attempt to elucidate where in the naming process odor naming fails.

A note should also be made about what we here mean with the term *odor naming*. Holley (2002) argued that in an ecologically valid sense, odors are nothing but "attributes of objects and substances whose natural function is to reveal the presence of those objects and substances in the environment. As a consequence, odor naming turns out to be odor-source naming (p.19)". Holley further argued that in some cultural practices, for example perfumery, odors have become dissociated from their sources meaning that their qualities rather than their referents (the objects) are brought into focus. In this work we only utilize the term odor naming as the ability to correctly name an odorous object.

The naming process

In a review, Johnson et al., (1996) proposed that picture naming occurs in three broad stages: (1) object identification, (2) name activation and (3) response generation. They argued that first an object has to be identified as a member of a particular class of objects. Then an appropriate name (or names) stored in memory must be activated. The third stage entails the articulation of one of these names as an overt response. According to Johnson et al. these stages are usually assumed to occur (more or less) sequentially. This is because an object can be recognized without its name being activated, and a name can be activated without it being overtly responded. This is in contrast to the object identification stage which is obligatory, that is, recognition of a familiar object cannot be prevented. Their view is built around the dual coding theory (Paivio, 1991). In Study III we utilized a similar differentiation between the different stages in the naming process, but applied to odor and person naming.

The connection between odors and emotion

The olfactory sense modality is often assumed to be more closely linked to our emotions and evoke more emotional memories than other modalities and empirical evidence gives some support to this assumption. The anatomy of the olfactory system is often brought forward as evidence for this close link to emotion and especially its unique and direct projection to the amygdala (Zald & Pardo, 2000). The amygdala is only three synapses from the olfactory receptor neurons in the nose (Wilson et al., 2004). In a recent functional magnetic resonance imaging (fMRI) study, Andersson et al. (2003) found that amygdala activation is associated with the emotional intensity of odors, but not their valence (i.e., pleasantness), which was instead associated with activation in orbitofrontal cortex. Herz et al. (2004) presented participants with personally significant odors, control odors of equal intensity and visual representations of the experimental and control odors. Utilizing fMRI they

found significantly greater activation in the amygdala and hippocampal regions during recall of a memory triggered by a personally significant odor than by any of the other cues.

For six common objects, Hinton and Henley (1993) let three separate groups of participants either smell its odor, see the object or its name and then write down whatever association(s) that came to mind. The olfactory presentation led to the shortest answers and the largest affect. Herz (1998b) compared emotional ratings of odors, music, and paintings. Music and paintings were chosen as comparison stimuli because they are also often considered to be emotionally laden. When asked, the participants believed that music could affect their emotions and moods more than odors, but the autonomic nervous system (ANS) measure used (heart rate) suggested that the odors were actually more arousing than the other stimulus types. However, the subjective ratings of emotion included in the study did not evidence any significant difference between the modalities, rendering a somewhat contradictory picture. One possibility is that the ANS measures were implicit, as opposed to the explicit subjective ratings. That ANS measures can be used to measure emotion in response to odors has been shown before. Alaoui-Ismaïli and colleagues (1997a; 1997b) compared subjective ratings of emotion in response to odors with six different ANS measures. They found several intercorrelations between the physiological measures and the subjective ratings; in particular pleasantness.

The hypothesized ability of odors to cue personal memories that are emotionally colored, vivid, and old is often referred to as the Proust phenomenon after a well known anecdote in a novel by the author Marcel Proust (1919). In their review of odor-cued autobiographic memories, Chu and Downes (2000) argued that there is some evidence that olfactory stimuli can cue autobiographical memories more effectively than cues from other sense modalities. They further argued that arousing stimuli may be especially effective retrieval cues and that the olfactory sense modality is often assumed to be the most affectively activating among our sense modalities. In contrast to this a recent study of autobiographic memories by Willander and Larsson (2005) showed that visual cues evoked significantly more emotional memories than both olfactory and verbal cues. Applying a paired associate learning paradigm, Herz and colleagues (Herz, 1998a; Herz & Cupchik, 1995) let participants encode olfactory, visual, tactile, lexical, and musical stimuli (cues) together with pictures of paintings. In a later test phase each cue was presented again and the task was to remember the associated painting and rate the memory of it on several scales. All modalities triggered equally accurate memories (recall of the paintings), but odor-evoked memories were consistently rated as being more emotional. It is not clear, however, whether the emotionality was an inherent aspect of the memory or if it was just due to simultaneously smelling the odor cue.

Other studies also imply that odors can evoke emotionally colored memories (although not necessarily more than other modalities). Odors can alter subjective preference ratings of neutral pictures of peoples' faces (Todrank et al., 1995) and neutral pictures of abstract paintings (van Reekum et al., 1999). Robin and colleagues (1999) showed that participants with previous negative experiences of visiting the dentist rated eugenol (an odor often encountered at the dentist) as unpleasant and showed ANS activation associated with negative emotion, whereas those that did not have such negative experiences instead rated it as pleasant and did not show such ANS activation. To conclude, research suggests that odors do evoke emotional memories. Moreover, anatomical, neuropsychological, and empirical evidence suggests that the olfactory sense modality is more closely connected to emotions than other modalities.

The basic metacognitive concepts

A recent definition of metacognition is that "*Metacognition* /…/ encompasses the knowledge and cognitive processes whose object is cognition and whose task is to control and verify one's own cognitive functioning" (Houdé, 2004, p. 227). A shorter definition is simply *cognition about cognition*. Conceptually, metacognition is closely related to consciousness (Roberts & Erdos, 1993). Although the term consciousness has lately increased in popularity among psychologists, it is closely related to centuries of philosophical thought, making a concrete definition of it difficult to make. The term metacognition might therefore be preferable. Schwartz (2002) has expressed similar thoughts, but also advanced the term phenomenology, that refers to a person's subjective experience. The basic idea is that people *monitor* (the metacognitive aspect) whatever information/cognitions that are available to consciousness and based on that information they *control* their actions. An example is if you had to solve a mathematical problem and based on your monitoring of your ability to do so within a reasonable timeframe you either go about trying to solve it or perhaps asks someone for help. In his article "Consciousness and metacognition", Nelson (1996) presents a more in depths introduction to, and discussion of, the psychological terms metacognition, the closely related term metamemory, and more importantly, their relation to philosophical thought.

The term *metamemory*, the one most often referred to here, is included in the broader term metacognition, so its definition is more specific. Metamemory can be defined as the cognitions we have about memory (rather than any cognition in general).

Metamemory judgments

Several different kinds of metamemory judgments have been investigated over the years. Some judgments are predictions of the learning of new information, so called *ease of learning* judgments. For example, when a student is about to learn a new material for an upcoming exam, different learning strategies may be applied based on an assessment of how difficult or easy a to-be-learned material is or the relative difficulty of different parts of the material. Then, while learning the material people make *judgments of learning* (JOL). For example, a student studying a new language might judge that some word pairs in a Swedish-English glossary list (e.g., "psykolog"-psychologist or "sjuksköterska"-nurse) have already been well learned, whereas others have not. It is central that such judgments accurately reflect actual level of learning so that attention can be directed towards the word-pairs that are in need of additional learning on the expense of those already well taught (Son & Metcalfe, 2000). Accurate monitoring is the basis for proper allocation of study time and termination of study. For example, early studies showed that peoples' JOLs are not very accurate in predicting later recall (Vesonder & Voss, 1985), but later it was found that the timing of the JOL is important. If the judgment is made immediately after learning its predictive validity is low, but if the JOL is delayed until shortly *after* the learning the judgment is instead a highly accurate predictor of later retrieval (Nelson & Dunlosky, 1991; Dunlosky & Nelson, 1992, 1994; Kelemen & Weaver, 1997; Thiede & Dunlosky, 1994; Weaver & Kelemen, 1997).

The current thesis focuses on three types of metamemory judgments about odor names, namely (i) *feeling of knowing* judgments (i.e., predictions of the future retrieval or recognition of a currently unrecalled memory), *tip of the tongue* experiences (strong and imminent feelings that a currently unrecalled word is soon to be retrieved), and (iii) *retrospective confidence* in retrieved answers (i.e., how sure a person is that some retrieved information from memory is correct). It should be noted that in the thesis the focus is only on the *monitoring* aspect and not the control aspect of metamemory judgments. The judgments central to this work, will now be introduced in more detail[1].

Feeling of knowing

One interesting and intriguing metacognitive aspect is that we can with some accuracy monitor the availability of a memory, even though we might fail to retrieve it (Hart, 1965; Koriat, 1993, 1994, 1995; Koriat & Levy-Sadot, 2001; Metcalfe, 2000; Nelson, 1996; Nelson & Narens, 1990), referred to as feeling of knowing (FOK) judgments. The FOK is a subjective feeling that a

[1] In addition to the just mention judgments, the thesis focuses on an olfactory special case of TOT experiences, called *tip of the nose* (TON) experiences. This judgment will be introduced and further discussed in the section dealing with research on olfactory metacognition.

currently unretrieved memory is stored in memory. If we are given, for example, general information questions we can above chance tell whether an answer to a question is in memory *before* we are able to articulate it, so called rapid FOK judgments or pre-retrieval FOKs (Nhouyvanisvong & Reder, 1998; Reder, 1987). Although most FOK studies have focused on FOKs for unretrieved memories only, Reder and Ritter (1992) argued that the FOK is a general process that operates whenever memory is queried. In this thesis we only focus on the former type, namely, FOK judgments that follow retrieval failures.

A FOK can vary from no FOK to a very strong FOK. A special case is the strong and imminent feeling of retrieval people can have when they have a missing word on the tip of the tongue (see the next section). The most common way of investigating FOK judgments is to present participants general information questions (e.g., what is the capital of Australia?). For all unretrieved questions the participants judge how strong their FOK is on a scale and are then given a multiple alternative forced choice recognition test (e.g., four alternatives: Sidney, Melbourne, Canberra, Perth)[2]. The FOK ratings are then correlated with actual recognition performance and there is an extensive body of FOK data showing that people are, in general, moderately accurate in monitoring their knowledge (see Metcalfe, 2000, for a review).

As is the case with memory research in general (Annett, 1996) most metamemory research as well as metamemory theories are based on verbal stimuli. Much less attention has been directed towards FOK judgments for non-verbal stimuli (but see, e.g., Perfect & Hollins, 1999; Peynircioglu et al., 1998), and only one previous study focused on FOKs in the context of olfaction (Cain et al., 1998).

Tip of the tongue

Although tip of the tongue (TOT) experiences have vexed researchers since the days of William James (1890), the first systematic investigation of them was not made until the late 1960's (Brown & McNeill, 1966). Brown and McNeill described the TOT experience as follows: "If you are unable to think of the word but feel sure that you know it and that it is on the verge of coming back to you then you are in a TOT state" (p. 327). Most subsequent studies have relied on this description or some version of it. Brown (1991) pointed out that a definition of TOTs almost always include (i) that a missing word is known but at the moment unavailable and (ii) recall of that word feels imminent. A TOT is about a word and a failure to retrieve that word from memory. TOT experiences are often associated with partial knowledge of structural-phonological aspects of the unretrieved word (e.g., first letter), but people can also access semantic attributes of the words (Koriat, Levy-Sadot, Edry, & Marcas, 2003). Typical ways of eliciting TOTs in the labora-

[2] The correct answer is Canberra.

tory is to present participants with word definitions or show pictures of famous persons.

Brown also differentiates between immediate and delayed resolution (i.e., eventual correct recall of the missing word). TOT experiences are resolved within a minute or two in about 50% of the cases (Brown, 1991), but others are resolved up to several days after their occurrence, long after the actual TOT experience has dissipated (i.e., the missing word just pops up in consciousness). The present thesis only focuses on immediate TOTs. TOT experiences seem to be universal across both languages as well as cultures (Schwartz, 1999; Schwartz, 2002) and can therefore easily be explained to the layperson or research participant. The scientific term is, at least in Swedish and English, identical to the term used in everyday language.

TOTs are closely related to FOK experiences and most probably, any FOK study includes some TOTs, disguised in the shape of strong FOKs. However, all strong FOKs are not TOT experiences as indicated by results of Yaniv and Meyer (1987). They used word definitions to elicit FOK responses for the defined words. Following all FOKs the participants also indicated whether they had a TOT experience or not. Although the mean gamma correlation (G = .92) between the FOK and TOT ratings was indeed very high, they noted that all strong FOKs did not lead to TOT responses. Nhouyvanisvong and Reder (1998) further noted that whereas a person in a FOK experiment is merely asked about the likelihood of being able to recognize (or retrieve) the missing answer at a later time, people who report TOT experiences are always very confident that they will eventually retrieve the answer and the feeling in itself motivates continued memory search.

Retrospective confidence judgments

Retrospective confidence refers to the belief a participant has in the veracity of his or her memory. This type of judgment does not, like FOKs and TOTs, refer to later retrieval or recognition of some unretrieved memory, but is instead a judgment made about the correctness of already retrieved information.

Most memories are associated with a more or less conscious interpretative aspect evaluating the correctness of them. Imagine that you have a disagreement with someone about some important fact or occurrence (e.g., discussing with your partner when you met the first time), and although you are both quite confident in that you remember the date correctly, at least one of you is obviously incorrect. Another example is if you forgot your PIN code to the ATM machine and you can only try three times or your original PIN code will be blocked. You really need to withdraw some money and although you are sure about the first two numbers, how about the two last ones; are they 43 or 34? Because you really need those money right now and is confident that 43 is more likely than 34 you give it a go. Good luck! Everyday life is abundant with situations when peoples' confidence in their

memories plays a role and sometimes a bad relationship between confidence and memory can be detrimental. One recent "real life" story is about the man McKinley Cromedy in USA, who was identified by a woman as her rapist. Based on her confident identification of him, and despite a lack of physical evidence, he was sentenced to 60 (!) years in prison. However, soon after the trial DNA evidence proved him innocent (Hoffman, 2005). Scientific studies typically find a low relationship between witness confidence and identification performance (e.g., Bothwell et al., 1987; but see Lindsay et al., 1998). Many studies also find that people are overconfident in their judgments about the accuracy of their memories (e.g., Griffin & Tversky, 1992; Koriat et al., 1980; Sniezek et al., 1990) and it has been debated whether the overconfidence found reflects a "real" bias in peoples' estimates or if people are actually rather well calibrated and the overconfidence found are due to, for example, statistical artefacts or item sampling errors (Juslin et al., 2000). This will be further discussed in the general discussion in relation to our studies.

The difference between absolute and relative metamemory accuracy

Several different measures of metamemory accuracy have been used in the literature. They can be put into two distinct categories referred to as *relative* and *absolute* metamemory accuracy. Relative accuracy refers to how well a person judges the likelihood of correct recall or recognition for one item relative to another item. For example, is a high FOK judgment (e.g., 70%) associated with a greater recognition performance on the following memory test, than an item given a lower FOK (e.g., 50%)? To the extent that there is a positive relationship between FOK strength and memory performance is measured with a correlation. Nelson (1984) compared several correlational measures and argued that the best available measure of FOK accuracy is the non-parametric Goodman-Kruskal gamma correlation.

Absolute metamemory accuracy is about the exact deviation between a person's judgments of the likelihood of correct recall for a particular set of items and the actual percentage correct recall or recognition for these items. For example, does a person who for a set of items is 60% confident in his or her answers also have about 60% correct on the memory test? This is measured with an over-/underconfidence (O/U) score. A person that is more confident than correct (e.g., 80% confident in his or her answers, but actually only has 60% correct) is said to be overconfident. A person that is exactly as confident in his or her answers as actual memory performance (e.g., 60% confident and 60% correct recall) is said to be perfectly calibrated, whereas a person that is generally less confident than correct is labeled underconfident. The focus thus lies on the exact match between subjective belief in memory

performance and actual performance. This technique to measure the accuracy of the metamemory judgments have been widely applied in other fields (for reviews see Juslin et al., 2000; Lichtenstein et al., 1982; McClelland & Bolger, 1994; Yates, 1990). Examples of research include decision making (e.g., Yates, 1990), confidence in one's general knowledge (e.g., Juslin et al., 2000), confidence in perceptual and sensory discrimination (e.g., Juslin & Olsson, 1997), and witness psychology (e.g., Juslin et al., 1996).

The relative and absolute metamemory accuracy measures can be used independently depending on what the researcher is interested in, but do complement each other if used in combination. For example, there might be a perfect correlation between a person's confidence in his or her odor naming attempts (i.e., the correlation is 1 or -1), but at the same time the person may be highly overconfident.

Metamemory theories

When investigating metamemory it is important to find out how accurate different metamemory judgments are in different situations and populations, as well as what impact these judgments have on behavior. However, a fundamental question for researchers is not only to understand when and how accurate the judgments are, but also why they are as accurate or as inaccurate. It is therefore necessary to develop theories and hypotheses about the underlying bases of these judgments. Although theoretical developments in the field of metamemory are relatively recent, substantial advancements have been made. A very brief overview of the main theories will be presented here.

The theories proposed have often been divided into two main classes, although possibly not mutually exclusive: namely, *direct-access* and *inferential* perspectives (Nelson et al., 1984; Schwartz, 2002). In addition, they have mostly been developed and applied in the context of FOK judgments (Koriat, 1993; Schwartz & Metcalfe, 1992) and TOT experiences (Schwartz, 2002; Schwartz et al., 1997; including the olfactory analogue to TOTs, so called tip of the nose experiences; see study I of this dissertation), but not with regards to retrospective confidence judgments. The theoretical focus on the current thesis is therefore not so much on retrospective confidence, but rather FOK and TOT judgments.

The direct-access theories assume a close connection between the metamemory judgment and the retrieval process. Essentially this means that the participants can somehow "sense" or monitor the objective strength of the unrecalled memory and this is the basis of the FOK or TOT judgment. The inferential theories do not claim this direct link between the objective memory status and the subjective experience. The inferential views claim that metamemory judgments are based on the clues associated with the target memory (e.g., partial knowledge about the unretrieved word), those associ-

ated with the cue triggering the FOK or TOT (e.g., the familiarity of the cue) or a combination of both (Koriat & Levy-Sadot, 2001). Because the judgments are inferences, rather than direct links to objective memory strength, the subjective experience is dissociable from the objective memory performance. This means that some variables may affect the subjective experience (e.g., increase or decrease the FOKs) without affecting the memory and vice versa. A direct-access theory would instead predict that because the subjective experience is so closely linked to the actual objective memory strength, any manipulation of the memory, should also correspondingly affect the judgments about that memory (Schwartz et al., 1997). The latter perspective also implicates the assumption that the metamemory judgments should be relatively accurate, again because they are assumed to be based on information about the target (e.g., a missing word) in memory. The inferential perspectives instead assume that because memory and metamemory (i.e., the subjective experience) is dissociable, the metamemory judgments may sometimes be systematically incorrect. Extensive evidence today exists of that metamemory and memory are dissociable (e.g., Metcalfe et al., 1993; Nelson, 1996; Metcalfe, 2000; Reder & Ritter, 2002). The two main inferential theories that have been proposed are the *cue familiarity theory* (e.g., Koriat & Levy-Sadot, 2001; Metcalfe et al., 1993; Reder, 1987) and the *accessibility theory* (e.g., Koriat, 1993; Koriat, 1995; Koriat & Levy-Sadot, 2001). One central difference between the two accounts is that cue familiarity refers to the familiarity of the *cue*, whereas the accessibility theory refers to the amount of accessible information about the *target*.

The cue familiarity theory assumes that people base their metamemory judgments on the perceived familiarity of the cue, for example the familiarity of the odor used to cue its proper name. For pre-retrieval FOKs (Reder, 1987, 1988, Reder & Ritter, 1992), Koriat and Levy-Sadot (2001) defined cue familiarity as perceptual fluency, that is, the ease with which information is perceived. For FOKs that occur *after* unsuccessful retrieval attempts familiarity has not been as clearly defined in the literature. Metcalfe et al. (1993, p.851) merely defined it as "familiarity with or knowledge about the cue". A prediction that springs from the cue familiarity theory is that manipulation of the cue (e.g., familiarization via cue priming), but not the target memory, will lead to changes in the FOK.

The accessibility perspective instead assumes that FOK is based on the amount and intensity of the partial information that is retrieved while searching for the missing memory (Koriat & levy-Sadot, 2001). This partial information can according to Koriat and Levy-Sadot (2001; see also Koriat 1993; 1995) be fragments of the target (e.g., the first letter if it is a word or the first two numbers of a four-digit PIN code), semantic and/or episodic attributes related to the target (e.g., a person remembers having studied the unrecalled information) or any other information related to the unretrieved memory. One important aspect of the accessibility perspective is that the judgments

are based on the total amount of information, independent of its correctness. In line with this it has been shown that FOK judgments are significantly higher after commission errors (incorrect answers) than after omission errors (i.e., no answers; Koriat, 1993; Krinsky & Nelson, 1985; Nelson & Narens, 1990). Recently, Koriat and Levy-Sadot (2001) presented an integration of the cue familiarity and the accessibility perspectives. They demonstrated that a certain degree of familiarity is necessary to activate search and retrieval of accessible information, that is, familiarity and accessibility interact. Study I and III of the current thesis include an investigation of the latter two theories, but only Study III investigates the cue familiarity theory.

Olfactory metacognition

Although metacognitive and metamemory research (e.g., Brown & McNeill, 1966; Hart, 1965) goes back to the earliest years of cognitive psychology, such research pertaining to the olfactory modality has been very scarce. We know very little about people's awareness of their olfactory ability, how tuned they are to actual objective ability, and what impact more or less functional awareness has on their behavior and well-being in different situations.

Previous research has highlighted several examples of how important it is that the beliefs (i.e., metacognitions) people have about their olfactory ability is on par with actual ability, that is, that the olfactory metacognitive ability is fully functional (White & Kurtz, 2003; Nordin et al., 1995). White and Kurtz (2003) argued that accurate estimations of olfactory ability is essential when determining how much to rely on olfactory information. Nordin et al. (1995) argued for the importance of the olfactory system as pleasure-seeking system and as a warning system. People that are not aware of that their olfactory ability is dysfunctional may not buy smoke alarms, gas detectors or adequately attend to the edibility of spoiled food via other channels than the sense of smell (e.g., to ask someone else to smell it).

Nordin et al. (1995) compared the self-reported and actual smell sensitivity (as assessed by a test) of three groups of elderly: normal elderly, patients with Alzheimer's disease (AD) and patients with sinusitis (an inflammation of the nasal sinuses). Among the elderly *with smell loss* as much as 74% of the AD patients and 77% of the normal elderly reported normal smell sensitivity. In contrast, only 8% of the sinusitis patients with loss reported normal smell sensitivity. Regarding the two first groups the authors argued that unawareness of actual smell dysfunction could be because the loss in smell sensitivity degrades gradually over a long time, which make it less noticeable (in contrast to an immediate and sizeable smell loss due to, for example, an accident). Another yet uninvestigated possibility is that the elderly simply compared themselves with other elderly and for that reason overly often erroneously reported having a normal sense of smell.

The potential danger of a discrepancy between judged and actual ability is illustrated by results of Stevens et al. (1987). They reported that 7 out of 21 healthy elderly failed to detect a warning agent in gas (ethyl mercaptan) at a concentration where the accompanying propane would actually explode. It is well documented that olfactory ability (e.g., odor identification) decrease with age (Larsson et al., 2004). Apart from age, cued odor identification was in the Larsson et al. study found to be related to sex (females were generally better), education, cognitive speed, and vocabulary. Murphy et al. (1991) demonstrated that elderly are also less sensitive to odors than young, with about 10 times higher odor thresholds.

White and Kurtz (2003) investigated young and elderly patients with different smell complaints (dysmosmics). Their odor identification ability was tested with the Odorant Confusion Matrix (OCM) and compared to self-reported assessments of ability as measured by a four alternative category scale. After dividing the OCM results in four categories to match the category scale (e.g., a score of 80-90% was considered to reflect a normal sense of smell) they compared the OCM results to the self-assessments. They found that 41% of the old adults and 42% of the young adults were unaware of the degree of the deficit, as measured by either over- or underestimations of actual performance at the OCM. Although this corroborates the finding of Nordin et al. (1995) in that estimates of olfactory ability in some populations is dysfunctional, a potential problem with the White and Kurtz study is that any change in the criteria set for the different categories of odor identification ability on the OCM, would also change the degree of over- and underestimations. Nevertheless, the evidence up to date indicates that olfactory metacognition is in some populations impaired and, arguably, this could have an impact on their lives.

Feeling of knowing

Cain and colleagues (1998, Exp. 3) were the first to investigate FOK judgments about odor names. They let young adults try to (i) name a set of common odors and (ii) answer general information questions. The participants made FOK judgments on the first seven odors not named and the first seven questions not answered. Each such trial was followed by an eight-alternative forced choice recognition test. Somewhat surprisingly they found no significant predictive validity for FOK judgments about odor names. They did, however, find a typical correlation between FOKs and recognition performance for the general information questions. In her review of metamemory Metcalfe (2000) showed that the typical correlation lies between .45 and .55.

The tip of the nose phenomenon

Lawless and Engen (1977; Experiment 3) performed a study modeled after a classical TOT study (Brown & McNeill, 1966), but instead of using word definitions they used odors to trigger TOT experiences for the veridical names of the odors, or more precisely, the names of the odorous objects. In the experiment participants made familiarity ratings and odor identification attempts for a set of 48 odors. If the participants were unable to name an odor, but made a high familiarity rating and felt that they had the odor name on the tip of the tongue, they filled out a questionnaire asking for partial information (also called structural-phonological clues; Koriat et al., 2003) about the sought-for odor name (e.g., first letter, number of syllables) and other associated information (e.g., if they could form a visual image of the odorous object or give a plausible odor category, for example, "fruit" for orange). They found that the participants had virtually no partial information about the name. This led them to give the experience a separate name, namely the "tip of the nose" experience as an olfactory parallel to the TOT experience.

The lack of structural-phonological information about the missing word, typical of the TOT experience, has been taken as evidence for the above suggested poor link between the odor and its name (Engen, 1991; Engen, 1987; Herz & Engen, 1996).

Retrospective confidence judgments

Whereas many studies have investigated odor naming per se, the subjective knowledge people have about the veracity of their odor identifications have not received much attention in the literature. If you give people a list with the names of common odors and ask them to what extent they believe that they would be able to identify these odors by smell alone, they would be quite confident in their ability (Cain, 1982), but actual performance will be quite poor (Cain, 1979; Cain, 1982; Cain et al., 1998; Desor & Beauchamp, 1974; de Wijk, 1994; de Wijk & Cain, 1994; de Wijk et al., 1995; Distel & Hudson, 2001; Lawless & Engen, 1977; Olsson & Fridén, 2001). Anyone having performed laboratory tests of odor naming knows that participants are often surprised by the difficulty of identifying even very familiar odors.

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 relationship between confidence and odor naming performance, as measured by a correlation between confidence and the correctness of the answers (relative

accuracy) and the study does not give any information about whether the participants were over- or underconfident (absolute accuracy). Taking into consideration the difficulty of naming even common odors, an interesting venue is to find out exactly to what extent people know whether they have identified an odor correctly or not. In the current context, participants tell how confident they are in the correctness of an odor name just retrieved. This is investigated in study I and II of the current thesis. Study II also focused on a possible interaction between confidence and emotionality.

The aims of the thesis

One of the most studied aspects of olfactory cognition has been odor identification, typically measured with odor naming tasks (Herz & Engen, 1996). Therefore, it is surprising that the metacognitive aspect has received so little attention. To get a complete picture of odor identification, it is central to highlight all its aspects. This includes the relationship between odor intensity, detection, discrimination and identification, but also includes the metacognitive aspects. When people encounter and identify odors this involves some degree of reflective processing, judgments and subjective analysis about what it is that smells. This is particularly evident in the laboratory where participants often are deprived of the contextual cues they might have in everyday life. Participants are vexed and sometimes frustrated over why they can not identify or name an odor they clearly recognize as familiar. Almost no research has focused on olfactory metacognition or metamemory. The overall aim of the thesis was to investigate aspects of metamemory in odor identification. In other words, focus is on the subjective knowledge people have about the odors they encounter in everyday life, both in the sense of how accurately we monitor our odor naming ability and the underlying basis of this monitoring. This overall aim can be subdivided into several aims.

Predictive validity

One central question is to what degree metamemory judgments about odor names reflect actual naming performance. When people fail to name an odor, to what extent is the feeling of knowing the odor related to actual knowledge about the odor (FOK judgments and TON experiences)? This is investigated in Study I and III.

Metamemory theory

Another motive for investigating olfactory metamemory is theoretical. Several theories have been developed to explain the underlying basis of

metamemory judgments. As with memory research in general (Annett, 1996), metamemory research and theories have to a large extent been based on verbal and visual material (but to some extent also, for example, auditory material). Whereas the strength of a theory lies on the one hand in its ability to explain a particular phenomenon yet another of its strengths is its ability to generalize across situations and modalities. As a consequence, one of the goals of the thesis was to evaluate current metamemory theories' potential as explanations of the underlying basis of metamemory judgments in an olfactory context. The theories under scrutiny here are the accessibility theory (Koriat, 1993, 1995; Koriat & Levy-Sadot) and the cue familiarity theory (Reder, 1987; Schwartz & Metcalfe, 1992; Koriat & Levy-Sadot, 2001). Study I and III relate to these theories.

Retrospective confidence and emotionality

Little is known about the accuracy of peoples' confidence in their odor naming attempts. When people try to identify odors via their names, to what extent are they aware of when they do so correctly and when they do not (retrospective confidence judgments)? This is investigated in Study I and II. Another aim was to investigate whether the emotional attributes of an odor can affect the confidence in its name (Study II). In a more general sense this study investigated if it is at all possible that metamemory may interact with emotionality or if metamemory judgments should be viewed as "cold", non-emotional judgments. The idea that emotionality can interact with confidence is not new (e.g., Brigham et al., 1983; Herz, 1998a; Herz, 2000; Hosch & Bothwell, 1990; Talarico & Rubin, 2003), but research is scarce. In a related vein of research, Herz (2000; 1998a) showed that odors do not evoke more accurate memories than other modalities and suggested that people might be overconfident in odor-evoked memories due to the close connection between emotion and the olfactory modality.

The odor naming process

In study III a metacognitive perspective is utilized to further our understanding of why it is so difficult to name even common odors. We focused on the naming process and argued that the relatively poor odor discrimination found in studies of olfaction (Olsson & Cain, 2000; Wise et al., 2000) and its strong relationship with odor naming performance (Cain et al., 1998; de Wijk and Cain, 1994a; Eskenazi et al., 1983; Eskenazi et al., 1986; Rabin, 1988) predicts that most odor naming failures occur in the object identification stage, rather than the name activation stage. The popular idea of a poor link between an odor and its name, either due to weak learning or if it is inherent, would instead predict that naming failures predominantly are due to weak activation of the veridical name. Odors were compared with pictures of

famous persons and person naming was to a larger degree assumed to fail in the name activation stage than odor naming.

Is odor metamemory special?

Many olfactory researchers have focused on potential differences between cognition associated with the olfactory sense modality and that of other modalities (e.g., Herz & Engen, 1996). Apart from a general focus at metamemory judgments about odor names we were particularly interested in whether those judgments are special in some respect compared to findings from other modalities. For example, the Cain et al. (1998) study indicated that FOKs about odor names have a lower predictive validity than FOKs for answers to general information questions, and TON experiences may be phenomenologically different from TOT experiences (Lawless & Engen, 1977). Study I and Study III of the thesis focused on these issues. However, of these two, only Study III entailed a direct cross-modal comparison. More specifically, special focus is directed towards a comparison of the predictive validity of FOK judgments and TON experiences, the naming process, the ability of odors to elicit TOT experiences and metamemory theory.

Empirical studies

To enhance the readability most statistics have been deleted from the summary of the empirical studies. Detailed statistics are presented in the enclosed articles at the end of this book. Also a word should be said about the presentation of the odorants. In all experiments the odors were common everyday items, that is, products that could be bought in any store (e.g., coffee, garlic, orange). The only exceptions were violet, geranium and peppermint, which were artificially produced. For example, to produce the smell of orange, oranges were bought, cut into small pieces and put in 160-ml tinted glass jars with screw lids. Cotton pads were placed over the odorous objects in such a way that no visual or other cues apart from the odorants were available to the participants. It should also be noted that all the participants had a functional sense of smell.

Study I: Jönsson and Olsson (2003) – Exp. 1

Aim

Experiment 1 aimed to investigate the predictive validity of FOK judgments about odor names. Previous research had shown that such FOKs were non-significantly related to subsequent memory performance in a multiple forced choice recognition task (Cain et al., 1998).

Procedure

Thirty-nine young men and women were presented with 48 common odorants. The task was to smell each odor once and try to name it. If a name was not retrieved a FOK judgment was made about the certainty of subsequent recognition of the veridical odor name (on a percentage scale from 0% to 100%, with 20% intervals). Directly after each FOK judgment an odor name was shown, either the correct or the incorrect, with 50% of the trials being associated with a correct name and vice versa. The presentation order of the stimuli and whether a correct or incorrect name was shown was fully randomized.

Main results

57% of the naming attempts made were correct (i.e., not including the FOK trials) confirming previous findings of the difficulty of naming even familiar odors. The participants FOK judgments were moderately, but significantly, correlated with their recognition of the correct odor name (G = .34), as measured by the Goodman-Kruskal gamma correlation (Nelson, 1984).

Conclusion

For other modalities the gamma correlation typically ranges between .45 and .55 (see Metcalfe, 2000, for a review). In a comparison with FOKs for answers to general information questions and odor naming Cain et al. (1998) found a typical FOK accuracy for the latter, but a non-significant predictive validity for the former. The current experiment utilized another method and demonstrated that odor-elicited FOKs are indeed indicative of later recognition performance, but compared to what is typically found for other modalities possibly less so.

Study I: Jönsson and Olsson (2003) – Exp. 2

Aim

The aims of the second experiment were (1) to replicate and further investigate the findings of Lawless and Engen (1977) concerning the TON phenomenon, and (2) to investigate peoples' confidence in the veracity of their odor naming attempts (i.e., how accurately they monitor their odor naming performance). The main points of interest were:

(i) How predictive is the TON experience about later retrieval of the missing odor name?

(ii) Is the subjective strength of the TON experience related to subsequent retrieval of the name (with stronger TON experiences being more predictive, i.e., leading to more retrievals, than weaker)?

(iii) What information is related to the predictive validity of the TON experiences?

It should be noted, that although FOK judgments were also included in this experiment they were mainly included to give the participants the option to sort away the unnamed odors that did not elicit TON experiences. They will therefore not be reported here.

Procedure

We presented 70 common odorants to 40 young men and women, all with a functional sense of smell. The task was to name each odor with its veridical label. The main procedure is outlined in Figure 1.

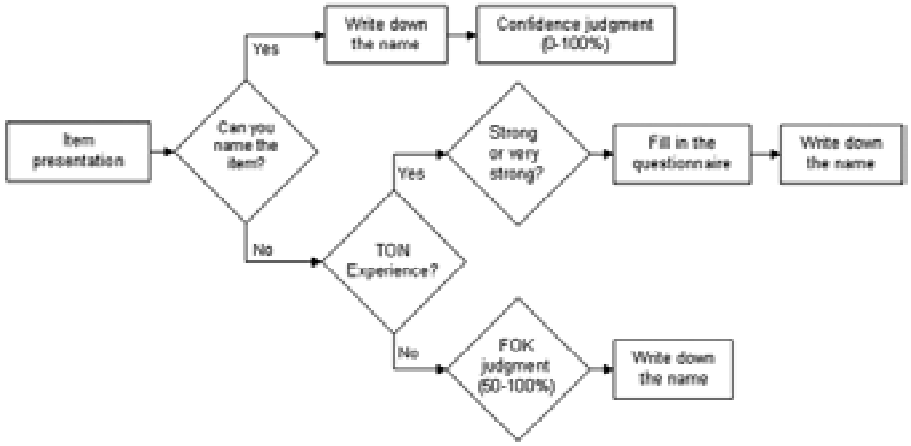

Figure 1. The procedure of Experiment 2 in Study I is presented. If the participants named an odor, it was always followed by a confidence judgment (0-100%, with 20% intervals) about its correctness. Following retrieval failures, one of two outcomes was possible: Either they had a TON experience (i.e., a strong feeling that they would soon be able to retrieve the name of the odor) or if not, they made a FOK judgment as in Experiment 1. Directly after having made a TON experience they judged its strength (strong or very strong) and while searching for the name of the odor, they filled in a questionnaire similar to that in Lawless and Engen (1977; see Jönsson & Olsson, 2003 for a more detailed description of the questionnaire). The questionnaire consisted of two parts, one part asking for partial information about the sought-for odor name (e.g., first letter) and the other part asking for other associated information (e.g., odor category like "fruit" for orange). Following TON experiences, the participants were given a maximum of 90 seconds to fill in the questionnaire and retrieve the name. If the name was retrieved before the allotted time had ended it was written down, all further activity terminated and the next trial commenced.

Main results

Retrospective confidence judgments

The naming performance was similar to that in Experiment 1 (see Table 1, where results from study II is also shown). The relative metamemory accuracy, measured by the Goodman-Kruskal gamma correlation was high, but the absolute accuracy measure showed that the participants had an overly high belief in the veracity of their odor naming attempts (Table 1). The high

overconfidence is also illustrated by Figure 2, where proportion correct odor naming is plotted as a function of the confidence categories (group data).

Table 1. M*ean naming performance, confidence, overconfidence and correlation between confidence and naming performance. All means are across participants with standard deviations in parentheses.*

	Study II	Study I, Exp. 2
	Mean (SD)	Mean (SD)
Naming performance	.49 (.18)	.54 (.16)
Confidence	.63 (.14)	.80 (.10)
O/U-index	.14* (.17)	.26* (.16)
Gamma correlation	.66* (.30)	.64* (.30)

*Indicates that the value is significantly different from zero as measured by a one-sample t-test (all *p's* < .001; only calculated for the gamma correlation and the O/U-index).

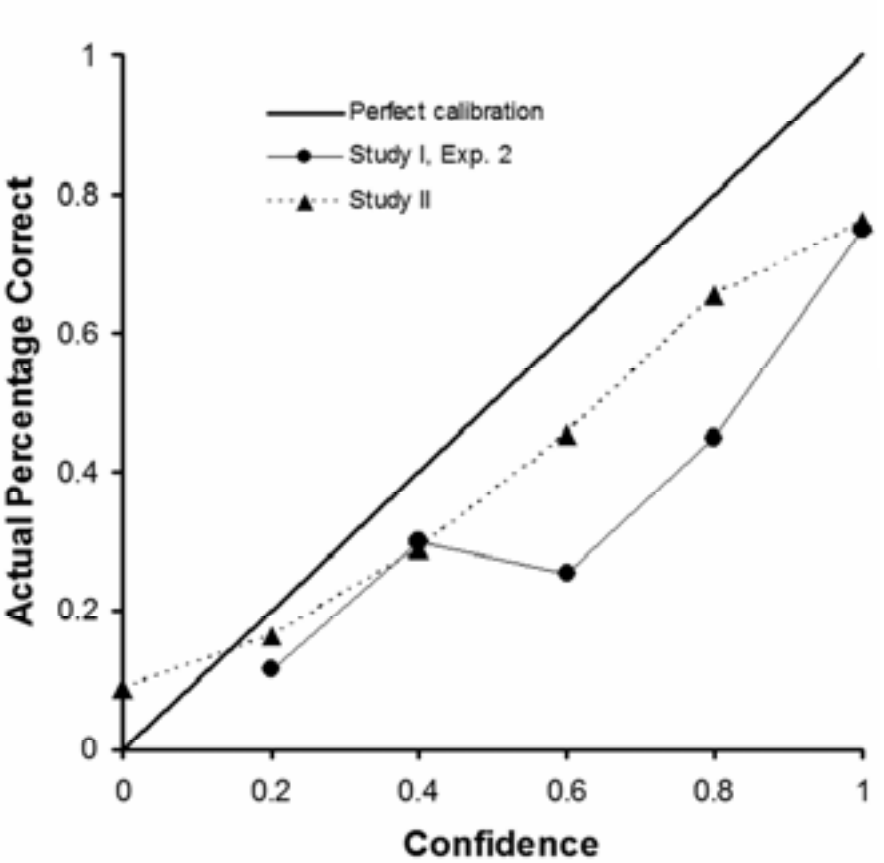

Figure 2. Confidence judgments plotted against actual proportion of correct answers per confidence category for Study I (Exp. 2) and Study II. The diagonal line represents perfect calibration.

Tip of the nose phenomenon

The TON experiences were followed by retrieval of a veridical odor label in 23.8% of the cases. It is noteworthy that as much as 34.9% of the TON experiences were "resolved" incorrectly. Very strong TONs were significantly more often resolved (36.7%) than the strong (19.0%), indicating that the participants did base their judgments on some valid cues. Because the participants had no partial information about the missing odor name the TON experience can not be based on differential access to the sought-for name (i.e., the stronger the experience the more access). However, they could significantly more often categorize the odors when the experience was very strong (22%) compared to when it was strong (12%). This means that the stronger the experience, the more knowledge they had about the unnamed odors.

Conclusion

Although the participants' confidence in their naming attempts was relatively accurate as measured by the gamma correlation, the absolute accuracy measure showed that they had an overly high belief in the veracity of their naming performance. The TON results replicated those of Lawless and Engen (1977). It is unclear why TON experiences are not associated with partial access to any aspects of the odor name (as for TOTs; Brown, 1991). It is often assumed to be due to a poorly learned or inherent (inborn) poor odor-name association (e.g., Engen, 1987; Herz & Engen, 1996). This will be further discussed in a later section.

Both the resolution rate as well as odor categorization performance were related to the subjective strength of the TON experiences. This renders support for the accessibility perspective of Koriat (1993) that people's metamemory judgments are based on whatever relevant partial or associated information we have about the sought-for memory.

Study II: Jönsson, Olsson, and Olsson (2005)

Aim

Study I demonstrated that the participants were overconfident in the veracity of their odor identifications. This means that their confidence is, expressed as subjective probabilities, on average, higher than actual proportion correct odor naming. Several reasons may lay behind this overconfidence (see Jönsson & Olsson, 2003). The present experiment tested whether the emotionality of the odor could be one such factor.

Procedure

42 young adults of both sexes were presented 16 common odorants. They first smelled each odor once and rated it for valence and arousal on the Self-Assessment Manikin (SAM; Bradley & Lang, 1994), which is a non-verbal method of assessing emotion. Then they smelled the odor once more, tried to name it and following all naming attempts they also judged how confident they were in the veracity of their naming attempts. The trials that did not lead to odor naming attempts were not included in the analyses.

Main results

As can be seen in Table 1, this experiment closely replicated the results of Study I, with the only difference that the degree of overconfidence was lower. As can be seen in Table 1, the naming performance was similar in the two studies, but the degree of confidence differed very much. The overconfidence was, also in this study, significantly different from zero.

The relationship between confidence and arousal

First, mean O/U-index was calculated across participants as a function of valence category (positive, neutral and negative) and arousal (high and low). Valence and overconfidence were not significantly related, but high arousing odors led to significantly higher overconfidence (mean *O/U* = .18), than low arousing odors (mean *O/U* = .10). Second, the O/U-index was calculated as a function of each level of arousal across the whole group of participants, showing that the relationship was high, $r\,(9) = .92$; $p < .001$.

We also made some further analyses not included in the Jönsson et al. (2005) paper. Because only 16 odors were included in our study a possibility is that the results appeared due to a sampling error in the choice of items (i.e., odors). When correlating mean arousal per odor with O/U-index per odor (i.e., the analysis was made across odors rather than across participants) we found a zero correlation ($r\,(16) = .03$), which would indicate that the results may very well have been due to a sampling error. However, such an analysis does not take into account the individual differences in how the odors were perceived. A better analysis, handling these individual differences, would be to rank the 16 odors with respect to how arousing they were for each individual. Then, the across participants ranking is correlated with the O/U-index for each rank. Such an analysis showed that the correlation between ranked arousal and O/U-index was significant with a one-tailed test, $r\,(16) = .45$; $p = .04$. This is in line with the hypothesis and the other analyses. Because different ways of analyzing the data give somewhat different results it must be underscored that the arousal-overconfidence finding should be replicated.

A relevant question is whether arousal and confidence interacted independently of any changes in the odor naming performance. The data set was divided into two parts (correctly and incorrectly named odors). Two paired t-tests showed that there was no difference in confidence between high (M = .80) and low (M = .81) arousing odors when the participants named them correctly. However, following incorrect naming attempts the participants were significantly more confident in the high (M = .53) than in low arousing odors (M = .46), t (40) = 2.07; p = .04. Remember that these levels of confidence were observed despite the fact that the correctness was zero.

Conclusion

We replicated both the gamma correlation and the overconfidence of Study I (Exp. 2). It was also shown that the overconfidence increases as a function of arousal, but not valence. It was demonstrated that when the participants in fact knew the correct answer, their confidence was unaffected by how arousing the odor was. However, when they named the odors incorrectly, their confidence judgments were higher for high-arousal odors. The results are important, because they indicate that there is reason to consider emotional variables when researching metamemory, but they also indicate that this relationship only appear when the participants are less sure in their answers (because mean confidence was lower for incorrectly compared to correctly named odors).

Study III: Jönsson, Tchekhova, Lönner, and Olsson (2005) – Exp. 1

Aim

The primary aim of the first experiment was to delineate where in the naming process odor naming fails. It was hypothesized that odor naming primarily fails due to failures to identify[3] the odorous object (i.e., odor naming fails already in the object identification stage in the naming process), and more so than for names of famous persons. This hypothesis is in contrast to the often encountered assumption that odor naming typically fails due to a poor activation of the odor name (i.e., in the name activation stage and not earlier in the naming process).

[3] The difference between odor *identification* and odor *naming* was dealt with in the introduction. Remember that odor naming performance is the most common measure of odor identification performance and the two terms are often used synonymously. However, *for this particular study*, we found it useful to differentiate between the two. Here odor naming refers to the ability to name an already identified odor and odor identification refers to the ability to identify an odor by any means (e.g., a description).

A second aim was to investigate the ability of common odors to elicit TOT experiences (i.e., strong and imminent feelings of retrieval of a missing word). If odor naming typically fails in the object identification stage, then TOT experiences should be unusual (because, presumably, they are name activation failures as indicated by the lack of partial access to the sought-for word). Third, and as indicated above, a cross-modal comparison was made, with names of famous persons as reference modality. TOT experiences were in this study defined as strong and imminent retrieval of an identified odor or person.

Procedure

Forty young adults (20 women) with a functional sense of smell viewed 30 pictures of famous persons and smelled 30 common odorants. The pictures and odors were equally difficult to name. The main procedure is outlined in Figure 2.

Figure 3. The main procedure of Experiment 1 of Study III is depicted. The overall task was to name the odors and persons. The persons were named by their surname and were presented on a computer screen and the odors were presented in the same way as in all other experiments. For each trial the participants had a questionnaire in front of them that consisted of three parts (A, B and C in the figure). Only one of these sections was filled out per trial. If the participants thought they immediately could name the stimulus they did so in section A and the experiment continued with the next trial. If not, they instead filled in section B (if they had identified the stimulus, but could not name it) or section C (if they could neither name the stimulus nor identify it). More detailed instructions for when a stimulus was identified or not is presented in Jönsson et al. (in press). For section B, they first judged whether they had the missing name on the tip of the tongue, whether that experience was strong or very strong, or if they did not have a TOT experience. Then they filled in a plausible odor category (e.g., fruit), or if a picture was presented, they filled in the person's profession (e.g., actress). If they were not able to do this they left this part blank. Finally they wrote down the name of the stimulus (if retrieved). In section C they first made a FOK judgment on a three-point scale (1 = no or weak; 2 = strong; 3 = very strong) about how sure they were that they would be able to *identify* the stimulus, either by a description of it or by retrieving the odor's or person's proper name.

Then they wrote down an odor category or a profession as in section B, after which they were asked to describe the odor (as an alternative way of demonstrating that the odor or person was identified). Finally they wrote down the name if available. In both section B and C, the participants always wrote down the missing name as soon and if it was retrieved and discontinued filling in the questionnaire. From the initial onset of each stimulus a maximum of 90 s was allotted, after which the trial was terminated. Each stimulus could be sampled repeatedly during this time. The presentation order of the stimuli was fully randomized for each participant, with the only exception that every second trial was an odor and every other a picture trial.

Main results

Naming performance

Two measures of naming performance were calculated. To test whether odors and persons were equally difficult to match an *overall* measure of their naming performance was calculated (i.e., correct naming after section A, B and C in the questionnaire). The two stimulus sets were equally difficult to name (p = .71; Table 2). Second, the immediate naming performance was calculated, excluding the B and C sections of the questionnaire. The persons were significantly better named than the odors (Table 2). We also analyzed to what extent the participants used the immediate naming option, referred to as *immediate naming attempts (Table 2)*. This analysis showed that the participants more often thought they could immediately name an odor than a person, but this difference only reached the level of a statistical tendency.

Table 2. (1) Proportion correct overall naming, (2) proportion correct immediate naming, (3) proportion immediate naming attempts and (4) proportion subjective identifications are shown for Experiment 1 and 2. (5) Proportion correct odor categorizations and person professions is shown for Experiment 1.

	Experiment 1		Experiment 2	
	Odor	Picture	Odor	Picture
1. Correct overall naming	.30 (.12)	.29 (.17)	.29 (.12)	.27 (.20)
2. Correct immediate naming	.68 (.16)***	.96 (.08)	.59 (.20)***	.87 (.18)
3. Immediate naming attempts	.28 (.14)*	.22 (.16)	.44 (.16)***	.28 (.19)
4. Subjective identifications	.08 (.09)***	.43 (.18)	.25 (.20)***	.46 (.18)
5. Odor category/person profession	.29 (.15)**	.40 (.18)	--	--

Note. Correct overall naming and immediate naming *attempts* (line 1 and 3) are calculated as proportions of all trials (i.e., including the FOK and TOT trials). Correct immediate naming at line 2 only refers to the proportion of the immediate naming attempts that were correctly named. Subjective identifications at line 4 and odor category/person profession at line 5 refer to the proportion of the unnamed stimuli that were reported to be identified. The latter proportion thus excludes the immediate naming attempts.
* $p = .06$; ** p < .001*** $p < .0001$

Subjective identifications

They should fill in section B (the TOT section) of the questionnaire if they had identified the stimulus and section C (FOK judgments) if they had not. The proportion subjective identifications (the number of section B responses divided by the total number of responses, that is, 30) was very low for the odors and significantly lower than for the persons (Table 2). This is in line with the hypothesis that odor naming failures are typically due to failures to identify the odors.

Other semantic information

When the participants failed to name the stimuli they significantly more often knew the persons' profession than they could categorize the odors (Table 2). This corroborates the subjective identification measure, because odors, which were to a lesser extent subjectively identified, were also associated with a lower degree of knowledge, indicated by the low odor categorization performance.

TOT experiences

It was hypothesized that the odors would elicit very few TOT experiences and significantly less than the persons. This hypothesis was dependent on the outcome of the identification judgments. It was hypothesized that if odor naming failures mostly are due to odor identification failures, rather than failures to activate the name of a known item, TOTs should as a consequence be very infrequent. It was also hypothesized that odor-elicited TOTs would be more infrequent than TOTs for person names. These hypotheses were all confirmed. In a review of the TOT phenomenon, Brown (1991) noted that the typical TOT incidence across studies vary between 13±5%. The amount of odor-elicited TOTs was well below that range (5%), whereas the pictures were in the upper part of the range (18%). Note that the latter percentages, as opposed to the proportions in Table 3, were calculated on all trials, because this is the standard measure used in TOT research. How the different judgments were distributed is shown in Table 3.

Table 3. *The table illustrates how the metamemory judgments were distributed. This is shown as mean proportions (SD) across participants and is 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 judgments).*

Modality	Metamemory judgments		
	non-TOT	*strong TOT*	*very strong TOT*
Odor	.02 (.04)	.03 (.04)	.04 (.05)
Picture	.19 (.14)	.09 (.10)	.15 (.13)
	weak FOK	*strong FOK*	*very strong FOK*
Odor	.37 (.16)	.30 (.12)	.24 (.14)
Picture	.51 (.17)	.05 (.06)	.01 (.02)

Conclusion

Whereas unnamed odors elicit few TOT experiences they do seem particularly apt in eliciting strong feelings of knowing (see Table 3). These strong FOKs for odors are typically associated with identification failures, rather than a failure to activate the name. Based on the present data, conclusions can be made about the TON experience. Based on the distribution of the metamemory judgments in Table 2, it can be concluded that the TON experience is actually a mix of mainly strong FOK judgments about the retrieval of odor identity and a few TOT experiences. Compared to similar judgments for person names the participants are further back in the naming process. Summing up the current results and those of Jönsson and Olsson (2003), it seems that although accessible information does seem to play some role as an underlying basis of prognostic metamemory judgments about odor names, such information may be less valid or exact. This is because, although FOKs are strong, the (unnamed) odors are only rarely identified, which was not the case for the famous persons.

Study III: Jönsson, Tchekhova, Lönner, and Olsson (2005) – Exp. 2

Aim

The primary aim was to perform a cross-modal comparison of the predictive validity of FOK judgments about odor and person names. Only one such cross-modal study exists in the literature (Cain et al., 1998), but then general information questions were used as comparison. 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. A third aim targeted metamemory theory, more precisely (i) the cue familiarity theory and (ii) the accessibility theory. Familiarity has in studies of other modalities been found to be related to FOK judgments (e.g., 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 about the unnamed odors, the familiarity might play an important and possibly increased role as underlying basis for FOK judgments about odor names. We here used a familiarization procedure to evaluate the role of cue familiarity and in comparison to the reference modality. It was expected that the familiarization procedure should increase the strength of the FOKs as compared to the non-familiarized stimuli. Also, the accessibility perspective was further evaluated. It was hypothe-

sized that the more subjective identifications reported by the participants, the stronger the FOK judgments.

Procedure

The procedure was similar to that of Experiment 1, but (i) only FOK judgments about the missing names were gathered, (ii) the time for each trial was decreased to 60 seconds and (iii) a new and improved questionnaire was used. Instead of judging whether they had identified the stimuli before the metamemory judgments, this was now done after those judgments.

Thirty-eight young men and women were, as in Experiment 1, individually exposed to 30 common odors and 30 pictures of famous persons. The experiment consisted of three separate phases: (1) familiarity judgments of the stimuli (familiarization), (2) short break with filler activities, and (3) test phase. In the first phase the participants were presented with half the stimulus set (uniquely randomized for each participant from the set of 30). The task was to smell each odor or watch each picture during three seconds, immediately followed by a familiarity rating of the stimulus (Scale: 1 = not at all familiar; 10 = extremely familiar). The presentation order was 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. 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 procedure of the test phase is described in Figure 4.

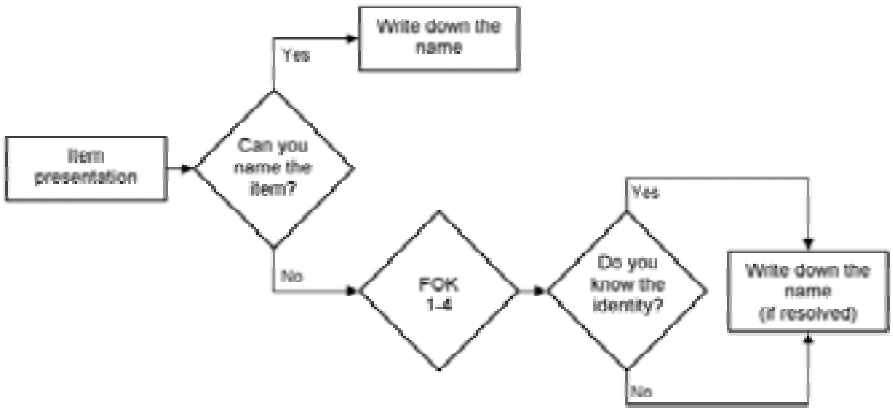

Figure 4. The procedure of the test phase of Experiment 2 of Study III is shown. All pictures and odors were presented one at the time, 50% of them from each modality being repeated from the first phase (old) and fifteen being new. The task was to name the stimuli. If they could name the odor or person within 15 seconds (so called immediate naming) they wrote it down and the next trial commenced. If the partici-

pants could not name the stimulus within the 15 seconds, they instead made a feeling of knowing (FOK) judgment about how sure they were of subsequent retrieval of the correct name (within about 60 seconds, with the time spent for the immediate naming attempt subtracted). The FOK judgment was made on a scale from (1) no/weak, (2) average strength, (3) strong, to (4) absolutely sure. Immediately after the FOK judgment they answered whether they had or had not identified the unnamed stimulus. This judgment was identical to that of Experiment 1. Whenever they retrieved the missing name they wrote it down, interrupted all further activities and continued with the next trial. The presentation order for the stimuli was fully 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).

Main results

Familiarity judgments in the first phase

In the first phase, the participants judged the familiarity of half the stimulus material twice. The analyses showed that odors and pictures were equally familiar and that the second presentation was associated with significantly higher familiarity judgments than the first.

Naming performance

First *correct overall naming* was analyzed to test whether the odors and pictures were equally difficult to name. A Modality (odor/picture) x Type (old/new; where old represents the stimuli also presented in the first phase) repeated measures ANOVA with proportion correct naming performance as dependent variable showed that there was no main effect of Modality. This means that the odor and picture sets were well matched for difficulty of naming (See Table 2). A main effect of Type showed that old stimuli were better named than new. Post-hoc comparisons of a significant Modality x Type interaction showed that whereas odor naming was unaffected by the repetition procedure, the old pictures (M = .32) were significantly better named than the new (M = .22). To conclude, the familiarization procedure resulted in better correct overall naming of the persons, but not the odors.

Second, we made two analyses of immediate naming (i.e. excluding the FOK trials), namely Modality x Type repeated measures analyses of variance of (i) the proportion *immediate naming attempts* (i.e. how often the respective stimulus type triggered an immediate naming response), and (ii) the proportion correct such naming attempts. The odors led to significantly more immediate naming attempts than the persons. A main effect of Type showed that familiarized stimuli (M = .39) more often led to immediate naming attempts than those not familiarized (M = .33). Post-hoc comparisons of a significant Modality x Type interaction showed that whereas the odors were unaffected by the repetition procedure (M_{new} = .43; M_{old} = .45), newly

presented pictures (*M* = .24) were associated with significantly less immediate naming attempts than old (*M* = .33). The analysis of proportion *correct immediate naming* showed that, despite the fact that the odors led to more immediate naming attempts, the participants were significantly poorer in naming the odors than the pictures (see Table 1). There were no other main or interaction effects.

Cue familiarity

It was expected that the repetition procedure would enhance the familiarity of the odors and pictures. Because FOK judgments are supposed to be based on the familiarity of the stimuli those should also be enhanced. However, the repetition procedure had no significant effect on the FOKs for neither modality. The only effect of the repetition procedure was that the participants named the old pictures better than the new (which was not the case for the odors). It is unclear why the familiarization procedure failed to enhance FOKs.

In a second analysis the familiarity judgments in the first session were correlated with the FOKs in the test phase. The relationship (mean gamma correlations across participants) was high for both odors (*M* = .51) and pictures (*M* = .74) indicating that familiarity is indeed related to FOKs, but significantly higher for the pictures..

It is unclear why the correlation was higher for the persons. We therefore investigated the reliability of the familiarity ratings for the two modalities. The participants made two consecutive familiarity ratings (on half the stimuli) in the first phase, which made it possible to calculate the test-retest reliability of those ratings. This analysis showed that the odor familiarity ratings were significantly less reliable (mean r_{xx} = .66) than the person familiarity ratings (mean r_{xx} =.79). 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 and accessibility

Directly after each FOK judgment the participants answered whether they had or had not identified the stimulus in question. A Modality x FOK repeated measures ANOVA, with proportion subjective identifications (i.e., "yes"-responses) as dependent variable showed the following:

The pictures were significantly more often identified than the odors (Table 2). A main effect of FOK strength showed that the higher the FOK, the higher the proportion of subjective identifications. This is perfectly in agreement with the accessibility theory (Koriat, 1993, 1995). Post-hoc comparisons of a significant Modality x FOK interaction showed that for the stronger FOK judgments, pictures (*M* = .82) were significantly more often than odors (*M* = .47) associated with identification responses. This was also

true for the weaker FOK judgments ($M_{picture}$ = .09; M_{odor} = .03), but the difference between the modalities was much smaller.

In sum, (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

Several correlational analyses were utilized to test how predictive the FOKs were of later correct name retrieval. These analyses demonstrated that the FOKs about person names were significantly more predictive of later retrieval than the FOKs about odor names. This replicates and extends the finding of Cain at. (1998).

Conclusion

FOKs about odor names are predictive of later memory performance, but to a significantly lesser extent than for other modalities (see also Cain et al., 1998). The finding extended the generalizability of a lower predictive validity to a new cross-modal comparison and another task (free retrieval of the odor name instead of recognition). Experiment 2 replicated the first experiment and it was shown that failures to name an odor is mostly due to failures in the object identification stage. This may, presumably, be due to identification failures caused by that the odor is unknown, but also due to failures to discriminate between odors. An example of the latter would be smelling orange, followed by an uncertainty about whether it is orange, lemon, lime or grapefruit. The results indicate that when participants make predictions of recalling odor names, including strong such predictions, they are typically further back in the naming process than is the case for person names. Because odors elicited few TOT experiences(i.e., name activation failures), this indicates that the often cited TON experience (Lawless & Engen, 1977), is mostly about object identification failures rather than failures to activate the name.

Concerning the two metamemory theories tested, the experiment was non-conclusive with respect to the cue-familiarity theory. However, the accessibility perspective was further supported, because the more subjective identifications reported by the participants, the stronger the FOK judgments.

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 thesis utilized that insight, particularly in the sense that people's metamemory judgments can broaden our understanding of the odor naming process or why it is so difficult to name even common odors.

Predictive validity

Feeling of knowing judgments

It is quite fascinating that even though we sometimes can not retrieve a memory, we do with some accuracy have the ability to monitor whether we will soon be able to do so or not (Metcalfe, 2000; Nelson, 1996).

One central question is why we, in Study I, found a significant FOK-recognition correlation for odor-elicited FOKs, whereas Cain et al. (1998) did not. There are a number of methodological differences between the two studies, which makes it rather speculative to pinpoint exactly what led to the difference. That methodological differences may indeed be the reason for the differences in predictive validity between our studies is indicated by the fact that mean gamma was not only low for the odors (G = .07; *n.s.*) but also for the general information questions (G = .19; $p < .01$). Gammas for general information questions are typically higher (see Metcalfe, 2000, for a review). These gammas were calculated on the rank ordering of seven unanswered items. However, Cain et al. also calculated an aggregated gamma across all items (rather than across participants) which equaled G = .60, indicating a predictive validity of FOKs for odors as well (but still lower than for the questions: G = .89). In addition to the rank-order procedure, they also let the participants rate their FOKs on an 11-graded category scale, but no *mean* FOK-recognition gammas were reported for this measure. This would have been the most precise comparison between our two studies. However, the most interesting aspect is the cross-modal comparison in the Cain et al. study, which is perfectly in line with our data.

Three articles, including two cross-modal comparisons have now been published about how predictive FOK judgments about odor names are (Cain

et al., 1998; Study I; Study III). Based on the data available it can be concluded that such FOKs are, in a statistically reliable sense, predictive of later retrieval (Study III) and recognition (Study I) of the missing name, but to a lesser extent than for other modalities. The latter has now been shown to be true compared to FOKs for answers to general information questions (Cain et al., 1998) and FOKs for names of famous persons. Hence, it does not seem to be a qualitative, although possibly a quantitative, difference between olfaction and other senses with respect to the predictive validity of FOK judgments.

The tip of the nose experience

The fate of TON experiences and definitions of it has been one of frequent citation, but no empirical attention. Up to now, the scientific literature has only consisted of a single article by Lawless and Engen (1977; but see also Engen, 1987). Concerning the predictive validity of such experiences, the resolution rate (i.e., the percentage of resolved experiences) seems to vary between 24% (Study I) and 37% (Lawless & Engen, 1977). This is on average lower than the typical finding for other modalities. Brown (1991) made an extensive review of existing research about TOT experiences and concluded that in laboratory studies the probability that a target will be retrieved in a minute or two varies around 50%. The predictive validity is typically calculated as a percentage resolved TOT experiences, but the fact that we in Study I differentiated between strong and very strong TON experiences let us calculate a gamma correlation as well. A further analysis of the latter data showed that the Goodman Kruskal gamma correlation was moderate but significantly different from zero [$G(31) = .38$; $SD = .76$; $t(30) = 2.79$; $p = .01$], similar to the FOK correlation in Study I.

To conclude, given the research so far, the consistent picture is one of a lower predictive validity for prognostic metamemory judgments about odor name retrieval. Possible reasons for this will now be discussed in the following section dealing with metamemory theory.

Metamemory theory

The metamemory theories tested here were developed in the context of FOK judgments (Koriat, 1993, 1995; Metcalfe et al., 1993 Reder, 1987; Schwartz & Metcalfe, 1992), and were later extended to include TOT (Schwartz et al., 1997) and TON experiences (Study I). The accessibility perspective basically postulates that metamemory judgments are based on any relevant information available about the sought for memory. A central feature of that theory is that judgments are assumed to be based on the sheer amount of information, independent of its correctness. The accessibility to pertinent

information has, for olfaction, now been shown to be related to the strength of both FOK judgments (subjective identifications in Study III) and TON experiences (odor categorization performance in Study I).

However, although accessibility do seem to be important it is possible that it plays a less important role for metamemory judgments about odor names than in other types of tasks or for other modalities (e.g., FOKs for answers to general information questions). The clues underlying the metamemory judgments can be more or less discriminating. The optimal clues would allow participants to perfectly discriminate between FOKs that are weak and non-predictive of later memory performance and those that or strong and predictive (as well as a finer continuum in between the extremes). If the knowledge about unnamed odors is generally vague it could be more difficult to discriminate between more and less predictive FOKs. Essentially, the accessible information would range from no information to more but vague information. That the knowledge of unnamed odors is relatively modest is illustrated by the, at best, moderate ability to categorize them (Study I; Study III), by the low degree of subjective identifications in Study III and by the absence of structural-phonological information about the non-retrieved odor name (Lawless & Engen, 1977; Study I). The subjective identifications in Study III showed that most strong predictions of later retrieval, as well as a majority of those of lesser strength, were made for odors of which the source object had not yet been identified. Henceforth, the proposed vagueness of the accessible information could be a possible explanation for the lower predictive validity of FOKs about odor names.

In Jönsson and Olsson (2003) it was reasoned that if the overall knowledge about the odors is low, other clues than the accessibility could have an increased role as a basis for FOKs. One possibility is that FOKs about odor names are to a higher degree based on the familiarity, due to a lower availability of accessible information. However, the test of the cue familiarity hypothesis (Study III, Experiment 2) failed to show any support for an increased role of familiarity in metamemory for odor names, compared to the persons. Actually, the familiarity-FOK correlation was significantly higher for person than odor names, in opposition to the above reasoning. In Experiment 2 of Study III, it was shown that the odor and person stimulus sets were equally familiar, so this can not be the reason for the difference. However, it was also shown that the participants' odor familiarity ratings were significantly less reliable (i.e., the measurement error was higher for the odors) than the person ratings, which could at least in part, explain the lower familiarity-FOK correlation for the odors.

Koriat and Levy-Sadot (2001) suggested three alternative explanations of how cue familiarity and accessibility theory could contribute to FOKs. First, they could represent competing, alternative accounts. That is, cue familiarity is the underlying basis of FOKs and *not* accessibility or vice versa. This is how the two perspectives have been treated in most articles. Second, the two

accounts could be highly related and constitute different facets of the same mechanism. The authors argued that either it could be the ease with which information comes to mind (retrieval fluency) that underlies familiarity or a familiar cue could also be a cue that triggers many associations and vice versa. Essentially, this type of familiarity could be hard to distinguish from accessibility. Third, Koriat and Levy-Sadot argued that cue familiarity and accessibility could be two separate mechanisms, both uniquely contributing to the FOK. The familiarity is here based on the perceptual fluency rather than retrieval fluency. Perceptual fluency refers to the ease with which information is perceived. For example, Reder and Ritter (1992) showed that participants could within 850 ms predict whether they would either immediately after the response retrieve an answer to an arithmetic problem or whether they would need more time to calculate it with the help of pen and paper. Reder and Ritter also showed that these pre-retrieval FOKs were more related to the familiarity of the cue shown (i.e., the arithmetic problem) than with knowledge of the answer. Koriat and Levy-Sadot (2001) favored this third account and distinguished between pre-retrieval FOKs based on cue familiarity and post-retrieval FOKs based on accessibility. They hypothesized that the cue familiarity based pre-retrieval FOKs guide search and retrieval of accessible information, whereas post-retrieval FOKs (i.e., after a retrieval attempt has been made) are also based on the amount of accessible information available about the target memory. Their data rendered support for such a notion. Although it seems reasonable to distinguish between cue familiarity (i.e., perceptual fluency) and accessibility as two separate mechanisms when discussing the basis of pre-retrieval as opposed to post-retrieval FOKs, this distinction may be more diffuse when trying to compare the two accounts as an underlying basis for post-retrieval FOKs only. The point is that for post-retrieval FOKs the perceived familiarity of the cue may very well be highly correlated with the total amount of (accessible) information a cue triggers.

Leibert and Nelson (1998) used word pairs (cues and targets) to investigate the cue familiarity hypothesis. Similar to our results they failed to show any effect of repetition on the participants' FOK judgments when only the cue was familiarized. Mean FOK strength was identical to a condition where the cue was not repeated. However, it should be noted that Metcalfe et al. (1993), who used a similar methodology to Leibert and Nelson, did find such an effect of cue repetition. In the Leibert and Nelson study, only familiarization of both cues and targets at the same time increased the FOK ratings, but it also increased recall. They concluded that their results were inconsistent with the hypothesis that FOKs solely depend on cue familiarity, rather accessibility and familiarity seems to interact. To conclude, more research is needed to evaluate the role of cue (i.e., odor) familiarity for FOKs about odor names and a different methodology might be needed to further clarify the issue.

Retrospective confidence and emotionality

The evidence at hand indicates that people are overconfident in the veracity of their odor naming attempts. These results are consistent with the more informal notions of overconfidence that has been present in the olfactory research literature up to now (e.g., de Wijk et al., 1995; Cain, 1982). We also demonstrated (Study II) that a part of this overconfidence may be explained by how arousing the odors are (or possibly some yet unknown correlate to arousal), but this effect was mainly for the unnamed or incorrectly named odors.

Herz and colleagues (Herz, 1998a; Herz & Cupchik, 1995) showed that odors do not evoke more accurate memories than other modalities, but they tend to be more emotional. Herz (1998, 2000) proposed a close interaction between the confidence people have in the correctness of odor-evoked memories and their emotionality in stating that "odors are no better than other sensory cues at eliciting an accurate recollection. Rather, it seems, the emotional intensity of odor-evoked memories leads to the false impression that such memories are especially accurate. In other words, it is emotional intensity, not accuracy that accounts for the impression that odors are the best memory cues (Herz, 2000, p. 37)". The idea is interesting and the findings of Study II support such a notion. Future research should focus on (i) whether people are overconfident in the accuracy of odor-evoked memories (and more importantly, more so than for memories evoked by other types of stimuli) and (ii) whether the emotionality does interact with such overconfidence.

The idea that metamemory judgments may be influenced by emotion is not new (e.g., Brigham et al., 1983; Hosch & Bothwell, 1990; Talarico & Rubin, 2003), but research is very scarce. Possibly, the most promising field of investigation for such findings is in the context of highly emotional memories, like for example flashbulb[4] and eyewitness memories. Several recent comparisons of everyday and flashbulb memories (Schmolck et al., 2000; Talarico & Rubin, 2003; Weaver, 1993) have shown that the decline in the consistency (i.e. accuracy) of those memories are equal in size, but with striking differences in the subjective judgments of those memories. Ratings of vividness, recollection, and belief in the correctness of the memories remained high for flashbulb memories, but decreased over time for everyday memories. Talarico and Rubin (2003) concluded that flashbulb memories are not special in their accuracy, only in their perceived accuracy, and Schmolk et al. (2000) argued that "the fact that individuals were frequently as confident in their inaccurate recollections as they were of their accurate recollections, and failed to say that they did not remember, suggests that some of the findings reflect a difficulty in

[4] Flashbulb memories refer to memories of surprising and shocking events and are supposed to be recalled over long periods of time (e.g., Talarico & Rubin, 2003).

metamemory (p. 44)". Both these statements mirror what Herz (2000) postulated about odor-evoked memories, but the former authors did not specifically pinpoint emotionality as the possible reason for the observed overconfidence. However, Talarico and Rubin (2003) did find a negative correlation between confidence in the accuracy of flashbulb memories and subjective arousal ratings (see also Hosch & Bothwell, 1990, for a similar finding in the context of confidence in eyewitness memory). Although the latter findings may seem contrary to what we found in Study II, they do support the notion of a confidence-emotionality interaction.

To conclude, the findings of Study II demonstrated that emotional variables should be taken into account also when investigating metamemory. Apart from the importance of replicating this finding, future research should investigate if it is unique to confidence in odor naming or if it, as previously proposed (Herz, 2000) can be extended to confidence in odor-evoked memories, as well as to other modalities and contexts. For example, in a study of the effects of arousal on facial recognition Brigham et al. (1983) did not get any support for their hypothesis that the confidence-accuracy relationship should get weaker with higher arousal.

The finding of a general overconfidence in many types of tasks and what causes it has been something of an enigma in the research literature and many different explanations have been proposed (e.g., Griffin & Tversky, 1992; Juslin et al., 2000; Koriat et al., 1980; Sniezek et al., 1990). For example, Koriat et al. (1980) argued that people are inclined to search for confirmatory evidence and neglect conflicting or contradictory evidence. This confirmation bias then creates overconfidence. Zakay and Tuvia (1998) and others (e.g., Kelly and Lindsay, 1993) have suggested that confidence is based on the ease of retrieval of a memory (i.e., retrieval fluency), which need not be related to its correctness. Zakay and Tuvia showed that the faster the confidence judgments were made, the higher the confidence. However, response latency was not significantly correlated with accuracy. They argued that retrieval fluency might therefore sometimes be a misleading cue. Apart from the arousal finding in Study II, the cues underlying confidence in odor naming is at present unknown and a full explanation of the overconfidence in odor naming is yet to be found.

Cain and Potts (1992) proposed that participants sometimes misapprehend the source of the odor (e.g., grapefruit is erroneously perceived as lemon), but that their ability to realize this is compromised. If participants repeatedly are highly confident in incorrectly identified odors due to misapprehension of them, this would inflate the mean confidence without an accompanying increase in proportion correct identification. This *confusion hypothesis*, as it was termed in Jönsson and Olsson (2003), predicts overconfidence in odor naming. Olfactory researchers are well aware of the ambiguity in people's odor identification responses and often calculate two measures of odor identification (read "naming") performance; one strict measure including only

fully correct responses and one including so called "near misses". That is, responses that are almost correct are also scored as correct (e.g., citrus fruit or lemon is scored as correct for the odor of orange). At least some of these near misses are most probably errors of discrimination as proposed by Cain and Potts (1992; see also Cain et al., 1998). If the overconfidence in odor naming would only be due to such errors, the overconfidence should, as a consequence, be eliminated if near misses would be included as correct responses. In Experiment 2 of Study I odor naming performance was calculated using both a strict criterion and a more relaxed one including such near misses. A further analysis of these data showed that the O/U-index did decrease from .26 (*SD* = .16) to .17 (*SD* = .15), but it was still significantly different from zero as shown by a one-sample t-test, t (39) = .17; $p < .001$ (see also Figure 1 in Jönsson & Olsson, 2003). This means that a considerable portion of the overconfidence can not be explained by the confusion hypothesis.

Juslin et al. (2000) reviewed studies targeting confidence in answers to general information questions and made a thorough investigation of how methodological aspects can affect or even artificially create overconfidence in answers to general information questions. Based on a metaanalysis of the existing empirical studies of relevance they concluded that there is little support for the wide-spread belief in a general overconfidence in answers to general information questions. However, their article not only targeted another modality, but also another type of confidence judgment, namely confidence in the choice of the correct alternative (i.e., answer to a question) in a two-forced choice paradigm. It is therefore unclear to what extent their conclusions are applicable to the confidence measure and task used in our studies (confidence in free recall of odor names). Juslin et al. (2000) showed that overconfidence can arise from simple regression effects. However, this is not an issue in our experiments because proportion correct (about 50%), in both Study I and II lies exactly in the middle of the confidence scale (which ranges from 0 to 100%). Regression effects may be an issue when proportion correct is either high or low. Another potential problem is so called scale-end effects. If, as in Juslin et al., the scale ranges from 50% (guessing) to 100% (absolutely sure) a proportion correct of 50% would only allow overconfidence. A proportion correct of about 100% would instead only allow underconfidence. Therefore, any response error in the participants' judgments would artificially produce over- or underconfidence depending on proportion correct. Again, because odor naming performance in our studies was about 50%, such floor or ceiling effects are not an issue.

Another potential problem in many previous studies is that the choice of test items for inclusion in these studies had not been representative. To clarify, a representative sample of items would be a *random* rather than *selected* sample of all possible items available for a particular domain. In a metaanalysis of most of the studies in the field (up to 1998) they showed that there

was no general overconfidence in the studies using representative samples, but a significant overconfidence effect was found in studies using non-representative (i.e., selected) samples. This potential problem was highlighted in Study II, where a lower overconfidence was found than in Study I. We showed that when we in Study I analyzed only the odor sample included in Study II[5] (see Jönsson et al., 2005), the level of overconfidence was non-significantly different between the two studies. This indicates that sampling error could also be a potential reason for at least a part of the overconfidence in odor naming found here. Future studies should include cross-modal comparisons. If, for example, overconfidence would be found across modalities, but with a *significantly higher* overconfidence for odors, it would further strengthen the conclusion about overconfidence in odor naming.

The odor naming process

Engen (1991, p. 85) wrote that "At times one cannot come up with a word for an odor at all, even though one knows what the odor is /…/. We have referred to this as the 'tip-of-the-nose' state". The main argument here contrasts with this view. It has been proposed that most unnamable odors are largely also unidentified odors, that is, the participants still *do not know what odor it is*.

More precisely, regarding the three stages in the naming process, our data indicate that odor naming fails in the first stage, the object identification stage (Study III). Actually, because common everyday odors were used in Study III, it could be predicted that odor identification failures would increase even further if less familiar odors would have been used. It is unclear why odor naming so often fails in this stage. One reasonable cause can be a "fuzzy perception" of odors as indicated by our relatively poor odor discrimination ability. Cain and Potts (1996) argued that errors in odor naming are predominantly perceptual, since participants when their naming attempts were incorrect, 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. Similarly, the frequency of the immediate naming attempts in both experiments of Study III (Table 2) showed that the participants did not have any problem in emitting possible odor names, but compared to the person names they were significantly less accurate (despite instructions being identical). To visually discriminate between for example grapefruit, lemon, orange, and lime is not especially difficult, but to do so by smell is.

The absence of structural-phonological information in connection with odor naming failures is an olfactory peculiarity. However, the assumption

[5] All odors in Study II were also included in the larger odor sample in Study I.

that it is due to a poor *odor-name* link is challenged by the present findings. We showed that it can instead be due to the high amount of object identification failures. The logic is simple: What is not known can normally not be named. The finding that only about 5% of the odors elicited TOT experiences is inconsistent with the idea of a poor link and in particular with the assumption that the odor-name association is uniquely weak, compared to other modalities.

It should be noted that rare and infrequent words are better in eliciting TOTs than highly frequent words (Vigliocco, 2004). Because common everyday odors are typically used in odor naming experiments (because otherwise the performance would drop dramatically) the names of the odorous objects (e.g., orange, apple, tar, baby powder, tobacco, coffee) hardly represent highly infrequent words. It can therefore be questioned whether the odor-elicited TOT experiences in Study III are phenomenologically identical to TOTs that elicit structural-phonological clues (see e.g., Brown, 1991). Future studies should utilize smells representing objects whose names are infrequent, if it is at all possible to achieve such a combination. Yet another possibility is that some of these TOTs were not "real" TOTs as defined here, but again very strong FOKs associated with object identification failures (which would mean that the TOT incidence of 5% is actually an overestimation). That as least some of the reported TOTs were actually still unidentified is supported by a further analysis of the data in Experiment 1 of Study III. For example, four out of the 56 reported TOT experiences were for the fruits orange, apple and lemon. It is quite unlikely that these TOTs mirror simple name activation failures (but rather object identification failures). If one would present these fruits to participants visually, they surely would be able to name them.

Is odor metamemory special?

Odor memory has often been proposed to be special compared to memory for other modalities. Can it, based on the results so far, be concluded that olfactory metamemory is special in some respect? The picture that emerges is one of lower predictive validity of the odor-elicited metamemory judgments. More cross-modal studies have to be made, for example with auditory stimuli and pictures of objects. Future research should also further delineate the reason for why it is so difficult to name odors and in particular why the structural-phonological information is absent when people report TON experiences. The present research has led to some suggestions concerning this, as noted above.

The specificity of the tip of the nose phenomenon

The TON experience (Lawless & Engen, 1977) has often been proposed to be an olfactory analogue of the TOT experience. The two phenomena are clearly not synonymous, but what is the essence of the difference? Lawless and Engen (1977) put forward two motivations for their invention of a separate term. First, TON experiences are not associated with access to structural-phonological information about the missing word (see also Study I). Second, if the participants had still not resolved their TON experience after filling out a questionnaire, they were given the dictionary definition of the missing odorous object. This hint led to correct odor name retrieval in 16 out of 23 instances. The authors argued for this to be another difference between the TOT and TON experience, because reading "the definition of a word can induce the TOT state, but usually solves the problem in the TON state (p.57-58)". This, however, is not a valid difference, because giving a person with a TOT experience an additional cue, would most probably inflate the resolution rates also for TOTs, as it does for TONs.

Apart from the absence of structural-phonological information what signifies a TON experience seems to be that it typically is not an imminent feeling of retrieving a missing odor name (which would make it a TOT experience). We propose that a more accurate definition of the TON experience is a *strong and imminent feeling of resolving the identity of a familiar odor*. As can be seen in and Table 3 (Study III), odors seem to be particularly apt in eliciting strong feelings of knowing, despite the fact that the naming attempts fail early in the naming process (as much as 54% of the unnamed odors led to strong to very strong FOKs about the identity of the odor). The TON experience thus seems to consist of mainly strong FOKs and a few TOT experiences.

Concluding remarks

Future research should also focus on more applied issues. As previously noted, empirical evidence indicates that in some populations, assessments of olfactory ability deviates from actual ability (Nordin et al., 1995; White & Kurtz, 2003). It is important to further corroborate these findings with stronger methodology, and fully understand when those judgments deviate and why they do so, as well as the potential impact this has on peoples' everyday life. In this thesis work I have tried to elucidate different aspects of olfactory metacognition, a research field that up to now has been close to absent. Hopefully, the thesis has filled in a part of that gap and it has been shown that a metacognitive perspective can aid in understanding odor identification.

Acknowledgements

First and foremost I would like to thank my family and parents for their strong and invaluable support during the years. I am deeply grateful to my wife (Natalya) for always encouraging me and for enduring the constant conflict between being a "family man" and a scientist. No words are enough to describe the support I got from my daughter (Nikolina), always being such a sunshine, so wonderful and happy and simply for being the best child in the world! Without them nothing of this would be possible.

I am deeply indebted to Mats Olsson, my supervisor and mentor, not only for being the great guy he is, but also for always having time for questions (even when he does not have it) and giving extremely valuable and helpful feedback on the manuscripts. I have really enjoyed the thought-provoking discussions we have had over the years as well as the company.

I would also like to thank Henrik Olsson (my assistant supervisor), Peter Juslin (my reviewer at the "full-time seminar"; special thanks for the insightful comments about the statistical analyses), Erik Börjesson (my reviewer at the "half-time seminar"), Maria Larsson and all the anonymous reviewers of my manuscripts for ingenious and helpful comments and constructive criticisms during different stages of the thesis work. Special thanks should be directed to Henrik Olsson for interesting discussions and helpful comments on the thesis, but also for always having time for chatting about things that are rather "off-topic" when I got bored sitting next to the computer (quite often).

I am indebted to Professor Thomas Nelson at the University of Maryland (who is no longer with us due to a too early death in the beginning of 2005), with whom I spent some months in 2001, for advancing my understanding of metacognition and for being really supportive during my time in the USA. Thanks for all the good tips I got about what it means to be a scientist.

Thanks also to the leadership of the department and all the administrative and technical personnel, without whom the department would not function. I also very much enjoyed the pleasant company of and talks with my colleagues Christina Blåvarg, Johan Lundström (my research group teammate), Nils Olsson, Leo Poom, Nazar Akrami, Timo Hursti and many others. Thanks also to all, whom I unintentionally forgot to mention. The list could be made so much longer…

Fredrik Jönsson, Uppsala, June 2005

References

Alaoui-Ismaïli, O., Vernet-Maury, E., Dittmar, A., Delhomme, G., & Chanel, J. (1997a). Odor hedonics: connection with emotional response estimated by autonomic parameters. *Chemical Senses*, 22, 237-248.

Alaoui-Ismaïli, O., Robin, O., Rada, H., Dittmar, A., & Vernet-Maury, E. (1997b). Basic emotions evoked by odorants: comparison between autonomic responses and self-evaluation. *Physiology and Behavior*, 62, 713-720.

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., Gabrieli, J. D., & Sobel, N. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature and Neuroscience*, 6, 196-202.

Annett, J. M. (1996). Olfactory memory: A case study in cognitive psychology. *Journal of Psychology, 130*, 309-319.

Bothwell, R. K., Deffenbacher, K. A., & Brigham, J. C. (1987). Correlation of eyewitness accuracy and confidence: Opitmality hypothesis revisited. *Journal of Applied Psychology*, 72, 691-695.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25, 49-59.

Brigham, J. C., Maass, A., Martinez, D., & Whittenberger, G. (1983). The effect of arousal on facial recognition. *Basic Applied Social psychology*, 4, 279-293.

Brown, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 5, 325-337.

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

Chu, S., & Downes, J. J. (2000). Odour-evoked autobiographical memories: psychological investigations of Proustian phenomena. *Chemical Senses*, 25, 111-116.

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

Cain, W. S., de Wijk, R., Lulejian, C., Schiet, F., & See, L. C. (1998). Odor identification: Perceptual and semantic dimensions. *Chemical Senses, 23*, 309-326.

De Wijk, R. A., & Cain, W. S. (1994a). Odor identification by name and edibility: Life-span development and safety. *Human factors*, 36, 182-187

De Wijk, R. A., & Cain, W. S. (1994b). Odor quality: Discrimination versus free and cued identification. *Perception and Psychophysics*, 56, 12-18.

De Wijk, R. A., Schab, F. R., & Cain, W. S. (1995). Odor identification. In F. R. Schab, & R. G. Crowder (Eds.), *Memory for odors* (pp. 21-37). Mahwah, NJ: Lawrence Erlbaum associates.

Distel, H., & Hudson, R. (2001). Judgement of odor intensity is influenced by subjects' knowledge of the odor source. *Chemical Senses*, 26, 247-251.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (jol) and the delayed-jol effect. *Memory & Cognition*, 20, 374-380.

Dunlosky, J., & Nelson, T. O. (1994). Does the sensitivity of judgments of learning (jols) to the effects of various study activities depend on when the jols occur. *Journal of Memory and Language*, 33, 545-565.
Engen, T. (1987). Remembering odors and their names. *American Scientist*, 75, 497-503.
Engen, T. (1991). *Odor sensation and memory*. New York: Praeger Publishers.
Eskenazi, B., Cain, W. S., Novelly, R. A., & Friend, K. B. (1983). Olfactory functioning in temporal lobectomy patients. *Neuropsychologia*, 21, 365-374.
Eskenazi, B., Cain, W. S., & Friend, K. B. (1986). Exploration of olfactory aptitude. *Bulletin of the Psychonomic Society*, 24, 203-206.
Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, 24, 411-435.
Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, 56, 208-216.
Herz, R. S., & Cupchik, G. C. (1995). The emotional distinctiveness of odor-evoked memories. *Chemical Senses*, 20, 517-528.
Herz, R. S., & Engen, T. (1996). Odor memory: Review and analysis. *Psychonomic Bulletin and Review*, 3, 300-313.
Herz, R. S. (1998a). Are odors the best cues to memory? A cross-modal comparison of associative memory stimuli. In C. Murphy (Ed.), *Olfaction and Taste XII*, 855, pp 670-674, New York Academies of Sciences.
Herz, R. S. (1998b). An examination of objective and subjective measures of experience associated to odors, music and paintings. *Empirical Studies of the Arts*, 16, 137-152.
Herz, R. S. (2000). Scents of time - A Proustian memory triggered by an odor can be emotionally overwhelming - but it can also be misleading. *Sciences*, 40, 34-39.
Herz, R. S., Eliassen, J., Beland, S., & Souza, T. (2004). Neuroimaging evidence for the emotional potency of odor-evoked memory. *Neuropsychologia*, 42, 371-378.
Hinton, P. B., & Henley, T. B. (1993). Cognitive and Affective Components of Stimuli Presented in 3 Modes. *Bulletin of the Psychonomic Society,* 31, 595-598.
Holley, A. (2002). Cognitive aspects of olfaction in perfumer practice. In C. Rouby, B. Schaal, D. Dubois, R. Gervais, & A. Holley (Eds.). *Olfaction, taste and cognition* (pp.16-26), Cambridge University Press.
Hosch, H. M., & Bothwell, R. K. (1990). Arousal, description and identification accuracy of victims and bystanders. *Journal of Social Behavioral and Personality*, 5, 481-488.
Houdé, O. (2004). *Dictionary of Cognitive Science: Neuroscience, Psychology, Artificial Intelligence, Linguistics, and Philosophy.* New York: Taylor and Francis.
Hoffman, J. (2005). Suspect memories. *Legal Affairs Magazine*, 5.
Hughson, A. L. (2003). Wine expertise: current theories and findings regarding its nature and bases. *Food Australia*, 55, 193-196.
Ikeda, K. (2002). New seasonings. *Chemical Senses*, 27, 847-849.
James, W. (1890) *Principles of psychology*. New York: Holt.
Jehl, C., Royet, J. P., & Holley, A. (1995). Odor discrimination and recognition memory as a function of familiarization. *Perception and Psychophysics*, 57. 1002-11.
Johnson, C. J., Paivio, A., & Clark, J. M. (1996). Cognitive components of picture naming. *Psychological Bulletin*, 120, 113-139.
Juslin, P., & Olsson, H. (1997). Thurstonian and Brunswikian origins of uncertainty in judgment: A sampling model of confidence in sensory discrimination. *Psychological Review*, 104, 344-366.

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

Jönsson, F. U., & Olsson, M. J. (2003). Olfactory metacognition. *Chemical Senses*, 28, 651-658.

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

Jönsson, F. U., Tcheckova, A., Lönner, P., & Olsson, M. J. (in press). A metamemory perspective on odor naming and identification. *Chemical Senses*.

Kaitz, M., Good, A., Rokem, A.M., & Eodelman, A.L. (1987). Mothers' recognition of their newborns by olfactory cues. *Developmental Psychobiology*, 20, 587-591.

Kelemen, W. L., & Weaver, C. A. (1997). Enhanced metamemory at delays: Why do judgments of learning improve over time? *Journal of Experimental Psychology: Learning Memory and Cognition*, 23, 1394-1409.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general information questions. *Journal of memory and Language*, 32, 1-24.

Koriat, A., Lichtenstein, S, & Fischoff, B. (1980). Reasons or confidence. *Journal of Experimental Psychology*, 6, 197-118.

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

Koriat, A. (1994). Memory's knowledge about its own knowledge: The accessibility account of the feeling-of-knowing. In J. Metcalfe & A. P. Shimamura (Eds). *Metacognition: Knowing about knowing* (pp. 115-135). Cambridge, MA: Bradford.

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General*, 124, 311-333.

Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 34-53.

Koriat, A., Levy-Sadot, R., Edry, E., & de Marcas, S. (2003). What do we know about what we cannot remember? Accessing the semantic attributes of words that cannot be recalled. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1095-1105.

Krinsky, R., & Nelson, T. O. (1985). The feeling of knowing for different types of retrieval failure. *Acta Psychologica*, 58, 141-158.

Larsson, M., Nilsson, L-G, & Olofsson, J. K., & Nordin, S. (2004). Demographic and cognitive predictors of cued odor identification: Evidence from a population-based study. *Chemical Senses*, 29, 547-554.

Lawless, H., & Engen, T. (1977). Associations to odors: Interference, mnemonics, and verbal labeling. *Journal of Experimental Psychology: Human Learning and Memory*, 3, 52-59.

Leibert, T. W., & Nelson, D. L. (1998). The roles of cue and target familiarity in making feeling of knowing judgments. *American Journal of Psychology*, 111, 63-75

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of subjective probabilities: The state of the art up to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306-334). New York: Cambridge University Press.

Lindsay, D., Read, J. D., & Sharma, K. (1998). Accuracy and confidence in witness identification: The relationship is strong when witnessing conditions vary. *Psychological Science*, 9, 215-218.

Lorig, (1999). On the similiarity of odor and language perception. *Neuroscience and Biobehavioral Reviews*, 23, 391-389.

McClelland, A. G. R., & Bolger, F. (1994) The calibration of subjective probabilities: Theories and models 1980-1993. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 453-482). Chichester: Wiley.

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 851-864.

Metcalfe, J. (2000). *Metamemory: Theory and data*. In E. Tulving and F. I. M. Craik (Eds.), The Oxford handbook of memory (pp. 197-211). New York: Oxford University Press.

Moskowitz, H. R. (1979). Mind, body and pleasure: an analysis of factors which influence sensory hedonics. In J. H. A. Kroeze (Ed.). *Preference, behavior and chemoreception* (pp. 131-141). London: IRL Press.

Mozell, M. M., Smith, B. P., Smith, P. E., Sullivan, R. L., & Swender, P. (1969). Nasal chemoreception in flavor identification. *Archives of Otolaryngology*, 90, 367-73.

Murphy, C., Cain, W. S., Gilmore, M. M., & Skinner, B. R. (1991). Sensory and semantic factors in recognition memory for odors and graphic stimuli: Elderly versus young persons. *American Journal of Psychology*, 104, 161-191.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and some new findings. In G. H. Bower (Ed.). *The psychology of learning and motivation* (pp.1-45). New York: Academic Press.

Nelson, T. O., & Dunlosky, J. (1991). When peoples judgments of learning (jols) are extremely accurate at predicting subsequent recall - the delayed-jol effect. *Psychological Science*, 2, 267-270.

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

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist,* 51, 102-116.

Nhouyvanisvong, A., & Reder, L. M. (1998). Rapid feeling-of-knowing: A strategy selection mechanism. In V. Y. Yzerbyt , G. Lories, & B. Dardenne (Eds.). *Metacognition: Cognitive and social dimensions* (pp.35-52). Thousand Oaks: Sage Publications.

Nordin, S., Monsch, A. U., & Murphy, C. (1995). Unawareness of smell loss in normal aging and Alzheimer's disease: Discrepancy between self-reported and diagnosed smell sensitivity. *Journal of Gerontology: Psychological sciences,* 50B, 187-192.

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

Parr, W. V., Heatherbell, D., & Geoffrey, W. K. (2002). Demystifying wine expertise: Olfactory threshold, perceptual skill and semantic memory in expert and novice wine judges. *Chemical Senses*, 27, 747-755.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45, 255-287.

Perfect, T. J., & Hollins, T. S. (1999). Feeling-of-knowing judgments do not predict subsequent recognition performance for eyewitness memory. *Journal of Experimental Psychology: Applied*, 5, 250-264.

Peynircioglu, Z. F., Tekcan, A. I., Wagner, J. L., Baxter, T. L., & Shaffer, S. D. (1998). Name or hum that tune: Feeling of knowing for music. *Memory & Cognition*, 26, 1131-1137.

Proust, M. (1919). *Du côté de chez Swann.* Paris: Gaillimard.

Rabin, M. D. (1988). Experience facilitates olfactory quality discrimination. *Perception and Psychophysics*, 44, 532-540.

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology,* 19, 90-138.

Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 435-451.

Roberts, M. & Erdos, G. (1993). Strategy selection and metacognition. *Educational Psychology*, 13, 259-266.

Robin, O., Alaoui-Ismaili, O., Dittmar, A., & Vernet-Maury, E. (1999). Basic emotions evoked by eugenol odor differ according to the dental experience. A neurovegetative analysis. *Chemical Senses*, 24, 327-335.

Schab, F. R., & Cain, W. S. (1991). Memory for odors. In D. G. Laing, R. L. Dory, & W. Breipohl (Eds.), *The human sense of smell* (pp. 217-240). New York: Springer.

Schmolck, H., Buffalo, E. A., & Squire, L. R. (2000). Memory distortions develop over time: recollections of the O.J. Simpson trial verdict after 15 and 32 months. *Psychological Science*, 11, 39-45.

Schwartz, B. L., Benjamin, A. S., & Bjork, R. A. (1997). The inferential and experiential bases of metamemory. *Current Directions in Psychological Science*, 6, 132-137.

Schwartz, B. L., & Metcalfe, J. (1992). Cue familiarity but not target retrievability enhances feeling-of-knowing judgments. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 18, 1074-1083.

Schwartz, B. L., & Metcalfe, J. (1994). Methodological problems and pitfalls in the study of human metacognition. In J. Metcalfe & A. P. Shimamura (eds.). *Metacognition: knowing about knowing* (pp. 93-113). London: The MIT Press.

Schwartz, B. L. (1999). Sparkling at the end of the tongue: The etiology of tip-of-the-tongue states: A diary study. *Psychonomic Bulletin & Review*, 6, 379-393.

Schwartz, B. L. (2002). *Tip-of-the-tongue states: Phenomenology, mechanism, and lexical retrieval.* Mahwah, NJ: Lawrence Erlbaum Associates.

Sniezek, J. A., Paese, P. W., & Switzer, F. S. C. (1990). The effects of choosing on confidence in choice. *Organizational Behavior and Human Decision processes*, 46, 264-282.

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 204-221.

Stevens, J. C., Cain, W. S., & Weinstein, D. E. (1987). Aging impairs the ability to detect gas odor. *Fire Technology*, 23, 198-204.

Sulmont-Rossé, C., Issanchou, S., & Köster, E. P. (2005). Odor naming methodology: correct identification with multiple-choice versus repeatable identification in a free task. *Chemical Senses*, 30, 23-7.

Talarico, J. M., & Rubin, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science*, 14, 455-461.

Thiede, K. W., & Dunlosky, J. (1994). Delaying students metacognitive monitoring improves their accuracy in predicting their recognition performance. *Journal of Educational Psychology*, 86, 290-302.

Todrank, J., Byrnes, D., Wrzesniewski, A., & Rozin, P. (1995). Odors can change preferences for people in photographs - a cross-modal evaluative conditioning study with Olfactory Uss and Visual Css. *Learning and Motivation*, 26, 116-140.

van Reekum, C. M., van den Berg, H., & Frijda, N. H. (1999). Cross-modal preference acquisition: evaluative conditioning of pictures by affective olfactory and auditory cues. *Cognition and Emotion*, 13, 831-836.

Vesonder, G. T., & Voss, J. F. (1985). On the ability to predict one's own responses while learning. *Journal of Memory and Language*, 24, 363-376.

Vigliocco, G. (2001). Tip-of-the-tongue, Psychology of. In N. J. Smelser & P. B. Baltes (Eds.). *International encyclopedia of the social & behavioral sciences* (pp.15759-15762). Oxford: Elsevier Science Ltd.

Weaver, C. A., & Kelemen, W. L. (1997). Judgement of learning at delays: Shifts in response patterns or increased metamemory accuracy? *Psychological Science*, 8, 318-321.

Weaver, C. A. (1993). Do you need a "flash" to form a flashbulb memory? *Journal of Experimental Psychology: General*, 122, 39-46.

White, T. L., & Kurtz, D. B. (2003). The relationship between metacognitive awareness of olfactory ability and age in people reporting chemosensory disturbances. *American journal of psychology*, 116, 99-110.

Willander, J., & Larsson, M. (2005). Smell your way back to childhood: Autobiographic odor memory. *Unpublished manuscript.*

Wilson, D. A., Best, A. R., & Sullivan, R. M. (2004). Plasticity in the olfactory system: Lessons for the neurobiology of memory. *The Neuroscientist*, 10, 513-524.

Wise, P. M., Olsson, M. J., & Cain, W. S. (2000). Quantification of odor quality. *Chemical Senses, 25*, 429-443.

Yates, J. F. (1990). *Judgment and decision making.* Englewood Cliffs, NJ: Prentice Hall.

Zakay, D., & Tuvia, R. (1998). Choice latency times as determinants of post-decisional confidence. *Acta Psychologica*, 98, 103-115.

Zald, D. H., & Pardo, J. V. (2000). Functional neuroimaging of the olfactory system in humans. *International Journal of Psychophysiology*, 36, 165-181.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 6

Editor: The Dean of the Faculty of Social Sciences

A doctoral dissertation from the Faculty of Social Sciences, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-5821

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2005