

Age-related changes in discrimination of unfamiliar odors

RICHARD J. STEVENSON, NINA SUNDQVIST, AND MEHMET MAHMUT
Macquarie University, Sydney, New South Wales, Australia

If odor perception involves mnemonic processes, differences in olfactory experience should affect discriminative ability. This was examined here by comparing discriminative performance in children and adults. Using an oddity test of discrimination, in Experiment 1 we tested 6-year-olds (G1), 11-year-olds (G6), and adults (A) on their ability to discriminate unfamiliar odors that varied either in quality (Q) or in quality and intensity (QI). G1 participants were poorer at discriminating the QI set, relative to G6 and A. In Experiment 2, we used an analogous visual procedure and confirmed that this age-related difference was olfactory specific. In Experiment 3, we repeated Experiment 1 but used an articulatory suppression task. G1 participants were poorer than G6 and A participants for both the Q and the QI sets. The implications of these findings for experiential accounts of odor perception and olfactory working memory are discussed.

Most natural odors are complex mixtures (Maarse, 1991), and the olfactory system has to recognize these against a continuously shifting chemical background. Several authors have suggested that the brain accomplishes this by matching the input pattern of olfactory stimulation against previously encoded patterns; that is, the olfactory system engages in a form of object recognition (Haberly, 2001; Hasselmo, Anderson, & Bower, 1992; Hudson, 1999; Stevenson & Boakes, 2003; Wilson & Stevenson, 2003). As the name *recognition* implies, learning and memory should play a central role in olfactory perception, and the empirical literature bears this out (see Wilson & Stevenson, 2003). If learning is important in olfactory perception by virtue of encoding new odors, one might expect that greater exposure to odors should translate into an improved ability to discriminate between them. The aim of the experiments reported in this article was to test this hypothesis by exploring whether adult participants, who have more experience with odors, are better at odor discrimination, relative to those with less experience—namely, children.

Odor-naming ability is significantly correlated with the ability to discriminate between odors (De Wijk & Cain, 1994), and odor naming has been consistently observed as poorer in children than in adults (e.g., Cain et al., 1995; Doty et al., 1984). That discrimination may account for some of this age-related difference in naming was suggested by Cain et al. (Experiment 3). They measured the speed with which children and adults learned labels for both familiar and unfamiliar odors. The youngest children (6-year-olds) were markedly worse at learning names for unfamiliar odors, relative to the other age groups, but performance was similar for familiar odors. Cain et al. interpreted this result as reflecting poorer dis-

criminative ability for the unfamiliar odors in the youngest children.

In children, the evidence favoring poorer discriminative ability is closely linked with the ability to give their olfactory experience a name. This ability may enhance performance on a test of discrimination, since a participant may encode each stimulus both as a percept and as a verbal label. This may have two consequences. First, some of the difference in naming ability between adults and children must result from lack of appropriate labels, rather than from lack of experience with the odor. Consequently, younger children's ability to solve a discrimination task will be constrained by greater reliance on perceptual encoding. Second, although odor naming is better in adults, it is still rather poor (Cain, 1979). If adults rely on both verbal labels and perceptual encoding to solve an olfactory discrimination problem, this joint approach may be worse than sole reliance on perceptual encoding. Exactly this type of effect has been observed (Melcher & Schooler, 1996).

The problem that we address here is whether odor discrimination is poorer in children than in adults. In Experiment 1, we examined this by using an oddity test of discrimination. In Experiment 2, we used the same design, but with visual stimuli, so as to compare children's and adults' performance on a task for which we had no reason to expect major differences. In Experiment 3, we focused on the role of odor naming by using an articulatory suppression task, thus forcing reliance on a perceptual strategy.

EXPERIMENT 1

In Experiment 1, we set out to test differences in olfactory discriminative ability between adults and children.

R. J. Stevenson, richard.stevenson@psy.mq.edu.au

Table 1
Odor Stimuli Used in Experiments 1 and 3

| Odorant | Manufacturer | Standard | Higher |
|------------------------|--------------|------------------------|------------------------|
| | | Concentration (g/L) | Concentration (g/L) |
| Patchouli | Oil Garden | 0.008 | 0.8 |
| Anisaldehyde | Sigma | 0.088 | 8.8 |
| Lavender | Oil Garden | 0.006 | 0.6 |
| Ylang ylang | Oil Garden | 0.03 | 3.0 |
| Water chestnut | Quest | 0.04 | 4.0 |
| Methyl anthranilate | Sigma | 0.00013 | 0.013 |
| Red bean | Quest | 0.036 | 3.6 |
| Linalyl acetate | Sigma | 0.016 | 1.6 |
| Acetyl methyl carbamol | Dragoco | 0.01 | 1.0 |
| Bornyl acetate | Sigma | 0.038 | 3.8 |

In an attempt to make naming difficult for all the participants, we selected and piloted a set of 10 unfamiliar, hedonically similar, hard-to-name odors, each at two concentrations, standard and higher (see Table 1). Pilot work ($n = 12$) established that the higher concentration differed significantly in intensity from the standard concentration [$F(1,11) = 31.66$, $MS_e = 3.89$, $p < .001$] but that odors did not differ significantly in intensity within the standard or the higher set. The standard set formed the first condition in Experiment 1, in which discrimination between odorants was based primarily on perceptual quality—the quality condition. Both the standard and the higher set were used in a second condition—the quality and intensity condition—so as to produce a more easily discriminable set. This condition was characterized by the inclusion either of two standard concentration odors (both of Quality A) and one higher concentration odor (of Quality B) or vice versa. This condition affords two cues to discrimination, perceptual quality and intensity. Prior studies have suggested, albeit with nonolfactory stimuli, that such additional cues may facilitate discrimination (e.g., Menier, Forget, & Lambert, 1996; Sinha, 1990).

Three age groups were selected: 6-year-olds, 11-year-olds, and adults (17+). Six-year-olds demonstrate poorer performance in acquiring odor labels, and 11-year-olds show poorer odor naming, relative to adults. Both groups, therefore, might show deficits in discrimination. Finally, we included an equal number of males and females, since females often outperform males on many olfactory tasks (Brand & Millot, 2001).

Method

Participants. One hundred thirty participants took part in Experiment 1 (see Table 2 for age and gender details). The adult participants were recruited from first-year psychology students and Macquarie University students generally and took part for course credit or a small cash payment (\$AU10). The child participants were recruited from two primary schools in the northern region of Sydney. These schools, along with the others participating in subsequent experiments, drew their children from middle class communities of primarily Northern European descent—as with the university sample too. No participant had served in a related study, as with all of the experiments reported here.

Stimuli. In total, 10 different odors were used, each at two different concentrations (see Table 1). A major constraint in stimulus selection was to ensure that no odorant could potentially harm or be perceived to harm any of the children taking part in the study—

hence, the rather eclectic selection. Each odor was initially dissolved in propylene glycol and then diluted to the requisite concentration with water. Odors were presented in colored opaque plastic squeeze bottles (250 ml), each of which contained 50 ml of the odorant. The snout of the bottle was attached to a 15-cm odorless polypropylene tube, enabling the experimenter to present all the stimuli to the participant in a standardized manner. In addition, the tube was required for the Grade 1 children, to enable the bottles to be disguised within a cloth frog puppet, with the tube protruding from its head (more below). Bottles were identified by three-digit codes written on their base. These were never visible during testing.

Procedure. All testing was conducted by two experimenters, due to child protection requirements. One administered the odorants, and the other recorded responses and timed the intervals within and between trials. For children, parental consent was obtained prior to testing at a given school. For Grade 1 children (6-year-olds), verbal assent was obtained after the procedure had first been explained. For Grade 6 children (11-year-olds) and adults, written consent was obtained. The participants were then asked whether they currently had a cold and, if so, whether they were still able to smell.

All the participants were then given the following instructions. “You will be asked to sniff three smells. Two of the smells will be the same and one will be different. Your job is to pick the odd one out.” The practice trial then began. This consisted of two bottles filled with water and a further bottle containing 50% water-diluted vinegar (2% acetic acid). This combination was selected because it relied on a trigeminal cue to discrimination that can be detected by anosmic participants (Laska, Distel, & Hudson, 1997). Thus, it would be easy for all the age groups to detect. On the first practice trial, the odd stimulus (vinegar) was always presented last. The spout of the first bottle was placed 7 cm from the participant’s nose, and the experimenter gave three deep squeezes of the bottle in quick succession. The participant was then asked whether he or she would like another sniff. If he or she said “yes,” a further two puffs were provided. If he or she said “no,” the bottle was placed in a display rack (numbered 1, 2, or 3) in front of the participant. If the participant asked for a further sniff, he or she was provided with two more puffs. The participant was then not allowed any further exposure to that stimulus.

Following the final puff, a 20-sec interval separated the first stimulus from the second. The second was then presented in the same way, followed by the third stimulus. The participant was then asked to select which of the three was the odd one out. This constituted a trial. Although this procedure was essentially the same for all the age groups, the Grade 1 children had a slightly different form of

Table 2
Participant Characteristics, Mean Age (in Decimals), and Number for Experiments 1, 2, and 3

| Group | Experiment 1 | | Experiment 2 | | Experiment 3 | |
|-----------------------|--------------|-----|--------------|-----|--------------|-----|
| | Age | No. | Age | No. | Age | No. |
| Quality and Intensity | | | | | | |
| Grade 1 male | 6.2 | 11 | 6.5 | 11 | 6.6 | 9 |
| Grade 1 female | 6.3 | 8 | 6.1 | 8 | 6.4 | 10 |
| Grade 6 male | 11.3 | 10 | 11.3 | 10 | 11.4 | 11 |
| Grade 6 female | 11.2 | 11 | 11.3 | 11 | 11.4 | 8 |
| Adult male | 23.0 | 11 | 24.4 | 8 | 26.3 | 9 |
| Adult female | 19.5 | 11 | 23.5 | 8 | 25.7 | 10 |
| Quality | | | | | | |
| Grade 1 male | 6.4 | 11 | 6.5 | 11 | 6.4 | 10 |
| Grade 1 female | 6.2 | 9 | 6.4 | 8 | 6.3 | 10 |
| Grade 6 male | 11.3 | 11 | 11.3 | 10 | 11.4 | 11 |
| Grade 6 female | 11.3 | 11 | 11.3 | 12 | 11.5 | 9 |
| Adult male | 23.2 | 11 | 24.8 | 8 | 23.5 | 10 |
| Adult female | 21.2 | 13 | 26.0 | 7 | 26.6 | 10 |

presentation, in that the discrimination task was presented as a game in which the children sniffed three frogs and had to pick which of the three smelled different. The same instructions were employed, but the word *frog* was substituted for *smell* in the first sentence and the word *frog* was inserted before the word *smells* in the second. In all other respects, the procedures were identical. Following completion of the first practice trial, a 1-min interval occurred prior to the start of the next trial.

If a participant successfully identified the odd one out on the practice trial, a further five triangle test trials, using the same procedure as that detailed above, were then completed. If, however, he or she incorrectly identified the odd stimulus on the practice trial, the practice was repeated a second time, but now with the odd stimulus (vinegar) presented first. If the second trial was completed successfully, the participant continued with the remaining five test trials. If he or she failed the second practice trial, the remaining five trials were still completed, but the participant's data were not used.

Presentation patterns of the odorants and their position on each trial were organized in the following way. First, a master plan composed of 24 patterns (where each pattern corresponded to the five trials completed by 1 participant) was devised (see Diagram 1). This was then employed in each of the six cells of the design (i.e., age group by condition; see below). All 10 odors were employed in each pattern, and whether a particular odor occurred once or twice within a given trial was equated across the 24 patterns (see Diagram 1 for examples). Second, for each participant, across the five trials, the odd stimulus always occurred, *on average*, in the second position (i.e., the odd stimulus on any trial could be presented first, second, or third). Twelve patterns had one presentation of the odd stimulus first, one third, and three second (e.g., Pattern 2 in Diagram 1), whereas the other 12 had two presentations of the odd stimulus first, two third, and one second (e.g., Pattern 1 in Diagram 1). The order in which each of these trial types appeared within a particular pattern was randomized for each of the 24 patterns. Third, on average, any odorant was approximately as likely to be presented with any other on a particular trial.

Finally, in addition to the three age groups, the design employed two separate experimental conditions: quality, and quality and intensity. In the quality condition, stimuli were selected only from the standard concentration column in Table 1 and were known from our pilot work to be of similar intensity (e.g., a trial might consist of standard anisaldehyde vs. standard patchouli vs. standard patchouli). In the quality and intensity condition, however, although two qualitatively different odors were again used on each trial, one (or two) was selected from the standard concentration column, whereas the other two (or one) were selected from the strong concentration column in Table 1 (e.g., a trial might consist of standard anisaldehyde vs. strong patchouli vs. strong patchouli). As with the choice of which odors to present together on a trial, whether a particular odorant was drawn from the strong or the standard concentration was again equated across the 24 patterns.

Results and Discussion

Nineteen participants reported having colds, but all claimed that they could still smell. Since their performance did not differ significantly from the remaining participants', who did not report colds, as with Experiment 3, all were included in the analyses. Two participants were excluded from the study (a Grade 1 and a Grade 6 student), since they failed both practice trials. Of the remaining participants, 2 Grade 6 students and 1 adult successfully completed a second practice trial.

Before turning to the results in full, it is instructive to see on which trials the participants made errors, since this might indicate whether the youngest participants had any task-related difficulties. This might manifest in at least

Table 3
Errors by Trial, for Each Condition, by Age Group
in Experiment 1

| Condition | Trial | Grade 1 (%) | Grade 6 (%) | Adult (%) |
|-----------------------|--------|-------------|-------------|-----------|
| Quality | First | 17.6 | 24.3 | 23.6 |
| | Second | 11.8 | 16.2 | 16.4 |
| | Third | 20.6 | 21.6 | 20.0 |
| | Fourth | 23.5 | 21.6 | 16.4 |
| | Fifth | 26.5 | 16.3 | 23.6 |
| | Total | 100.0 | 100.0 | 100.0 |
| Quality and intensity | First | 19.5 | 6.2 | 27.8 |
| | Second | 21.9 | 37.5 | 22.2 |
| | Third | 12.4 | 18.8 | 11.1 |
| | Fourth | 21.9 | 12.5 | 22.2 |
| | Fifth | 24.3 | 25.0 | 16.6 |
| | Total | 100.0 | 100.0 | 100.0 |

two ways: fatigue, so that errors might be more likely on the later trials, or not fully comprehending the instructions, in which case errors might occur predominantly in the earlier trials. We would expect an error rate of approximately 20% on each trial by chance alone, and this is what we observed (see Table 3). A further method for detecting any task-related difficulty was to see whether the youngest children showed a position bias in their response (irrespective of whether they were correct or incorrect). To detect any position bias, we calculated (1) differences in their response position irrespective of correctness, between groups, by condition and (2) the deviation from the expected number of responses in each position (i.e., correct responses) from those actually obtained. In both cases, a chi-square analysis revealed no significant effects. In sum, the youngest participants demonstrated on the practice trial that they understood the task; they showed a pattern of errors across trials that was similar to that expected by chance, and their actual responses were not biased toward any particular position.

Table 4 illustrates mean correct scores for each of the three age groups by gender and condition. For the participants in the quality condition, performance appears to have differed little between children and adults, with adults, in fact, recording the lowest mean score. In the quality and intensity condition, performance was notably better in the older children and adults. These data were analyzed using a three-way ANOVA, with condition (quality vs. quality and intensity), age group (Grade 1 vs. Grade 6 vs. adults) and gender (male vs. female) as between-participants factors.

The ANOVA confirmed the description above, in that there was a significant interaction between condition and age group [$F(2,116) = 4.54$, $MS_e = 1.21$, $p < .02$]. To test whether the age group effect was significant within each condition, we conducted two contrasts. The first, for age group in the quality condition, was not significant. The second, for age group in the quality and intensity condition, was significant [$F(2,59) = 8.77$, $MS_e = 1.21$, $p < .001$]. Tukey's HSD on these data revealed that Grade 1 participants were significantly different from both Grade 6 and adult participants, who did not signifi-

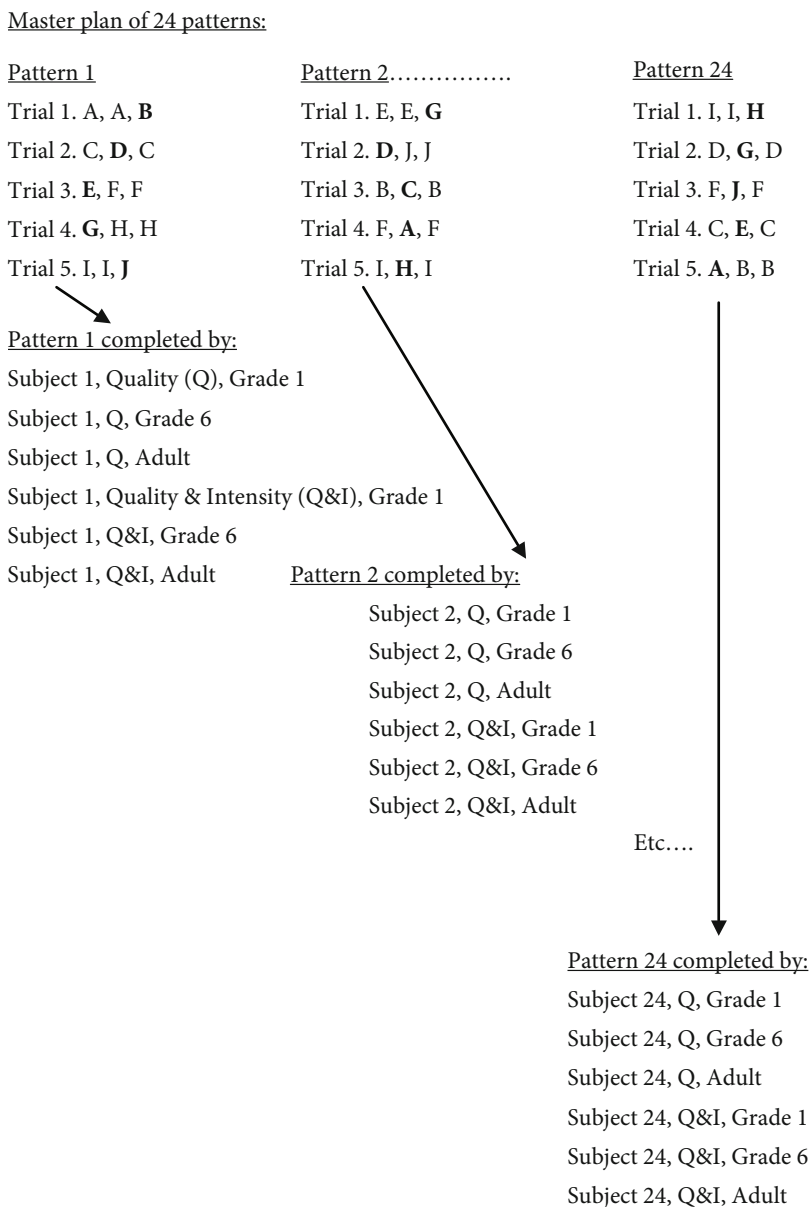


Diagram 1. Patterns of stimulus presentation. Letters refer to odorants, and bold-face type indicates the odd one out.

cantly differ from each other. Main effects were observed for condition [$F(1,116) = 12.58, MS_e = 1.21, p < .001$] and age group [$F(2,116) = 5.10, MS_e = 1.21, p < .01$], whereas gender approached significance ($F = 3.69, p = .057$). There were no other significant effects.

Three observations can be made about these results. First, the youngest children had little difficulty with the task. They performed adequately on the practice trial and did as well as the adults in the presumably more difficult quality condition. Although this should serve to reassure us that the discrimination task was not, in and of itself, a problem for the children, we attempted to resolve this experimentally by comparing (in Experiment 2) children's

and adults' performance, using the same discrimination technique, but on hard-to-label ambiguous visual stimuli (amoeboid shapes), which have been used before for this type of comparison (e.g., Lawless, 1978).

The second observation concerns the quality and intensity condition, in which differences in performance more consistent with those we expected in the quality condition were evident. It is intriguing to reflect on whether the 6-year-olds here were perhaps less sensitive to intensive cues for experiential reasons. Several studies suggest that more familiar odors are regarded as smelling stronger, especially if they can be named (e.g., Distel & Hudson, 2001). If older participants were more familiar with these

odors than were Grade 1 children, one might expect that the intensity cue would be more pronounced. Experiment 3 in this series tested the replicability of this effect.

Third, the obvious puzzle is why the adults were poorer, albeit not significantly, than the Grade 1 children in the quality condition. One possibility is that the adults (and the older children too) used a different response strategy, relative to the Grade 1 children, and that this strategy was markedly unsuccessful. Melcher and Schooler (1996) have shown that when regular wine drinkers, who are poorly equipped to label their wine experience, are asked to verbally describe it, this adversely affects their discriminative performance, when contrasted with performance in the absence of verbal description—that is, verbal overshadowing. Perhaps we could draw an analogy here to the performance of the adults and older children, in the quality condition. They may have attempted to label the stimuli, adversely affecting their discriminative ability because their label quality was poor. This is highly likely, because many of the odors we used did not have a veridical name, nor is naming easy even for familiar odors (Cain, 1979). The effect of minimizing naming with articulatory suppression was explored in Experiment 3.

EXPERIMENT 2

In Experiment 2, we set out to examine performance on an analogous visual task. More specifically, we wished to determine whether the age-related difference in the quality and intensity condition in Experiment 1 was odor or task specific.

Method

Participants. One hundred twelve participants took part in Experiment 2 (see Table 2 for age and gender details). Adult participants were recruited in the same manner as for Experiment 1. Child participants were recruited from three further primary schools in the northern region of Sydney. No participant was visually impaired, and those who normally wore correcting lenses did so.

Stimuli. In total, 10 different filled amoeboid shapes were used. Each shape came in two sizes: standard (approximately 6 × 6 cm across) and large (approximately 12 × 12 cm across), with each presented on a 21 × 21 cm paper card. The stimuli for the quality condition were drawn solely from the standard set, whereas the stimuli for the quality and intensity condition were drawn from both the standard and the large sets. Although each shape was different, all were similar, with six protruding lobes. An identifying code was written on the reverse side of each card and was hidden by a flap of paper, and this was never visible during testing. The practice trial used two four-lobed stars and an oval-like shape.

Procedure. The procedure was similar to that used in Experiment 1, with a number of minor differences. First, it was clear, after testing the youngest children, that they did not need the game format; consequently, in this and the subsequent experiment, this was dropped, and the same procedure was applied to all the age groups. Second, the participants viewed each stimulus for 3 sec and were then offered a second view of 3 sec if they wanted to see it again. The stimulus was then placed facedown in a holding rack. Third, in the quality condition, all the stimuli were of the same approximate dimensions, whereas in the quality and intensity condition, the area of the stimulus was increased, although the shape remained the same. Finally, interstimulus and intertrial intervals were identical to those in Experiment 1, and so were the formats of the trials: one compul-

Table 4
Mean Correct Scores (and Standard Deviations) by Group, for Each of the Experiments

| Group | Experiment 1 | | Experiment 2 | | Experiment 3 | |
|-----------------------|--------------|-----------|--------------|-----------|--------------|-----------|
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> |
| Quality and Intensity | | | | | | |
| Grade 1 male | 2.7 | 1.3 | 4.4 | 0.9 | 3.1 | 1.2 |
| Grade 1 female | 3.1 | 1.5 | 4.5 | 0.9 | 3.5 | 1.3 |
| Grade 1 | 2.9 | 1.4 | 4.4 | 0.9 | 3.3 | 1.2 |
| Grade 6 male | 4.4 | 0.8 | 4.6 | 0.5 | 3.7 | 1.0 |
| Grade 6 female | 4.0 | 1.0 | 4.8 | 0.4 | 4.0 | 0.9 |
| Grade 6 | 4.2 | 0.9 | 4.7 | 0.5 | 3.8 | 0.9 |
| Adult male | 3.8 | 1.3 | 4.9 | 0.4 | 3.7 | 0.9 |
| Adult female | 4.4 | 0.7 | 4.9 | 0.4 | 4.5 | 0.5 |
| Adult | 4.1 | 1.0 | 4.9 | 0.3 | 4.1 | 0.8 |
| Quality | | | | | | |
| Grade 1 male | 2.6 | 0.9 | 4.0 | 0.6 | 2.3 | 1.3 |
| Grade 1 female | 3.4 | 1.0 | 4.1 | 0.8 | 2.3 | 0.7 |
| Grade 1 | 3.0 | 1.0 | 4.1 | 0.7 | 2.3 | 1.0 |
| Grade 6 male | 3.1 | 1.1 | 3.9 | 0.7 | 2.6 | 1.4 |
| Grade 6 female | 3.5 | 1.0 | 4.2 | 1.0 | 3.7 | 1.3 |
| Grade 6 | 3.3 | 1.1 | 4.0 | 0.9 | 3.1 | 1.4 |
| Adult male | 2.6 | 1.1 | 4.6 | 0.5 | 3.5 | 1.3 |
| Adult female | 3.0 | 1.2 | 4.4 | 0.5 | 3.4 | 0.7 |
| Adult | 2.8 | 1.2 | 4.5 | 0.5 | 3.5 | 1.0 |

sory practice trial, one optional practice trial if the first was failed, and five actual trials, all using the triangle test procedure.

Results and Discussion

No participant failed the practice test, and nobody required a second practice trial. Table 4 presents the mean correct scores for each of the three age groups by condition and gender. Overall, performance here was far superior to that in Experiment 1. For those in the quality condition, performance appears to have differed little between the younger and the older children, but the adults were somewhat better. In the quality and intensity condition, performance was slightly worse in the youngest children and was similar in the older children and adults. Notably, adult performance was near ceiling in this condition. These data were analyzed using a three-way ANOVA, with condition (quality vs. quality and intensity), age group (Grade 1 vs. Grade 6 vs. adults) and gender (male vs. female) as between-participants factors.

The ANOVA revealed two main effects. First, there was an effect of condition [$F(1,100) = 11.88, MS_e = 0.49, p < .001$], with all the participants performing better on the quality and intensity condition, relative to the quality condition. Second, there was an effect of age group [$F(2,100) = 3.69, MS_e = 0.49, p < .05$], with better performance in the older participants. However, post hoc contrasts conducted separately by condition revealed no significant differences for age group in either the quality or the quality and intensity condition, even with a liberal alpha of .05. There were no other significant main effects or interactions.

We then directly tested whether the Grade 1 children would demonstrate a significantly larger difference in

performance for the visual quality and intensity condition versus the olfactory quality and intensity condition (M difference = 1.5), relative to the older children and adults (M difference = 0.6; these scores were pooled, since they did not differ significantly in either the olfactory or the visual condition). The difference in performance in the visual and olfactory conditions differed significantly by age group [$F(1,114) = 6.35$, $MS_e = 0.83$, $p < .02$]. Thus, we may conclude that poorer performance by the Grade 1 children in the quality and intensity condition in Experiment 1 is a probable consequence of the modality, not the task, but with the obvious caveat about cross-experimental comparisons.

In sum, performance was, overall, markedly better on the visual task than on the olfactory one. Most important, in the quality and intensity condition, this improvement was most marked for the youngest children, strongly suggesting an olfactory-specific discrimination deficit.

EXPERIMENT 3

We concluded from Experiment 1 that the absence of an age-related effect in the quality condition might have resulted from the use of different strategies to solve the discrimination problem. We attempted to resolve this issue here by forcing all the participants to engage in a perceptual strategy by minimizing naming with an articulatory suppression task (AST).

Minimizing naming in an olfactory task is complicated by the act of sniffing, which prevents overt speech. This clearly presents problems for the use of an AST, which involves saying out loud, repeatedly and as quickly as possible, “the, the, the . . .” (Baddeley, 1976). To address this problem, the participants were asked to start the AST as soon as they had experienced three sniffs of the odor and to keep the AST going for the whole of the interstimulus interval, to prevent verbal rehearsal. In addition, after sniffing the final odorant, the participants continued the use of the AST until *after* they had selected the odd one out. We cannot be sure that naming was absent, but we can be sure that all the participants complied with the AST, especially during the decision phase. Consequently, their discrimination choice should be based primarily on the perceptual encoding of the odors, rather than on the use of labels. A similar argument has been mounted before with the use of the AST in the decision phase of an olfactory serial position task (Reed, 2000).

Method

Participants. One hundred seventeen participants took part in Experiment 3 (see Table 2 for age and gender details). Adult participants were recruited in the same manner as in the preceding experiments. Child participants were recruited from two additional primary schools in the northern region of Sydney and from one demographically similar school in a western suburb of Sydney.

Stimuli. The stimuli were identical to those described for Experiment 1.

Procedure. The basic procedure remained the same as that in Experiment 1, apart from the following modifications. First, as with Experiment 2, the youngest children were now tested in a manner

identical to that for the older children and adults. Second, an AST was included. To limit the opportunity for verbal labeling and rehearsal, the participants were allowed only three deep puffs of the odor, after which they were told to immediately start saying “the, the, the . . .” as rapidly as possible until told to stop. This was continued for the full 20 sec of the interstimulus interval. The next odor was presented in the same way, followed again by the AST. The third odor was then presented, and after it had been smelled, the AST was again started immediately, and the participant was then instructed to point to the “odd one out.” The use of the AST during the decision phase of each triangle test was most important, since it should prevent the participant from using, or at least minimize the use of, a verbal strategy to identify the odd stimulus. During this phase, the participants were instructed to continue repeating the word *the* until told to stop. The importance of doing so was stressed prior to the practice trial, since the experimenter asked the participant to practice saying “the, the, the . . .” for 5 sec, until told to stop, and then for 20 sec, until told to stop. This procedure was strictly enforced, and even the youngest participants grasped the importance of keeping the AST going until a selection had been made and they were told to stop. Any failures to comply were carefully noted.

Results and Discussion

One adult participant’s data could not be used, since this participant failed to use the articulatory suppression task correctly, and one Grade 1 student was omitted, since the student failed both practice tests. One Grade 6 and 3 Grade 1 participants had to repeat the practice trial, and all were successful on the second occasion.

Mean performance on this experiment is presented in Table 4. The Grade 1 children performed more poorly, relative to the older participants, in both the quality and the quality and intensity conditions. It is noteworthy that mean performance in the Grade 1 children in the quality and intensity condition was not significantly different (independent $t < 1$) from their performance on the basic task—that is, in Experiment 1. Consequently, it is difficult to see how alteration in task difficulty alone can explain any observed differences between age groups in this experiment.

We again used a three-way ANOVA with condition (quality vs. quality and intensity), age group (Grade 1 vs. Grade 6 vs. adults), and gender (male vs. female) as between-participants factors. The ANOVA revealed three main effects but no significant interactions. There was a main effect of age group [$F(2,105) = 8.21$, $MS_e = 1.18$, $p < .001$], indicating, overall, superior discriminative ability in the adults and older children, relative to the Grade 1 children. A Tukey HSD confirmed this, with a significant difference between the Grade 1 children and the Grade 6 children and the adults, who did not differ. There was also a main effect of condition [$F(1,105) = 15.13$, $MS_e = 1.18$, $p < .001$], with better performance in the quality and intensity condition, relative to the quality condition, in all the participants. Finally, there was a main effect of gender [$F(1,105) = 4.03$, $MS_e = 1.18$, $p < .05$], with females outperforming males, a finding consistent with their generally superior performance on olfactory tasks. In sum, under conditions in which all the participants were forced to rely primarily on a perceptual strategy, the Grade 1 children were poorest at discriminating the odors.

GENERAL DISCUSSION

The experiments reported here tested for age-related changes in odor discrimination. The results suggest that Grade 1 children are poorer at odor discrimination than are Grade 6 and adult participants. The basis for this conclusion rests on the following. First, in Experiment 1, discrimination was poorer in Grade 1 children when the stimuli varied in quality and intensity, but when only quality varied, performance was similar across ages. The latter effect was ascribed to strategic inconsistency in the older participants' approach to the discrimination task. Second, in Experiment 2, in which a task visually analogous to that in Experiment 1 was employed, adults performed, overall, slightly better than children. Most important, however, was the finding that Grade 1 children were selectively worse on the olfactory quality and intensity task, relative to adults and older children, when Experiments 1 and 2 were directly compared. That is, the age-specific effect in Experiment 1 was unlikely to have resulted from problems with the task per se. Third, in Experiment 3, in which the participants were forced to utilize a common perceptual strategy, Grade 1 children demonstrated poorer performance, relative to older children and adults, in both the quality and the quality and intensity conditions. Before considering these results, we will examine the most important issue facing any developmental study: Could the Grade 1 children comprehend the task at hand?

There are four grounds for thinking that the Grade 1 children understood the task. First, a comparison of Experiments 1 and 2 indicated a differential deficit in performance in Grade 1 children for a similar task that differed primarily in modality, relative to older children and adults. Second, across the two olfactory experiments, the vast majority of the participants successfully completed the practice trial on their first attempt. For the visual experiment, all the participants completed the first practice trial successfully. Third, in Experiment 1, where we reported the pattern of errors, there was no indication of differences either in the position chosen for the odd one out (i.e., position bias) or in the trial on which errors were made, indicating that the youngest children were not making more errors early on (poor task comprehension) or later on (loss of attention/fatigue). This also holds true for the remaining experiments, in which these data were not presented. Fourth, performance in the quality and intensity condition for Grade 1 children actually improved, in mean terms, in Experiment 3, relative to Experiment 1, even though the former was arguably a more demanding task.

The prediction that children would evidence poorer discriminative performance than adults was based on the premise that previous experience with an odor enhances its discriminability from other odors, by virtue of the odor's being encoded into memory (Wilson & Stevenson, 2003). When the same pattern of stimulation is encountered again, this automatically recovers the matching odor memory, resulting in a more distinctive percept. Contrast this with encountering an unfamiliar odor for the first time, where its pattern of stimulation may activate many encodings with which it shares some similarity. It will then be redolent of

many smells, but none uniquely, as with a familiar odor. Since adults and older children have arguably had more opportunity to encounter the odorants used here before, when they smell the target odors, they should experience a more unique representation. In Grade 1 children, on the other hand, who are less likely to have encountered these odorants before, the percept should be redolent of many odors, producing considerable overlap in perceived quality and, therefore, greater confusability between them. Consequently, they will perform more poorly on the discrimination task than do adults and older children.

A further perspective on these results is that they reflect developmental changes in working memory. Neuroimaging studies of olfactory processing in adults reveal overlaps with brain areas believed to be involved with working memory (e.g., Savic, 2001). Developmental neuropsychological data indicate that although the basic modularity of working memory is well established in 6-year-olds, its capacity continues to expand linearly into adulthood (Gathercole, Pickering, Ambridge, & Wearing, 2004). Whether olfactory memory has a dissociable short-term store is currently in contention (e.g., Stevenson & Boakes, 2003; White, 1998), but our results can be taken to suggest developmental improvements in olfactory short-term storage.

Two experimental observations require further comment. The first concerns strategic inconsistency, which we claim adversely affected the older participants' performance in the quality condition in Experiment 1. Although we cannot rule out the possibility that this was a chance perturbation, it does resemble, as we noted before, Melcher and Schooler's (1996) finding of the detrimental effects of verbal labels on discriminative performance. In our postexperimental discussions with participants, especially adults, many reported trying to name the stimuli in Experiment 1, and all were curious afterward about what they had been smelling. We can only presume that this tacit-naming approach was unsystematic and that it would have been further compromised by the inherent difficulty in naming odors (Cain, 1979) and would have been worsened further by the fact that many of the odors here did not even have proper names. It is plausible that the results from Experiment 1 reflect verbal overshadowing.

The second observation concerns why the Grade 1 children were poorer to a similar degree in the quality and intensity condition, relative to older children and adults, in both odor experiments. First, the Grade 1 children might have benefited less from intensive cues, because of their greater unfamiliarity with the odors (e.g., Distel & Hudson, 2001). Second, and consistent with the verbal overshadowing account above, is that the Grade 1 children might have been less prone to verbally label the stimuli and, therefore, did not benefit from the potentially easy labeling scheme of *strong* or *weak*.

In conclusion, the experiments reported here establish that discriminative ability for odors is poorer in Grade 1 children than in older children and adults. We suggest that this reflects the greater experience that older participants have with odorants and/or developmental improvements in odor short-term memory.

AUTHOR NOTE

This research was funded by a grant from the Australian Research Council to the first author. The authors thank Dragoco and Quest for supplying some of the odorants used in this study and Betty Repacholi for suggestions concerning the design of Experiment 1. Correspondence concerning this article should be addressed to R. J. Stevenson, Department of Psychology, Macquarie University, Sydney, NSW 2109, Australia (e-mail: richard.stevenson@psy.mq.edu.au).

REFERENCES

- BADDELEY, A. D. (1976). *The psychology of memory*. New York: Basic Books.
- BRAND, G., & MILLOT, J. L. (2001). Sex differences in human olfaction: Between evidence and enigma. *Quarterly Journal of Experimental Psychology*, **54B**, 259-270.
- CAIN, W. S. (1979). To know with the nose: Keys to odor and identification. *Science*, **203**, 467-470.
- CAIN, W. S., STEVENS, J. C., NICKOU, C. M., GILES, A., JOHNSTON, I., & GARCIA-MEDINA, M. R. (1995). Life-span development of odor identification, learning, and olfactory sensitivity. *Perception*, **24**, 1457-1472.
- DE WIJK, R. A., & CAIN, W. S. (1994). Odor identification by name and by edibility: Life-span development and safety. *Human Factors*, **36**, 182-187.
- DISTEL, H., & HUDSON, R. (2001). Judgement of odor intensity is influenced by subjects' knowledge of the odor source. *Chemical Senses*, **26**, 247-251.
- DOTY, R. L., SHAMAN, P., APPLEBAUM, S. L., GIBERSON, R., SIKSORSKI, L., & ROSENBERG, L. (1984). Smell identification ability: Changes with age. *Science*, **226**, 1441-1443.
- GATHERCOLE, S. E., PICKERING, S. J., AMBRIDGE, B., & WEARING, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, **40**, 177-190.
- HABERLY, L. B. (2001). Parallel-distributed processing in olfactory cortex: New insights from morphological and physiological analysis of neuronal circuitry. *Chemical Senses*, **26**, 551-576.
- HASSELMO, M. E., ANDERSON, B. P., & BOWER, J. M. (1992). Cholinergic modulation of cortical associative memory function. *Journal of Neurophysiology*, **67**, 1230-1246.
- HUDSON, R. (1999). From molecule to mind: The role of experience in shaping olfactory function. *Journal of Comparative Physiology*, **185A**, 297-304.
- LASKA, M., DISTEL, H., & HUDSON, R. (1997). Trigeminal perception of odorant quality in congenitally anosmic subjects. *Chemical Senses*, **22**, 447-456.
- LAWLESS, H. T. (1978). Recognition of common odors, pictures, and simple shapes. *Perception & Psychophysics*, **24**, 493-495.
- MAARSE, H. (ED.) (1991). *Volatile compounds in foods and beverages*. New York: Dekker.
- MELCHER, J. M., & SCHOOLER, J. W. (1996). The misremembrance of wines past: Verbal and perceptual expertise differentially mediate verbal overshadowing of taste memory. *Journal of Memory & Language*, **35**, 231-245.
- MENIER, C., FORGET, R., & LAMBERT, J. (1996). Evaluation of two-point discrimination in children: Reliability, effects of passive displacement and voluntary movements. *Developmental Medicine & Child Neurology*, **38**, 523-537.
- REED, P. (2000). Serial position effects in recognition memory for odors. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 411-422.
- SAVIC, I. (2001). Processing of odorous signals in humans. *Brain Research Bulletin*, **54**, 307-312.
- SINHA, S. P. (1990). Tactual discrimination among children. *Indian Journal of Current Psychological Research*, **5**, 12-16.
- STEVENSON, R. J., & BOAKES, R. A. (2003). A mnemonic theory of odor perception. *Psychological Review*, **110**, 340-364.
- WHITE, T. L. (1998). Olfactory memory: The long and the short of it. *Chemical Senses*, **23**, 433-441.
- WILSON, D. A., & STEVENSON, R. J. (2003). The fundamental role of memory in olfactory perception. *Trends in Neurosciences*, **26**, 243-247.

(Manuscript received April 4, 2005;
revision accepted for publication May 5, 2006.)