Older and younger adults' strategies in approximate quantification

Delphine Gandini, Patrick Lemaire*, Stéphane Dufau

Laboratoire de Psychologie Cognitive, Centre National de la Recherche Scientifique, 3 Place Victor Hugo, Bâtiment 9, Case D, Marseille 13331, France

Abstract

In two experiments, participants were asked to provide a quick and rough estimate of the number of items in collections of 4–79 items. In Experiment 1 verbal strategy reports and performance on each item were collected, and in Experiment 2 performance and eye movements were collected, while young and older participants were tested in strategy-instructed conditions. Results showed that: (a) participants used six different estimation strategies, (b) overall, young and older participants used the same set of strategies, but varied in how often they used each strategy, (c) older adults' strategy repertoire was smaller than young adults' (i.e., inter-individual differences in strategy repertoire), (d) strategy use, participants' performance, and eye movements varied as a function of numerosities and configurations of items, (e) in both the age groups, each strategy was associated with distinctive performance measures and eye movement patterns. These findings show that different processes are available for approximate quantification in both young and older adults and that aging is associated with strategic variations.

1. Introduction

Everyday, we are bombarded by a lot of numerical information. Due to environmental constraints (e.g., time pressure), we often process this numerical information in an approximate manner. For instance, in a supermarket, to choose the checkout where we would wait for, we generally approximate the number of persons waiting for each checkout and choose the one with the fewer persons. Similar processes are involved when we want to determine if the available place to park our car is sufficiently large or if the quantity of food in our plate is reasonable. This ability is called approximate quantification. It consists in providing a quick and rough estimate of a magnitude. The goal of this research was to understand the approximate quantification processes and how these processes change with adults' age. Addressing these issues is important both because of the pervasiveness of this ability and because of its centrality for understanding the social and physical environment.

In this research, we have adopted a strategy perspective. Examining strategy changes during adulthood has already provided a rich and relevant amount of data to understand age-related differences in approximate quantification. Before outlining the logic of this project, we first review previous findings on age-related differences in cognitive strategies.

1.1. Age-related differences in adults' cognitive strategies

Previous works have already revealed that young and older adults use several strategies to accomplish cognitive tasks in a variety of domains (see Salthouse, 1991, for a review). For example, Dunlosky and Hertzog (1998) showed that young and older adults used three main strategies to memorize paired associates: interactive imagery to link the words, verbal repetition of the words, and making a sentence including two words. Similarly, Geary and his collaborators (Geary, Frensch, & Wiley, 1993; Geary & Wiley, 1991) found that young and older adults used several strategies to solve simple or complex arithmetic problems such as 42–49 (i.e., counting down, decomposition, columnar retrieval, or other strategies).

Previous works on cognitive strategies have also revealed that strategy distribution (i.e., mean percentage of the use of each strategy) may vary across age groups. For example, Dunlosky and Hertzog (2001) found that older adults used a sentence strategy for paired associates more often than young adults, a repetition strategy less often than young adults, and an imagery strategy equally often. In arithmetic, Geary et al. (1993) found similar results while people solved simple subtraction problems. Note that, as discussed by Salthouse (1991), age-related differences in strategy use are not always found. For example, similar strategy
distributions were observed by Cohen and Faulkner (1983) in mental rotation and linguistic verification tasks.

Concerning strategy speed and accuracy, in a wide variety of cognitive domains, young people have been found to be almost always faster and more accurate than older adults with each strategy. Different data suggest that these age-related differences depend on the resource-demanding nature of the strategies or the complexity of the problems. For instance, in arithmetic, Geary et al. (1993) found smaller age-related differences to solve simple subtraction problems with retrieval than with addition-reference strategy (i.e., retrieving $5 + 3 = 8$ to solve $8 - 5 = 7$; see also Charness & Campbell, 1988; Lemaire & Lecacheur, 2001; Siegler & Lemaire, 1997).

Finally, several studies investigating strategy adaptivity showed that older adults were less efficient at choosing the best strategy on a given problem than young individuals (Duverne & Lemaire, 2004; Duverne & Lemaire, 2005; Lemaire, Arnauer, & Lecacheur, 2004). This difficulty stems from age changes in variables influencing strategy choices.

In the specific domain of quantification, previous studies tested only small numerosities (i.e., less than 10 elements). The general finding is that counting (more than 4 items) is unaffected by normal aging, but subtitizing speed decreases with increasing age (Geary & Lin, 1998; Kotary & Hoyer, 1995; Nebes, Brady, & Reynolds, 1992; Sluisinsk, 1997; Trick, Enns, & Brodeur, 1996; Watson, Maylor, & Bruce, 2005; Watson, Maylor, & Manson, 2002). No studies compared young and older adults' approximate quantification performance. The present studies aimed at addressing these issues by directly assessing and manipulating young and older adults' estimation strategies and by manipulating stimulus features. Looking at age-related differences in approximate quantification with a strategy perspective enabled us to further examine the general issue of aging and strategic variations. Above and beyond observing age-related differences and similarities in cognitive performance, previous research in cognitive aging has asked whether aging is systematically associated with strategic variations. This study contributed to this issue by documenting the conditions under which aging is systematically associated with strategic variations.

1.2. Previous findings on approximate quantification

Previous works on approximate quantification suggest that participants use different strategies (e.g., Crites, 1992; Lemaire, Lecacheur, & Farioli, 2000; Luwel, Lemaire, & Verschaffel, 2005; Luwel, Verschaffel, Onghena, & De Corte, 2003a; Luwel, Verschaffel, Onghena, & De Corte, 2003b; Siegel, Goldsmith, & Madson, 1982). In particular, Siegel and colleagues showed that children (7–8 y.o.) and young adults used five strategies to solve verbal estimation problems (e.g., “about how many ridges are there on the handle of this whisk?”): (a) the perceptually-based strategy (e.g., do not know, guess, eyeball, range) requiring no specific knowledge stored in long-term memory, (b) the benchmark strategy consisting in applying a known standard to estimate the problem, (c) the fractional and multiple benchmark, corresponding to the use of a standard that is some multiple or fraction of the problems, and finally (d) the decomposition/recomposition strategy, consisting in subdividing into small parts so that a benchmark could be applied when no benchmark is available.

The perceptually based strategy has been extensively studied to determine whether visual features of stimulus influence individuals’ performance. In this case, estimation performance is often assessed in laboratory situations where people have to compare two collections of dots and have to quickly decide which collection has the largest number of dots (e.g., Frith & Frith, 1972; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004; Piazza, Mechelli, Price, & Butterworth, 2006). They are also tested in situations where people are shown collections of items and are asked to determine approximately how many items there are in each collection. In these studies, collections of dots generally contained more than 50 dots and were very briefly displayed (between 150 and 500 ms). These studies showed that participants tend to give larger estimates for regular patterns (e.g., dots are arranged as a circle or a rectangle) than for random patterns (e.g., Ginsburg, 1978; Ginsburg, 1980), for one large cluster of dots than for several small clusters (e.g., Frith & Frith, 1972; Ginsburg, 1991; Vos, van Oeijen, Tiboisch, & Allik, 1988) or when items occupy a more extended area on the display (e.g., Bevan, Maier, & Nelson, 1963; Krueger, 1972). Participant’s judgment of approximate numerosity is also highly influenced by the size of the item to be estimated (e.g., Ginsburg & Nicholls, 1988).

In this study, we directly assessed strategic variations in young and older adults’ approximate quantification performance. The existence of multiple strategy use in approximate quantification, how these strategies influence participants’ performance, how participants choose and execute these strategies remains question. Addressing these issues is important to understand how participants provide rough estimates of the number of items in a collection and to build theoretical models of approximate quantification that will account for the effects of stimulus features (e.g., numerosities, visual layouts) or participants’ characteristics (e.g., age, math skill).

1.3. Overview of the present study

In two experiments, young and older adults were asked to estimate numerosities of collections of 4–79 dots. In both the experiments, manipulations of the problem features included numerosities of items and configurations of dots, so that items varied in the number of dots and in how dots were displayed. In the first experiment, trial-by-trial strategy reports, accuracy and latency were collected. This enabled assessment of the participants’ strategy repertoire, the examination of strategy selection as a function of item characteristics in each age group, and the effects of participants’ age and strategy on approximate quantification performance. Despite their limits regarding reactivity and validity (see Ericsson & Simon, 1993; Kirk & Ashcraft, 2001; Robinson, 2001; Smith-Chant & LeFevre, 2003), immediate retrospective verbal protocols provide fruitful data to understand participants’ strategies. Previous studies showed that participants are able to describe their strategies in a fairly accurate and reliable way, especially in numerical processing tasks (see Lemaire & Arnaud, 2004, for a discussion).

The second experiment studied participants’ performance and eye-movements as a function of strategies, problem features, and participants’ age when strategies are manipulated. Eye-movements were collected here as previous studies showed that they reflect on-line processing and to obtain relevant information about the dynamic of the encoding processes (McCarley & Kramer, 2006; Rayner, 1998).

These two experiments enabled to test the strategy variability hypothesis as well as age-related differences in cognitive strategies and in approximate quantification skills. According to the strategy variability hypothesis (a) participants use several strategies, (b) these strategies yield different levels of speed and accuracy, and (c) both strategy use and execution vary as a function of item characteristics. Moreover, age-related differences in strategic variations predict that young and old adults differ in (a) the strategy repertoire used to accomplish approximate quantification tasks, (b) how often they use each available strategy, (c) how fast they are able to provide estimates and how accurate their estimates are with each strategy, and (d) how problem features affect strategy distributions.
2. Experiment 1

2.1. Method

2.1.1. Participants

Twenty-four young adults (14 females) and 24 older adults (16 females) took part in Experiment 1. Young adults were undergraduate students from the University of Provence (Marseille, France), who participated voluntarily or received course credit for their participation; older adults were volunteers recruited from the community who received a short book on cognitive aging written by the second author (Lemaire, 1999) as acknowledgement for their participation. All had normal or corrected-to-normal vision. None of the volunteers reported any eye disease (e.g., cataracts, macular degeneration, diabetic retinopathy).

As can be seen in Table 1, older adults had slower processing speed, higher arithmetic skills, and comparable verbal ability. Thus, the present background data are consistent with the pattern of aging typically reported in the literature.

2.1.2. Stimuli

The stimuli were 144 configurations of 8-mm black dots displayed in a visible square grid on a white background; two-thirds of which were experimental stimuli (including 15, 20 or 25 dots) and one-third of which were fillers (including 4–79 dots, excluding the collections of 15, 20 or 25 dots). The set of 144 grids was divided into three sets of 48 trials each. Each grid was made of 81 (9 × 9) units; each unit had a size of 1 × 1 cm square (participants sat 60 cm away from the screen, meaning that each grid occupied 8.6 degrees of the visual angle). The minimum distance between the dots was 4 mm, and between the dots and the grid was 2 mm. Six types of grids were tested on the basis of the size of the numerosity and configuration of black dots. The ninety-six experimental stimuli always had 15, 20, or 25 dots. In a given grid, the arrangement of dots was either random or canonical (i.e., being composed of canonical patterns; see Fig. 1). These experimental trials were intermixed with 48 fillers containing between 4 and 79 randomly displayed dots. In each block of 48 trials, 16 had a random configuration, 16 had a canonical configuration and 16 were fillers.

Canonical configurations were constructed with several constraints: (a) We controlled the mean number of the groups of (1–5) dots composing an item, independently of the pattern types: On average, displays always contained 5, 7, and 9 groups of dots for the numerosities 15, 20 and 25, respectively; (b) Displays were matched on the number of different pattern types appearing on a grid: Canonical displays were always combinations of three types of canonical patterns (e.g., a canonical display could contain two patterns of five dots, two patterns of two dots, and one single dot); (c) We controlled the percentages of each pattern type: On average, each pattern type appeared on 20% of all the canonical trials, except for the patterns of one single dot and five dots, which appeared, respectively, on 18% and 22% of all the canonical items. These distributions were identical for the three target numerosities; (d) For each target numerosity, we controlled such that each pattern appeared on the same number of grids; (e) Following Mandler and Shebo (1982), no shapes other than shown in Fig. 1a were used for the groups of dots (e.g., a triangle of three dots always had its apex at the top; four dots could not form a diamond).

2.1.3. Procedure

Participants were tested individually in one session that lasted for approximately 90 minutes. They first performed the paper-and-pencil tasks (i.e., MHVS, Deltour, 1993; Arithmetic fluency, French, Ekstrom, & Price, 1963; MMSE, Folstein, Folstein, & McHugh, 1975; Letter comparison, Digit Symbol, and Digit Digit, Wechsler, 1981), and then the approximate quantification task. Stimuli were presented using a SONY G-FX201 PC computer with a 14-inch computer screen. The experiment was controlled using the E-Prime software. The program generated the displays and recorded latencies to the nearest millisecond. The display resolution was 640 × 480 pixels. Each trial was preceded by a blank screen (1000 ms) and a fixation point (”*”) in the center of the screen for 750 ms. The grid was then displayed in the center of the screen during 6 s. Participants were instructed to provide their estimates within 6 s. Six seconds was chosen to encourage approximate quantification, as pilot testing with three university students indicated that 6 s was the average time minus 30% to enumerate exactly between 10 and 79 dots. The grid remained on the screen until the participant responded or until 6 s elapsed. If an estimate was not provided after 6 s, the grid disappeared and four blue

![Random 20](image1.png)  ![Canonical 20](image2.png)

Fig. 1. (a) The five canonical patterns used in Experiments 1 and 2. (b) Examples of canonical and random items with 20 dots.

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Young adults</th>
<th>Older adults</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experiment 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>24</td>
<td>24</td>
<td>–</td>
</tr>
<tr>
<td>Age</td>
<td>24 (3.7)</td>
<td>72 (7.2)</td>
<td>–</td>
</tr>
<tr>
<td>Years of education</td>
<td>15.6 (2.4)</td>
<td>11.5 (2.8)</td>
<td>29.41*</td>
</tr>
<tr>
<td>Health</td>
<td>5.9 (0.9)</td>
<td>5.3 (1)</td>
<td>4.80</td>
</tr>
<tr>
<td>MHVS</td>
<td>26.2 (4)</td>
<td>27.8 (3.7)</td>
<td>2.04</td>
</tr>
<tr>
<td>Arithmetic fluency</td>
<td>75.8 (23.2)</td>
<td>103.9 (35)</td>
<td>10.76</td>
</tr>
<tr>
<td>MMSE</td>
<td>–</td>
<td>29.3</td>
<td>–</td>
</tr>
<tr>
<td>Letter comparison</td>
<td>32.1 (5)</td>
<td>22 (5.1)</td>
<td>47.74*</td>
</tr>
<tr>
<td>Digit symbol</td>
<td>65.2 (103)</td>
<td>50 (11.5)</td>
<td>17.77</td>
</tr>
<tr>
<td>Digit digit</td>
<td>64.4 (10.8)</td>
<td>42.5 (12.9)</td>
<td>45.18</td>
</tr>
<tr>
<td><strong>Experiment 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>15</td>
<td>15</td>
<td>–</td>
</tr>
<tr>
<td>Age</td>
<td>26.8 (2.6)</td>
<td>69.8 (7.7)</td>
<td>–</td>
</tr>
<tr>
<td>Years of education</td>
<td>16.8 (3.1)</td>
<td>11 (2.5)</td>
<td>64.32*</td>
</tr>
<tr>
<td>Health</td>
<td>5.7 (0.8)</td>
<td>5.2 (1.1)</td>
<td>0.65</td>
</tr>
<tr>
<td>MHVS</td>
<td>25.2 (3.6)</td>
<td>25.3 (5.1)</td>
<td>0.04</td>
</tr>
<tr>
<td>Arithmetic fluency</td>
<td>70 (22.6)</td>
<td>81.1 (26.5)</td>
<td>0.51</td>
</tr>
<tr>
<td>MMSE</td>
<td>–</td>
<td>29.3</td>
<td>–</td>
</tr>
<tr>
<td>Letter comparison</td>
<td>28.3 (6)</td>
<td>20.7 (5.7)</td>
<td>10.16*</td>
</tr>
<tr>
<td>Digit symbol</td>
<td>56.7 (9.4)</td>
<td>38 (10.9)</td>
<td>17.82</td>
</tr>
<tr>
<td>Digit digit</td>
<td>59.1 (8.2)</td>
<td>45.2 (10.1)</td>
<td>11.88*</td>
</tr>
</tbody>
</table>

Note. Dfs = 1,46 and 1,28 for Expts 1 and 2, respectively. p < .05; Health was self-rated by individuals on a scale from 1 (poor) to 7 (excellent). Arithmetic fluency = Arithmetic fluency corresponds to the number of correctly solved problems (e.g., 69 + 93 + 85 or 98 – 75) within two minutes. MHVS = Mill-Hill Vocabulary Score which is the number of items (max = 33) for which participants had to identify which of 6 words had the same meaning as a target word. None of older adults obtained an MMSE score lower than 27; therefore, none were excluded.
croses were displayed. After each grid, the experimenter recorded the participant's response and verbal protocol (i.e., participants were asked “how did you estimate the number of dots?”). Verbal protocols were fully written down by the experimenter for later coding. Instructions mentioned no particular strategies. Grids did not remain on the screen during verbal protocols. A timer was started when the grid appeared on the screen and ended when the experimenter pressed the space bar of the computer keyboard, which happened as soon as possible after the participants provided their response orally. The order of presentation of grids was randomized for each participant.

General instructions described the approximate quantification task (i.e., “providing approximate number of dots displayed on the square grid without counting the exact number”), and the participants were asked to respond as quickly as possible, but without sacrificing accuracy. Each participant was permitted a 5–10-min rest between blocks. Before the experiment started in earnest, participants received 12 practice (similar to experimental) trials to familiarize themselves with the apparatus, procedure, and task.

2.2. Results

Results are reported in three main parts. The first looks at which strategies participants used and strategy frequencies; the second examines age-related differences in strategy selection. Finally, we analyze strategy performance. In all the results, unless otherwise noted, differences are significant at least p < .05.

2.2.1. Strategy use

The analyses of strategy use aimed at answering the following questions: (a) What strategy did the participants use to provide estimates? (b) Did the individuals use a single or several strategies?, and (c) Which strategies were used most often? Two raters who independently classified strategy use agreed on 97% of them. Analyses of individual protocols revealed six strategies (see examples of verbalizations in Table 2): (a) **Anchoring**: Participants enumerated several dots (via counting), visually estimated the remaining dots based on the first enumeration, and then added the enumerated result and the estimated result. (b) **Benchmark**: Participants visually scanned the stimulus, retrieved a numerical representation in long-term memory (LTM), compared the difference between the encoded representation and the retrieved representation, and then adjusted their answer on the basis of this difference. (c) **Decomposition/recomposition**: Participants spotted one group of few dots, up to about four or five items, estimated the number of analogous groups, and then multiplied the number of items primarily subitized by the estimated number of groups. (d) **Approximate counting**: Participants perceived several groups of different sizes and approximately added these groups to produce estimates. (e) **Exact counting**: Participants counted all the dots displayed in the grids by systematically adding all the items (by ones, twos, or threes). (f) **Others**: These strategies included verbal reports that did not correspond to the previous categories.

Overall percentages of the strategy use revealed that young and older adults did not prefer the same strategies, $F(4,184) = 5.21$, $MSe = 469.6$. Young participants’ most favourite strategies were approximate counting (which participants used on 31% of all trials on average) and benchmark strategy (28%), followed by anchoring (17%), decomposition/recomposition (16%), exact counting (7%), and others (1%). Older participants preferred benchmark (33%) and exact counting (29%), followed by anchoring (17%), approximate counting (17%), and finally decomposition/recomposition (5%); they never used others.

2.2.2. Strategy selection

A second series of analyses aimed at examining the role of stimulus characteristics in young and older adults’ strategy choices. Mean percentages of the use of each strategy was analyzed separately, with ANOVAs involving 2(Age: young, older adults) x 2(Configuration: random, canonical) x 3(Numerosity: 15, 20, 25) designs, with age as the only between-subjects factor. As the other strategies were used too rarely, we focused on the other five strategies (see summary of significant effects in Table 3 and means shown in Fig. 2).

2.2.2.1. Benchmark strategy. Participants used the benchmark strategy more often while quantifying random configuration than canonical configurations (47% vs. 14%) and used it more often with increasing numerosities. These two variables interacted, showing larger random-canonical differences with increasing numerosities. These differences were 26%, 35%, and 37% for 15, 20, and 25 dots, respectively (all pairwise comparisons were significant, $F_{s} > 44.98$).

2.2.2.2. Anchoring strategy. ANOVA revealed a main effect of numerosity, which interacted with the configuration. On random configurations, participants used anchoring strategy less often on displays of 15 dots (14%) than on displays of 20 (21%) and 25 dots (21%), all pairwise comparisons were significant, $F_{s} > 3.97$. On canonical configurations, participants increased their use of the anchoring strategy with increasing numerosities: They used it more often on numerosity 25 (27%) than on numerosity 20 (13%).
2.2.2.3. Decomposition/Recomposition strategy. Young adults (16%) used decomposition/recomposition strategy more often than older adults (5%), and participants varied their strategy use with the size of collections: They used it less often to estimate 15 dots (8%) than 20 dots (12%) or 25 dots (11%, F_s > 7.45). It was used equally often to estimate 20 and 25 dots, F < 1.20.

2.2.2.4. Approximate counting strategy. Young adults used approximate counting strategy (31%) more often than older adults (17%). Moreover, participants used it more often on canonical configurations (39%) than on random configurations (8%). These two variables interacted, showing larger age differences on canonical configurations than on random configurations (26% vs. 1%, for canonical and random displays, respectively). Finally, participants used the approximate counting strategy less often with increasing numerosities (30%, 25% and 17%, for 15, 20, and 25 dots, respectively). This factor interacted with configuration, showing larger canonical-random differences with 15 or 20 dots (33%) than with 25 dots (25%).

2.2.2.5. Exact counting strategy. Older adults used the exact counting strategy (29%) more often than young adults (7%). Participants used it more often on canonical configurations (21%) than on random configurations (13%). These two variables interacted, showing larger age differences on canonical configurations (29%) than on random configurations (14%). Finally, participants used the exact counting strategy less often with increasing numerosities (27%, 17%, and 8%, for 15, 20, and 25 dots, respectively). This factor interacted with age, showing larger age differences with decreasing numerosities. These differences were 29%, 22%, and 14%, for 15, 20, and 25 dots, respectively.

2.2.3. Strategy execution

Mean solution times\(^3\) and percentages of deviation were analyzed with ANOVAs with a 2(Age: young, older adults) × 5(Strategy: anchoring, benchmark, approximate counting, exact counting, decomposition/recomposition), with age as the only between-subjects factor. Summaries of significant effects are presented in Table 4 and means are shown in Fig. 3.

\(^3\) In both experiments, analyses of solution times were also conducted on standardized solution times to control for artifactual interactions involving the age factor due to general slowing. As no interactions involving the age factor were artifactual in the present experiments, we report statistics for raw estimation times.

---

Table 3
Summary of significant effects on percentage of strategy use (Experiment 1)

<table>
<thead>
<tr>
<th>Effects</th>
<th>df</th>
<th>MSe</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark</td>
<td>1.46</td>
<td>1042.65</td>
<td>74.67</td>
</tr>
<tr>
<td>Numerosity</td>
<td>2.92</td>
<td>91.77</td>
<td>64.05</td>
</tr>
<tr>
<td>Configuration × numerosity</td>
<td>2.92</td>
<td>68.42</td>
<td>10.37</td>
</tr>
<tr>
<td>Anchoring</td>
<td>1.46</td>
<td>2051.77</td>
<td>6.76</td>
</tr>
<tr>
<td>Numerosity</td>
<td>1.46</td>
<td>771.93</td>
<td>88.63</td>
</tr>
<tr>
<td>Configuration × numerosity</td>
<td>1.46</td>
<td>101.67</td>
<td>35.93</td>
</tr>
<tr>
<td>Decomposition/recomposition</td>
<td>1.46</td>
<td>771.93</td>
<td>14.21</td>
</tr>
<tr>
<td>Age × Configuration</td>
<td>1.46</td>
<td>129.58</td>
<td>3.81</td>
</tr>
<tr>
<td>Age × Numerosity</td>
<td>2.92</td>
<td>54.95</td>
<td>5.44</td>
</tr>
</tbody>
</table>

Table 4
Summary of significant effects on latencies and percentage of deviation (Experiment 1)

<table>
<thead>
<tr>
<th>Effects</th>
<th>df</th>
<th>MSe</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latencies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>1.32</td>
<td>133.457</td>
<td>5.81</td>
</tr>
<tr>
<td>Strategy</td>
<td>4.128</td>
<td>214.348</td>
<td>10.82</td>
</tr>
<tr>
<td>Age × Strategy</td>
<td>4.128</td>
<td>214.348</td>
<td>2.54</td>
</tr>
<tr>
<td>Percentages of deviation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strategy</td>
<td>4.128</td>
<td>26</td>
<td>27.47</td>
</tr>
</tbody>
</table>

---

F(1,46) = 45.69, MSe = 109.89, and more often on numerosity 20 than on numerosity 15 (7%), F(1,46) = 14.71, MSe = 57.36.
The corresponding analysis has been conducted on the mean percentages of deviation. Following previous works on estimation (LeFevre, Greenham, & Waheed, 1993), the estimation accuracy was calculated, for each participant and each collection, as following 
\[
\frac{\text{Absolute} \ (\text{Estimate} – \text{Actual quantity})}{\text{Actual quantity}} \times 100.
\]
For instance, suppose a participant gave 20 as an estimate for 25 dots, that participant would be 20% \(\left(\frac{(20 – 25)}{25}\right) \times 100\) away from the exact numerosity4.

This analysis only revealed a significant main effect of strategy. Both young and older adults produced better estimates with exact (5.9%) and approximate counting (5.9%, \(F < 1\)) than with the other three strategies, all pairwise comparisons were significant, \(F_s > 10.03\). Next, participants were more accurate while using anchoring (11.8%) and decomposition/recomposition (10.7%, \(F < 1.33\)). Finally, benchmark was the less accurate strategy (16.7%, all \(F_s > 21.14\)).

### 2.3. Discussion

The results of Experiment 1 showed that participants used several strategies to accomplish the approximate quantification task: benchmark, anchoring, decomposition/recomposition, approximate counting, exact counting, and other strategies. Some strategies were used more often (e.g., benchmark or approximate counting) than others (e.g., decomposition/recomposition or exact counting). Strategies varied with problem characteristics, and most interestingly with aging.

The configuration of dots affected the use of all the strategies, except decomposition/recomposition, and numerosity influenced the use of all the strategies. The configuration \(\times\) numerosity interaction was significant when the participants used benchmark, anchoring, and approximate counting strategies, but not when they used decomposition/recomposition or exact counting. These different strategy distributions as a function of problem type show that the participants selected strategy on a problem-by-problem basis and adjusted their strategy use to both configurations and numerosities. Effects of configuration and numerosity on strategy use validate the distinctions among strategies as the use of different strategies was differently influenced by problem features.

Moreover, percentages of exact estimates were larger when participants used exact counting strategy than when they used non-exact counting strategies (see Table 5). Also, these percentages decreased with increasing numerosities while people used exact counting strategy. This pattern was most obvious in younger adults. Older adults provided exact responses on fewer trials when they used exact counting strategy than approximate counting strategy. Although mean percent use of each type of strategy differed in each age group, this suggests that older adults might have had more difficulties in discriminating among the strategies they were using. In particular, they might have used exact counting on some trials, and claim to have used approximate counting. Also, they might have used approximate counting on some trials while saying that they used exact counting. Note though that percentages of exact answers decreased with increasing numerosities in both young and older adults. This is consistent with the idea that, even if older adults were not as accurate as young adults in their

---

4 The signed values of the mean percentages of deviation (i.e., positive/negative range) were analyzed too. The ANOVA revealed that young and older adults tended to underestimate quantity but there was no difference between the two age groups, (young: -5.7% vs. older: -6%, \(F < 1\)).
verbatim protocols on each trial, they did not provide random verbal protocols.

Although young and older adults' strategy repertoire generally included the same set of strategies, young and older adults differed in how often they used each available strategy. In particular, older adults used the decomposition/recomposition and approximate counting strategies less often than young adults but used exact counting more often. These effects possibly occurred because these strategies involve visuo-spatial processes that decline with age. Recall that in these two strategies, participants were separating visual stimulus into different sub-collections of 3–5 dots (that they could subitize), and keeping track in visuo-spatial working memory the distinction between already enumerated dots and still to-be-enumerated dots. Moreover, especially for the decomposition/recomposition strategy, participants had to mentally transfer the first enumerated group of dots on the remaining dots to determine the approximate number of similar groups. If these processes are less efficient in older adults than in young adults, then older adults would use them less often than young adults. The examination of strategy performance partially supports this interpretation. Indeed, older adults took longer than young adults at executing the decomposition/recomposition strategy (+794 ms), but were equally fast with the approximate counting strategy (+193 ms).

In parallel, older adults used the exact counting strategy more often than young adults. Older adults' greater use of exact counting may result from older adults' willingness to be most accurate, even if the task was not to provide exact numerosities. This is possible as previous studies showed that older adults were as good as young adults at exact counting (e.g., Basak & Verhaeghen, 2003; Geary & Lin, 1998; Sliwinski, 1997; Trick et al., 1996; Watson et al., 2002, 2005).

One could argue that age differences found here in strategy distribution stem from the stimulus duration (i.e., 6 s). It is possible that this was insufficient for older people and modified their strategy use compared to younger adults. Note that older adults did not need 6 s to make their approximations as their mean response time was equal to 4.2 s. Also, it could be argued that most of our findings stem from the distribution of our stimuli. As two-thirds of the stimuli had numerosities of 15, 20, or 25 dots, participants may have estimated the number of dots based on the likelihood of these numerosities, shifting from an approximate quantification task to a categorization task during the course of the experiment. We tested this possibility by comparing strategy distributions in the first and second halves of the experiment. The shifting from an estimation task to a simple categorization task predicts that participants change their strategy use during the experiment. They should use the benchmark strategy much more often and exact counting less often during the second half of the experiment. Actually, this was not the case. Participants used the benchmark strategy on 30% and 32%, respectively. Actually, this was not the case. Participants used the decomposition/recomposition and approximate counting strategies less often than young adults but used exact counting more often. These effects possibly occurred because these strategies involve visuo-spatial processes that decline with age. Recall that in these two strategies, participants were separating visual stimulus into different sub-collections of 3–5 dots (that they could subitize), and keeping track in visuo-spatial working memory the distinction between already enumerated dots and still to-be-enumerated dots. Moreover, especially for the decomposition/recomposition strategy, participants had to mentally transfer the first enumerated group of dots on the remaining dots to determine the approximate number of similar groups. If these processes are less efficient in older adults than in young adults, then older adults would use them less often than young adults. The examination of strategy performance partially supports this interpretation. Indeed, older adults took longer than young adults at executing the decomposition/recomposition strategy (+794 ms), but were equally fast with the approximate counting strategy (+193 ms).

3. Experiment 2

Experiment 1 made it difficult to evaluate the impact of problem features and participants' age on performance as it was contaminated by strategy selection. This second experiment investigated strategy execution, uncontaminated by strategy selection. We manipulated strategies via task instructions and controlled that all participants use the same strategies on the same problems and with equal frequencies. Two specific strategies were tested, the anchoring and benchmark strategies, as Experiment 1 revealed that both young and older adults used them spontaneously and equally often. A further change in Experiment 2 was that eye movements were recorded while participants completed the task.

Using this tool to investigate perceptual aspects of cognitive tasks has already provided fruitful and relevant data to better understand the way individuals accomplish a given task or to explain individual differences (Charness, Reingold, Pomplun, & Stampe, 2001; Grant & Spivey, 2003; Green, Lemaire, & Dufau, 2007; Watson et al., 2005). In particular, in cognitive aging, eye-movement data helped to clarify the impact of age on enumeration and visual search tasks (Kramer et al., 2006; Watson et al., 2005). For instance, the possibility to make voluntary eye movements in enumeration tasks, but not in search tasks, produces age equivalence for enumeration (Watson et al., 2005). In their study, Watson et al. (2005) have found an age-related decrement in search rates for single targets both for RT and fixation frequency, but found no deficit in enumeration rates either with or without distractors even though serial enumeration rates were much slower than single target search rates. Their results suggested that the rate-limiting factor might be the speed with which eye movements could be programmed and executed. Indeed, previous work has shown that enumeration beyond about four items becomes difficult or less efficient when eye movements are prevented (e.g., Atkinson, Campbell, & Francis, 1976), whereas serial visual search may become more efficient (e.g., Klein & Farrell, 1989; Zelinsky & Sheinberg, 1997). Thus, eye movements may be necessary for accurate enumeration performance, but not for visual search. According to Watson et al., an age-related reduction in search efficiency may reflect an age-related reduction in attentional capacity. In contrast, enumeration rates are age equivalent because enumeration requires eye movements, and these eye movements are sufficiently slow (but age equivalent) to reduce or effectively mask the impact of any reduced attentional resources caused by old age (Watson et al., 2005).

Finally, it is important to note that different aspects of eye movements, such as fixation control (Kosnik, Kline, Fikre, & Sekuler, 1987) or detailed features of saccadic movements (Abrams, Pratt, & Chasteen, 1998), are spared by aging. So measuring eye movements appears as a promising approach to supplement reaction time, accuracy, and verbal report measures to investigate problem-solving processes and cognitive aging.
The hypothesis that approximate quantification performance is influenced by strategies predicts that the participants should be faster, though less accurate, with the benchmark strategy than with the anchoring strategy. Also, the effects of item characteristics were expected to be different when participants use each strategy. These outcomes should occur because, in the benchmark strategy, approximate numerosity is found through quickly scanning the stimulus and retrieving a corresponding approximate numerosity in LTM. Indeed, the benchmark strategy involved processes similar to those that have been described by the numerosity-accumulator model (Dehaene & Changeux, 1993; Gallistel & Gelman, 1992; Meck & Church, 1983). This system is composed of a pacemaker, an accumulator, a working memory buffer, reference memory, and a comparator. The pacemaker produces pulses at a constant rate that can be gated into an accumulator. When a response by the organism is rewarded, the accumulator value is transferred from working memory to be stored in reference memory. A comparator process enables the organism to compare the current working memory content to the reference memory content. The outputs of the accumulator are magnitudes that represent numerosity. So, as these encoding and retrieval processes are quickly executed, estimates should be found quickly and response times should not vary with item characteristics (or should vary little). In contrast, the anchoring strategy consists in enumerating groups of dots, estimating the remaining dots, and then adding these two results. As counting is more difficult on random configurations than on canonical configurations and is influenced by numerosities, estimation latency is expected to increase with numerosities and to be longer on random configurations than on canonical configurations while using the anchoring strategy. Also, the random-canonical difference should increase with increasing numerosities while using the anchoring strategy.

As all young and older adults used each strategy on all problems, the present experiments enabled us to test aging effects for both strategies and all problem types without running the risk of confounding aging and strategy selection effects. As older adults are known to be impaired on visuo-spatial processes, and not on enumerating processes, we tested Age × Strategy interaction that should show differences between young adults and older adults for the benchmark strategy only. Indeed, the age-impaired visuo-spatial processes are crucial for the benchmark strategy, and age-invariant enumeration processes are critical for the anchoring strategy. Moreover, we tested the Age × Numerosity × Configuration interaction expecting greater increase of random-canonical differences with increasing numerosities in older adults than in young adults while using the benchmark strategy. This outcome would result from older adults’ greater difficulties with encoding large numerosities made of random patterns configurations.

3.1. Method

3.1.1. Participants

Participants were 30 individuals who did not participate in Experiment 1: 15 young adults (7 females) and 15 older adults (9 females). Young and older adults were volunteers recruited from the community; older adults received a short book on cognitive aging (Lemaire, 1999) as acknowledgement for their participation (see participants’ characteristics in Table 1). As in Experiment 1, all had normal or corrected-to-normal vision. None of the volunteers reported any specific eye disease (e.g., cataracts, macular degeneration, diabetic retinopathy).

3.1.2. Stimuli

The stimuli were 108 square grids similar to those used in Experiment 1, except that they consisted of 10 × 10 little square units (as in the first experiment, participants sat 60 cm away from the screen, meaning that each grid occupied 9.5° of the visual angle). As in Experiment 1, two thirds of stimuli were experimental stimuli (including 15, 20, or 25 dots) and one third were fillers (including 11–79 dots, excluding collections of 15, 20, or 25 dots).

3.1.3. Procedure

Procedure differed from that used in Experiment 1 on three points. First, participants were tested only on two strategy-instructed conditions, benchmark strategy and anchoring strategy conditions. For the benchmark strategy, participants received the following specific instruction, “To do this estimation task, you can use one and only one strategy. We call it perceptual estimation. It consists in visually scanning the patterns of dots on the computer screen, finding an approximate numerosity in your memory, and possibly adding or subtracting a little bit from this retrieved numerosity to give your estimate. For example, if after seeing the whole collection of dots you feel that there are 30 dots or a little bit more, you can do 30 + 2 = 32 dots”. The specific instruction for the anchoring strategy was the following, “To do this estimation task, you can use one and only one strategy. We call it anchoring. It consists in first counting approximately part of the presented dots and then globally estimating the other dots as a function of the result of your first approximate counting. For example, if you approximately count a first group of 7 dots, and then estimate that twice as many dots remain, there are 7 + 14 = approximately 21 dots.” Second, for the anchoring strategy condition, participants had to count out loud so as to control that they used the required strategy. Third, eye-movements were recorded during the experiment with an iView® × Remote Eyetracking Device (Senso-Motoric Instruments). A camera located next to the participant’s visual display unit, on the participant’s right, tracked the location of the right pupil. Simultaneous tracking of the corneal reflex allowed the system to compensate for minor head movements. The iView system we used had a sampling rate of 50 Hz and a gaze position accuracy of 0.5°. The participants were asked not to make too much head or body movement, but no device restricted them from moving. Calibration was performed by requesting the participants to view nine crosses on the screen. Recalibration was performed between each block if necessary. Eye position was sampled every 20 ms and analyzed offline using customized software.

Within each age group, 8 participants saw the benchmark strategy condition first and the anchoring strategy second, and 7 participants saw the reverse order. Participants saw the same displays twice, once under each strategy condition. To avoid specific stimulus recognition, we created two versions of the same canonical and random collections of dots, one being the mirror-reversed of the other (i.e., the original stimuli were reversed with a vertical symmetry); the presentation of the two versions was counterbalanced across strategy conditions. Participants practiced each strategy on 12 problems. When the experiment started in earnest, all young and older adults were comfortable with distinguishing among and executing each strategy. Participants were tested individually in one single session that lasted for approximately 60 minutes.

3.2. Results

3.2.1. Performance

The first series of analyses aimed at examining the role of participants’ age and stimulus properties on strategy performance when this performance is not contaminated by strategy choices. Mean solution times and percentages of deviation for each strategy, in each group, are presented in Fig. 4. ANOVAs on solution times and mean percentages of deviation involved 2(Age: young, older adults) × 2(strategy: benchmark, anchoring) × 3(configuration: random, canonical) × 3(numerosity: 15, 20, 25) designs, with
age as the only between-subjects factor (see Table 6 for summary of significant effects).

### 3.2.1.1. Latencies

ANOVA revealed main effects of strategy, configuration, and numerosity. Participants were faster with benchmark (2319 ms) than with anchoring (3923 ms), on canonical displays (3031 ms) than on random displays (3211 ms), and took more time to estimate arrays of dots as numerosity increased (15: 2928 ms; 20: 3165 ms; 25: 3269 ms, all pairwise Fs > 12.35).

Moreover, the following interactions were significant: Age × Strategy, Strategy × Configuration, Strategy × Numerosity, and Configuration × Numerosity. These interactions showed that: (a) older adults took longer than young adults to execute benchmark strategy (+572 ms), \( F(1,28) = 5.87, MSe = 2507855 \), but were equally fast to execute anchoring strategy, \( F < 1.12 \), (b) configura-

tion only influenced the execution of anchoring since participants took more time to execute this strategy on random displays (4101 ms) than on canonical displays (3744 ms), \( F(1,28) = 46.41, MSe = 124169 \), (c) numerosity had a larger effect on anchoring than on benchmark, resulting in increasing latencies only for anchoring with increasing numerosities (i.e., 2228 ms vs. 3628 ms, 2369 ms vs. 3961 ms, and 2359 ms vs. 4179 ms, for 15, 20, and 25, respectively), and (d) numerosity had a larger effect on canonical displays than on random displays, resulting in decreasing random-canonical differences in latencies with increasing numerosities (i.e., these differences were 232 ms, 200 ms, and 107 ms, for collections of 15, 20, and 25, respectively).

### 3.2.1.2. Percentages of deviation

The corresponding analyses on mean percentages of deviation revealed main effects of age, strategy, configuration, and numerosity. These effects showed that (a) young adults (12.2%) were more accurate than older adults (17.2%), (b) participants' estimates were more precise when they used the anchoring strategy (12%) than the benchmark strategy (17.5%), (c) participants approximated more precisely canonical displays (12.8%) than random displays (16.7%), and (d) both young and older adults were more accurate when they approximated arrays of 15 (13.4%) and 20 dots (14.3%), \( F < 1.51 \), than arrays of 25 dots (16.5%, respectively \( F(1,28) = 21.50, MSe = 26.86 \), and \( F(1,28) = 9.52, MSe = 29.93 \)). No other effects were significant on either latencies or percentages of deviation.

### 3.2.2. Eye-movements data

Analyses of eye movements had the following two goals. First, patterns of eye movements were used as converging evidence for participants' using benchmark and anchoring strategies. Second, although no previous eye movement data have been reported in estimation task and no specific a priori predictions could be tested,
patterns of eye-movements were expected to inform about cognitive processes within each strategy more precisely than latencies or percent deviations. As no previous studies on approximate quantification collected such data, it is hard to make specific and precise predictions. Nevertheless, if different strategies are associated with different patterns of eye movements, then mean number of fixations and amplitudes of saccades should not be the same for each strategy. Moreover, interactions of age with strategy and/or problem types on eye-movement data were tested to determine if young and older adults execute approximate quantification strategies in comparable ways. Indeed, even if both the age groups were instructed to use both the strategies on all the problems, strategy differences in each age group may stem from the way each strategy is executed by young and older adults.

To achieve these ends, we ran two sets of analyses. The first examined effects of participants’ age and properties of stimuli on eye movements during the trial. The second examined effects of participants’s age and strategy on eye movements across the trial (the first and second bins of 500 ms). Firstly, we ran 2 × 2 × 3 (Age × Strategy × Configuration) ANOVAs on mean number and on mean amplitudes of saccades (see Table 7 for summary of significant effects). Illustrative patterns of eye movements are displayed in Fig. 5.

Participants’ point-of-regard (POR) was recorded every 20 ms by the IView-X® system. To classify POR samples according to fixations, we used the Dispersion-Threshold Identification (I-DT) algorithm based on spatial dispersion of consecutive POR (for the detailed algorithm, see Salvucci & Anderson, 2001; Salvucci & Goldberg, 2000).6

Mean number of fixations during the trial (see Fig. 6). ANOVAs on mean number of fixations showed the same effects as ANOVA on solution times. Correlations between solution times and mean number of fixations were r = .82 and .89 for the benchmark and anchoring strategies, respectively.

Mean amplitudes of saccades during the trial (see Fig. 7). Main effects of age, configuration, and numerosity were significant. Older adults (3.8°) made larger saccades than young adults (3.1°). Both the age groups made smaller saccades while quantifying random configurations (3.5°) than canonical configurations (3.9°), and when they estimated arrays of 20 (3.7°) or 25 dots (3.6°, F < 1) than arrays of 15 dots (3.9°, both Fs > 7.94). The Age × Strategy, Strategy × Numerosity and Strategy × Configuration × Numerosity interactions were also significant, showing that (a) young participants made smaller saccades while executing benchmark than anchoring, whereas older participants made saccades of equal amplitudes while quantifying with either strategy, (b) problem features had larger effects on the execution of benchmark than on the execution of anchoring, resulting in an increasing canonical-random difference with larger numerosities for the former (these differences were 0.8°, 0.4°, and 0.1°, for 15, 20, and 25 dots respectively, Fs > 9.83 for 15 and 20 dots, and F < 1, for 25 dots), and constant canonical-random differences with increasing numerosities for the latter (these differences were 0.5°, 0.6°, and 0.4°, for 15, 20, and 25 dots, respectively, all pairwise Fs > 6.20).

Mean amplitudes of saccades across the trial (see Fig. 8). A 2 × 2 × 2 (Age × Time × Strategy) ANOVA was run during the first half of each trial with the additional factor of time bin (first and second bins of 500 ms) on the mean amplitude of saccades. This latter analysis was expected to be most fruitful regarding early processes distinguishing the two strategies. We found main effects of age, time, and strategy. During this 1000-ms initial trial duration, older adults (4.2°) made larger saccades than young adults (3.7°). Both the age groups made smaller saccades during the first 500-ms bin (3.1°) than during the second time bin (4.1°), and when they used the benchmark strategy (3.3°) than when they used the anchoring strategy (3.8°). The Time × Strategy interaction was also significant, showing a larger benchmark-anchoring difference during the first than during the second time bin (these differences were 0.8° and 0.2°, both Fs > 7.20).

3.4. Discussion

In this second experiment, we focused our attention on the two main strategies used by both young and older adults in Experiment 1, the benchmark strategy and the anchoring strategy, and we controlled the strategy choices to examine the participants’ performance. Data showed the effects of problem features (e.g., participants were faster on canonical displays than on random displays), strategies (e.g., participants were faster with benchmark than with anchoring), age, and interactions among these factors on both approximate quantification performance and eye movements. These effects enable a better understanding of the previous findings from both Experiment 1 and other studies. They also help further specify cognitive processes in approximate quantification as well as loci of age-related differences.

Participants’ performance in Experiment 2 replicated most results from both Experiment 1 and previous studies. Participants performed better with canonical collections of dots than with random collections of dots, on small numerosities than on larger numerosities, and the random–canonical difference increased with increasing numerosities. It is possible that physical factors characterizing stimuli (e.g., distances between groups of dots, distance between items in a group, maximum number of dots that can appear in a group) contributed to differences in performance between random and canonical configurations. Such effects have been found in previous studies (e.g., van Oeffelen & Vos, 1982) testing briefly presented large numerosities. Such parameters were perfectly controlled in our canonical pattern, but not in our random collections. Future studies might/should replicate our canonical–random differences, while holding constant not only strategies (like here), but also physical features of random configurations.

The benchmark strategy yielded faster, but less accurate performance than the anchoring strategy, and strategy differences were

---

**Table 7**

Summary of significant effects on mean number of fixation and mean amplitudes of saccade (Experiment 2)

<table>
<thead>
<tr>
<th>Effects</th>
<th>df</th>
<th>MSe</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean number of fixation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strategy</td>
<td>1.28</td>
<td>27.99</td>
<td>109.43</td>
</tr>
<tr>
<td>Configuration</td>
<td>1.28</td>
<td>3.13</td>
<td>23.08</td>
</tr>
<tr>
<td>Numerosity</td>
<td>2.56</td>
<td>1.19</td>
<td>50.79</td>
</tr>
<tr>
<td>Age × Strategy</td>
<td>1.28</td>
<td>27.99</td>
<td>6.98</td>
</tr>
<tr>
<td>Strategy × Configuration</td>
<td>1.28</td>
<td>1.36</td>
<td>37.54</td>
</tr>
<tr>
<td>Strategy × Numerosity</td>
<td>2.56</td>
<td>1.23</td>
<td>14.50</td>
</tr>
<tr>
<td>Configuration × Numerosity</td>
<td>2.56</td>
<td>0.57</td>
<td>6.79</td>
</tr>
<tr>
<td><strong>Mean amplitudes of saccades</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>1.28</td>
<td>9.94</td>
<td>12.34</td>
</tr>
<tr>
<td>Configuration</td>
<td>1.28</td>
<td>0.70</td>
<td>21.96</td>
</tr>
<tr>
<td>Numerosity</td>
<td>2.56</td>
<td>0.25</td>
<td>7.48</td>
</tr>
<tr>
<td>Age × Strategy</td>
<td>1.28</td>
<td>3.05</td>
<td>5.52</td>
</tr>
<tr>
<td>Configuration × Numerosity</td>
<td>2.56</td>
<td>0.24</td>
<td>5.71</td>
</tr>
<tr>
<td>Strategy × configuration × Numerosity</td>
<td>2.56</td>
<td>0.15</td>
<td>5.69</td>
</tr>
</tbody>
</table>

---

5 No effects came out significant on mean duration of fixations (all Fs<2.42).
6 The I-DT algorithm requires two parameters, the duration threshold and the dispersion threshold (the velocity is not considered with this algorithm). The minimum duration threshold was set to 60 ms, i.e., three consecutive points of regard. To initiate a fixation, the first three consecutive points had to be within a 50 pixel square. A fixation ended when a POR was considered as a fixation outlier (calculation based on two standard deviations). The first and the last fixations were included in the analyses. The last fixation ended when the space bar was pressed. No minimal inter-saccade distance was used here. Eye blinks were not considered in the saccade calculation process.
larger with more numerous random configurations of dots. Most interestingly, these effects were comparable in young and older adults. One exception to this was the Age × Strategy interaction, showing effects of age on latencies for the benchmark, but not for the anchoring strategy.

As all participants used both strategies on all problems, estimates of both latencies and accuracy and effects of strategies and problem features on performance resulted from strategy execution. Thus, faster performance with benchmark than with anchoring cannot come from participants using the former on small numerosities and the latter on large numerosities. As the benchmark strategy involves a quick visual scanning process and anchoring a series of enumerations, benchmark is faster (though less accurate) than the multi-step anchoring strategy, and is not influenced (or to a lesser extent) by problem features.

The only significant effects involving the age factor was a main effect of age on accuracy and an Age × Strategy interaction. Young adults were more accurate than older adults when they used both the strategies on all the problems. They were faster than older adults when they used the benchmark strategy. This effect likely stems from the nature of the processes involved in the benchmark strategy. Indeed, this strategy mainly implies visuo-spatial processes, known to be age-impaired (e.g., Hale, Myerson, Faust, & Fristoe, 1995; Jenkins, Myerson, Joerding, & Hale, 2000; Lima, Hale, & Myerson, 1991; Vecchi & Cornoldi, 1999). So, older adults were less efficient while executing them on all problems. It is also possible that older adults’ memory representations for large numerosities are less accurate and precise than young adults’.

In sum, controlling effects of strategy choices on participants’ approximate quantification performance showed that effects of
problem features cannot be entirely accounted for by different populations using different sets of strategies or the same set of strategies with different proportions on different types of problems. Without discarding contributions of strategy repertoire and strategy distributions on participants' performance, using the same set of processes is differently efficient for different types of problems and in different age groups.

Concerning eye-movement data, the analysis of mean number of fixations during the trial showed similar effects of strategy, age, and configuration on performance, providing converging evidence for strategy distinction. When eye movements were more specifically examined during the first 1000 ms, where the maximum differences between strategies were expected, they showed clear and interesting effects of time bins and strategies for all the problems. Saccade amplitudes were larger during the second 500-ms time bin than during the first one for the benchmark strategy, but not for the anchoring strategy.

These data suggest that our two estimation strategies involved different encoding processes. When participants used the benchmark strategy, participants made first saccades of small amplitudes, and then larger saccades. This suggests that they gradually encoded configurations of dots, concentrating first on a small portion of the stimulus and then fixating the rest of the stimulus (see van Oeffelen & Vos, 1984, for similar findings). In contrast, when they used the anchoring strategy, participants began to explore broadly the collection of dots Consistent with what participants claimed during post-experimental debriefing, this suggests that they first tried to find groups of dots. This first pre-enumeration scanning process was mainly associated to saccades of larger amplitudes. The most important and interesting findings to note from Experiment 2 concern comparisons of eye movements in young and older adults. On average, older adults tended to make more fixations and larger saccade amplitudes when using the benchmark strategy. This effect seems to be the consequence of a reduced useful field of view (UFOV; Ball, Beard, Roenker, Miller, & Griggs, 1988; Scialfa, Kline, & Lyman, 1987; Sekuler, Bennett, & Mamelak, 2000; Watson et al., 2005). The decreased UFOV account is consistent with some evidence in the literature that older adults maintain a more spatially limited (i.e., narrower) spatial distribution of attention than young adults (see also Greenwood & Parasuraman, 1999; Hartley & McKenzie, 1991; Madden & Gottlob, 1997; McCalley, Bouwhuis, & Juola, 1995; Trick, Perl, & Sethi, 2005; see however, Hartley, Kieley, & McKenzie, 1992; Madden, 1992). This reduced UFOV may restrict the spatial zone from which visual information can be efficiently extracted. Atchley and Hoffman.

Fig. 7. Mean saccade amplitudes for each strategy as a function of configuration and numerosity, in young and older adults (Experiment 2).

Fig. 8. (a) The simple effect of age on the mean saccade amplitudes during a bin. (b) The strategy \times bin interaction on the mean saccade amplitudes (Experiment 2).
4. General discussion

This study contributes to our further understanding of the cognitive processes in approximate quantification and adults’ age-related differences in these processes. The most original features of the present experiments include a strategy approach, coupled with collection of eye-movement data. The present findings showed that young and older adults use multiple strategies and execute strategies differently on different types of problems, and that strategy use and execution are influenced by both participants’ age and problem features. They also showed age-related differences and similarities in this domain, where they had never been investigated before. We discuss the implications of these findings to further understand approximate quantification and age-related changes in approximate quantification in particular and in strategic behaviours in general.

4.1. Implications for further understanding approximate quantification

The central question on approximate quantification is how participants provide a quick and rough estimate of the number of items in a collection. Previous works investigated the perceptually-based strategy (e.g., surface covered by stimulus, size and arrangement of items) that is often (but not always) correlated with numerosity. This study is the first one to show directly that, in addition to base their estimates on visual features of stimuli, people use a wide variety of approximate quantification strategies. Like in other cognitive tasks, the participants used between two and six strategies, and they selected strategies on a problem-by-problem basis (see Siegler, 1996, for an overview).

Multiple-strategy use and effects of problem features on strategy use are consistent with, and complement, previous results in approximate quantification (Crites, 1992; Siegel et al., 1982). These studies suggested that the participants used different strategies in addition to visually-based strategies to resolve verbal estimation problems. Here, by directly assessing strategies on each trial in each participant, Experiment 1 provided a more exhaustive, accurate, and detailed description of what the participants actually do when accomplishing approximate quantification tasks.

Moreover, in all the previous research on young adults, visual guessing was viewed as a single strategy (e.g., Lemaire & Lecacheur, 2007; Luwel et al., 2003a; b). The present findings showed that visual guessing actually includes a whole range of different counting, decomposition/recomposition, benchmark, and anchoring strategies. The present findings also suggest that the well-documented effects of problem features may be the result of the participants using different sets of strategies or the same strategies in different proportions on each type of problems. Consistent with this, problem features interacted with strategies. Indeed, participants’ latencies were influenced by problem features only when they used the anchoring strategy, and not when they used the benchmark strategy. This implies that effects of problem features on participants’ performance in previous studies of approximate quantification were contaminated by the strategies participants used and by how often each strategy was used on each type of problems. More generally, this implies that analyzing effects of problem or situation features on accuracy of estimates and solution latencies is too limited if we want to understand how people accomplish approximate quantification tasks. It is also necessary to know which strategies are used on each kind of problems, the relative frequencies with which each strategy is used on different problems, and how participants execute each of these strategies while solving different problems in different experimental conditions.

One additional original feature of this study was the collection of eye-movements. Our eye-movement data confirm that they are fruitful measures to reveal different encoding processes specifically engaged by each strategy, as it is the case in many problem situations (e.g., chess, Charness et al., 2001; Reingold, Charness, Pomplun, & Stampe, 2001; tumor problem, Grant & Spivey, 2003; arithmetic, Green et al., 2007). In particular, our results showed that eye-movements are sensitive to problem features, strategies, and participants’ age and might also be fruitful measures to further understand effects of problem features that previous findings proved crucial in estimation tasks (e.g., regularity of patterns of dots, area covered by dots, density or texture of dots).

4.2. Implications for further understanding aging effects

The comparison of young and older adults’ approximate quantification skills revealed very surprising age-related differences and similarities. These speak of two issues, namely aging effects in approximate quantification and strategic variations.

The present findings showed age-decreased approximate quantification performance, either in latencies alone (Expt. 1) or in both latencies and accuracy (Expt. 2), especially when they used time-consuming strategies on the hardest problems. After controlling strategy use, some age differences remained. These differences may explain age-differences in strategy selection. Further aging research should help to clarify this issue in the present context of approximate quantification. Finally, these findings complement previous results on aging and quantification. Exact counting processes decline with age only when participants are asked to quantify collections including 0–4 elements if distractors are present in the field (i.e., subitizing range), but not for larger collections (e.g., Basak & Verhaeghen, 2003; Geary & Lin, 1998; Kotary & Hoyer, 1995; Nebes et al., 1992; Sliwinski, 1997; Trick et al., 1996; Watson et al., 2002; Watson et al., 2005). The present results showed that approximate quantification performance also declines with age. This makes approximate quantification one of the numerical activities that is age-sensitive, in contrast to other numerical domains or tasks (e.g., exact counting of 4–10 items, some arithmetic problem-solving tasks; Duverne & Lemaire, 2005; Geary & Lin, 1998). Several specific aspects of approximate quantification may contribute to making it age-sensitive, above and beyond general cognitive factors like processing resources. Potential age-related degradation of memory representations for numerosities, impaired specific visual-spatial components of approximate quantification strategies, and encoding processes are examples of those aspects that future aging research should investigate.

The most important age-related difference was in how often young and older adults used each available strategy. The present age-related differences in strategy distributions were also found in many previous studies that assessed strategy distributions in different age groups. A critical issue about the age-related difference in strategy distribution concerns the factors influencing this distribution. Preliminary correlation analyses were conducted to determine the link between speed measures and mean percentages of use of each strategy. Only the mean percentages of use of exact counting correlated significantly (r = .29) with our mean speed z-scores (i.e., the mean of our three standardized speed measures). Further studies should investigate this link between between processing resources and strategy distribution (see Duverne and Lemaire, 2004) to explain age-related difference in approximate
quantification. The present age-related strategic variations points to the necessity to take into account these variations if we want to understand age-related differences in cognition. Of course, this requires, like here, assessing strategies on each problem to know which strategies the participants use and how often they use each available strategy, and to manipulate the type of strategies to investigate participants’ performance independently of strategy use. Only then we would be able to provide mechanistic accounts of age-related differences in cognitive performance and gain deeper understanding of cognitive aging.

Acknowledgements

This research was supported in part by the CNRS (French NSF), a grant from the Agence Nationale de la Recherche to PL (Grant # BLAN07-1_196867), and a graduate fellowship to DG. This work was supported by a grant from the Agence Nationale de la Recherche (Grant BLAN07-1_196867). We would like to thank Neil Charney, Lana Trick, and two anonymous reviewers for helpful comments on previous drafts of this paper.

References


