
This item was submitted to [Loughborough's Research Repository](#) by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

The measurement of approximate number system acuity across the lifespan is compromised by congruency effects

PLEASE CITE THE PUBLISHED VERSION

<https://doi.org/10.1177/1747021818779020>

PUBLISHER

SAGE Publications © Experimental Psychology Society

VERSION

AM (Accepted Manuscript)

PUBLISHER STATEMENT

This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Norris, Jade E., Sarah Clayton, Camilla K. Gilmore, Matthew Inglis, and Julie Castronovo. 2019. "The Measurement of Approximate Number System Acuity Across the Lifespan Is Compromised by Congruency Effects". figshare. <https://hdl.handle.net/2134/32944>.

The Measurement of Approximate Number System Acuity across the Lifespan is Compromised by Congruency Effects

Jade Eloise Norris^{1*}, Sarah Clayton², Camilla Gilmore³, Matthew Inglis³, Julie Castronovo^{4*}

¹ Department of Psychology, University of Bath, Claverton Down, Bath, BA2 7AY, UK

² Department of Health Sciences, University of Leicester, University Road, Leicester, LE1 7RH, UK

³ Mathematics Education Centre, Loughborough University, Loughborough, Leicestershire, LE11 3TU, UK

⁴ School of Life Sciences, University of Hull, Cottingham Road, Hull, HU6 7RX, UK

*Correspondence: Jade Eloise Norris, Department of Psychology, University of Bath, Claverton Down, Bath, BA2 7AY, UK. j.norris@bath.ac.uk | 01225 384 841

- The data were collected at the Department of Psychology, University of Hull

Word count: 5,278

Keywords: Approximate Number System, Number Sense, Ageing, Convex Hull, Visual Cues, Numerosity Discrimination

1 **Abstract**

2 Recent studies have highlighted the influence of visual cues such as dot size and cumulative surface
3 area on the measurement of the approximate number system (ANS). Previous studies assessing ANS
4 acuity in ageing have all applied stimuli generated by the Panamath protocol, which does not control
5 nor measure the influence of convex hull. Crucially, convex hull has recently been identified as an
6 influential visual cue present in dot arrays, with its impact on older adults' ANS acuity yet to be
7 investigated. The current study therefore investigated the manipulation of convex hull by the
8 Panamath protocol, and its effect on the measurement of ANS acuity in younger and older
9 participants. Firstly, analyses of the stimuli generated by Panamath revealed a confound between
10 numerosity ratio and convex hull ratio. Secondly, although older adults were somewhat less
11 accurate than younger adults on convex hull incongruent trials, ANS acuity was broadly similar
12 between the groups. These findings have implications for the valid measurement of ANS acuity
13 across all ages, and suggest that the Panamath protocol produces stimuli that do not adequately
14 control for the influence of convex hull on numerosity discrimination.

15

16 Introduction

17 The Approximate Number System (ANS) supports the imprecise representation of numerosity, as
18 demonstrated by behavioural and neuronal indicators of Weber’s law: i) numerical representations
19 become less precise and more approximate with increasing magnitude (the size effect), and ii)
20 discrimination between two numerosities becomes more difficult as their ratio approaches 1 (the
21 ratio effect) (Dehaene, 1997; Gallistel & Gelman, 2000; Izard, Sann, Spelke, & Streri, 2009; Piazza &
22 Izard, 2009; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004). The acuity of the ANS is most often
23 measured with comparison tasks, whereby participants are shown two arrays of non-symbolic
24 numerosities (e.g. dots), and asked to select which array is most numerous without counting.
25 Comparing performance on easier and harder ratios provides evidence for the ratio effect. However,
26 the validity of such tasks in providing a pure measure of ANS acuity is contested, because
27 participants are found to be influenced by visual characteristics of the stimuli including convex hull
28 (the perimeter around a dot set, sometimes referred to as “area extended”), average dot size, and
29 cumulative surface area of the dots (Clayton & Gilmore, 2014; Gebuis & Reynvoet, 2012a, 2012b,
30 2012c; Gilmore, Cragg, Hogan, & Inglis, 2016; Leibovich & Henik, 2013; Szűcs, Nobes, Devine,
31 Gabriel, & Gebuis, 2013). It is generally accepted that when these visual characteristics are
32 uncontrolled, i.e. the more numerous set is also larger in terms of its non-numerical visual cues,
33 participants may make their decisions using visual cues alone (e.g. by choosing the array that
34 contains larger dots on average compared to the other array), without engaging the ANS (Gebuis &
35 Reynvoet, 2012a, 2012b). Here we explore how these visual cue characteristics are manipulated in a
36 commonly-used programme to generate dot array stimuli, and how this impacts on numerosity
37 judgements across the lifespan.

38 When creating dot array stimuli, researchers originally sought to address concerns about the
39 influence of visual cues by applying controls to average dot size and cumulative surface area, varying
40 the relationship of these visual cues with the number of dots in the array (Abreu-Mendoza, Soto-
41 Alba, & Arias-Trejo, 2013). For example, as described in the software guidelines for Panamath
42 (Halberda, Mazocco, & Feigenson, 2008), a commonly used method for generating stimuli for non-
43 symbolic numerosity comparison tasks, dot-size congruency is controlled by manipulating the
44 cumulative surface area of the arrays. During a congruent trial, cumulative surface area is positively
45 correlated with numerosity. The more numerous array therefore has a larger cumulative surface
46 area and a larger average dot size: cumulative surface area and dot size are both congruent to
47 numerosity. During what we will term a matched trial (to reduce confusion between an
48 ‘incongruent’ trial as defined by Halberda et al. (2008) and incongruent visual cues in the more

49 general sense), cumulative surface area is matched between arrays in order that the less numerous
50 array has a larger average dot size: dot size is incongruent to numerosity. Finally, during
51 anticorrelated trials, cumulative surface area (and therefore average dot size too) are negatively
52 correlated with, and so incongruent to, numerosity (Halberda et al., 2008). In short, only dot size is
53 incongruent to numerosity during matched trials, whereas during anticorrelated trials both
54 cumulative surface area *and* average dot size are incongruent to numerosity. The application of such
55 controls has varied between studies using a range of protocols to generate dot arrays: whilst some
56 have included congruent and matched trials (Cappelletti, Didino, Stoianov, & Zorzi, 2014), others
57 have compared congruent and anticorrelated trials (Clayton, Gilmore, & Inglis, 2015; Gilmore et al.,
58 2013; Hurewitz, Gelman, & Schnitzer, 2006; Inglis & Gilmore, 2014; Odic, Libertus, Feigenson, &
59 Halberda, 2013; Szűcs et al., 2013), with others using all three control conditions (DeWind &
60 Brannon, 2012; Fuhs & McNeil, 2013; Keller & Libertus, 2015; Rousselle & Noël, 2008) or matched
61 trials alone (Gray & Reeve, 2014). Varied methods of visual cue control, along with other
62 inconsistencies such as display time, number of trials, and numerosity ratio cause problems when
63 comparing ANS acuity across studies (Clayton & Gilmore, 2014; Clayton et al., 2015; Dakin, Tibber,
64 Greenwood, Kingdom, & Morgan, 2011; Dietrich, Huber, & Nuerk, 2015; Gebuis & Reynvoet, 2012c;
65 Gilmore et al., 2016; Inglis & Gilmore, 2013, 2014; Szűcs et al., 2013).

66 Several authors have argued that during trials with incongruent visual cues, participants must first
67 inhibit the influence of those visual cues in order to perform a numerosity judgement (Cappelletti et
68 al., 2014; Cappelletti, Pikkat, Upstill, Speekenbrink, & Walsh, 2015; Clayton & Gilmore, 2014; Fuhs &
69 McNeil, 2013; Gilmore et al., 2013, 2016). However, others find similar performance between
70 congruent trials and those with incongruent visual cues, arguing that performance on ANS tasks does
71 not require inhibitory control (Keller & Libertus, 2015; Odic, Hock, & Halberda, 2014; Odic et al.,
72 2013). This is important when considering ANS acuity in ageing. Inhibitory control declines with age
73 (Hasher & Zacks, 1988; Kramer, Humphrey, Larish, & Logan, 1994): if incongruent visual cues in ANS
74 tasks do indeed require inhibition, then older participants may be expected to show a greater
75 decline in performance on such trials. A limited number of studies have investigated ANS acuity in
76 ageing, with some examining the impact of congruency effects. Halberda, Ly, Wilmer, Naiman, and
77 Germine (2012) investigated ANS acuity across the lifespan, concluding that acuity declines with
78 increasing age beyond 30 years. However, it is difficult to draw conclusions from this study regarding
79 the impact of older age for three reasons. Firstly, participants aged 45-85 were categorised within
80 one age group, due to the small number of older adults included in the overall sample. Secondly,
81 Figure 3 (p.11119: Halberda et al., 2012) demonstrates highly variable ANS acuity and several
82 outliers within the older group. Thirdly, although congruent and matched trials were used, whether

83 age-related decline may be attributable to poorer performance during matched trials, where
84 inhibitory control may be required, is not reported. Indeed, in the first study to directly investigate
85 the impact of dot-size congruency on ANS acuity in ageing, Cappelletti et al. (2014) compared
86 younger and older adults' performances on an ANS task based on the Halberda et al. (2008)
87 Panamath protocol. Their findings initially indicated declined ANS acuity in older age (as in Halberda
88 et al., 2012). However, separate analyses for performances on congruent and matched trials
89 revealed that the older group's acuity was only declined compared to the younger group when
90 average dot size was incongruent to numerosity (matched trials). The authors concluded that
91 seemingly poorer ANS acuity in ageing may be accounted for by declined inhibitory control (Hasher
92 & Zacks, 1988) rather than deteriorated approximate numerical skills. In short, older adults' ability to
93 inhibit the influence of an incongruent visual cue (dot size) was found to be declined. A later study
94 by Cappelletti et al. (2015) administered ANS training paired with parietal stimulation to younger
95 and older adults to investigate whether ANS acuity could be enhanced. Acuity was similar between
96 groups at pre-training, with improvement in both groups after training. Crucially, older adults'
97 stronger ANS acuity post-training was driven by improved performance for matched trials (i.e. with
98 incongruent dot size), which was related to smaller interference effects on traditional inhibition
99 tasks. Moreover, older adults' success in learning to inhibit non-numerical magnitudes on the ANS
100 task led to poorer performance on tasks assessing the discrimination of such magnitudes (e.g. length
101 discrimination). These findings further support the existence of an inhibitory component to ANS
102 tasks, a finding which may be particularly evident in older adults due to age-related decline in
103 inhibitory control (Cappelletti et al., 2014; Hasher & Zacks, 1988). However, another study using the
104 same protocol and comparable methods found similar ANS acuity for younger and older adults, even
105 for matched trials (Norris, McGeown, Guerrini, & Castronovo, 2015). It is likely that other
106 methodological differences such as the use of intermixed vs separated dot displays contributed to
107 these contradictory findings (see Norris & Castronovo, 2016 for evidence of the impact of different
108 stimuli presentation methods in younger adults).

109 Although some of the studies investigating ANS acuity in ageing have examined the impact of dot
110 size congruency, recently the influence of convex hull congruency on numerosity discrimination has
111 been emphasised (Clayton & Gilmore, 2014; Clayton et al., 2015; Gebuis & Reynvoet, 2012c; Gilmore
112 et al., 2016). Convex hull refers to the smallest possible perimeter that can be drawn around an
113 array (Graham, 1972), and may affect the processing of numerosity to a greater extent than dot size,
114 even when convex hull and numerosity are not correlated (Gebuis & Reynvoet, 2012c). As with dot
115 size, convex hull can be congruent (the more numerous array has the larger convex hull), or
116 incongruent to numerosity (the more numerous array has the smaller convex hull). Crucially, the

117 studies reviewed above which investigated ANS acuity in ageing did not investigate convex hull
118 congruency effects, as they used stimuli generated by the Panamath protocol, which does not
119 control for convex hull congruency. Clayton and Gilmore (2014) investigated how manipulating
120 numerosity mediated the influence of visual cues on 7-9 year-olds' ANS acuity. As numerosity
121 increased, performance declined due to the increasing interference of convex hull. However for
122 smaller numerosities, performance was most strongly influenced by dot size. Therefore, the type of
123 visual cues used by participants appears to be mediated by numerosity (Clayton & Gilmore, 2014).
124 Crucially, below-chance performance during larger-numerosity trials demonstrated the greater
125 influence of convex hull over other visual cues. In a further study, Clayton et al. (2015) compared
126 ANS acuity when measured with two commonly-used protocols: Panamath, which controls total
127 cumulative surface area in order to manipulate dot size (Halberda et al., 2008), and a script by
128 Gebuis and Reynvoet (2011) which controls both the cumulative surface area and convex hull of dot
129 arrays. The authors not only found poorly correlated performance between the protocols, but also
130 diverging interactions between cumulative surface area and convex hull congruencies: for the
131 Gebuis and Reynvoet (2011) paradigm, accuracy was higher for convex hull incongruent trials when
132 cumulative surface area was congruent compared to when it was incongruent. However, cumulative
133 surface area did not significantly affect performance when convex hull was congruent. Therefore,
134 convex hull congruency appears to influence numerosity comparison performance to a greater
135 extent than cumulative surface area congruency with the Gebuis and Reynvoet (2011) paradigm.
136 Performance was also enhanced during convex hull congruent trials on the Panamath task. However
137 on the Panamath task, accuracy was higher for cumulative surface area incongruent trials compared
138 to cumulative surface area congruent trials regardless of convex hull. Finally, Gilmore et al. (2016)
139 demonstrated that although dot size influences children's performance on an ANS task, this effect
140 decreases into adulthood, whereas the influence of convex hull remains consistent from childhood
141 to adulthood. These findings emphasise the influence of convex hull during ANS tasks, highlighting
142 the necessity to investigate the effect of convex hull on protocols which do not control it, such as
143 Panamath (see also DeWind & Brannon, 2016). Indeed, as performance on the Panamath task is
144 significantly influenced by convex hull (Clayton et al., 2015; DeWind & Brannon, 2016), it is unclear
145 to what extent previous findings of age-related decline in ANS acuity may be due to poorer
146 performance on convex hull-incongruent trials.

147 The influence of convex hull on ANS acuity as measured by the Panamath protocol is therefore a
148 timely and important consideration in examining the inhibitory components of ANS tasks. Moreover,
149 as the studies to date investigating the impact of ageing on ANS acuity have all used stimuli
150 generated by Panamath (Cappelletti et al., 2014; Halberda et al., 2012; Norris et al., 2015), a

151 protocol which does not manipulate nor measure the impact of convex hull, it is important to
152 consider the influence of convex hull on the conclusions drawn in these studies: that older adults'
153 poorer inhibitory control leads to declined performance during dot-size incongruent trials
154 (Cappelletti et al., 2014, 2015). It is therefore necessary to examine whether similar mechanisms
155 may shape the impact of convex hull congruency on older adults' performances, especially
156 considering recent findings indicating that convex hull constitutes a more important predictor of ANS
157 task performance compared to other visual cues (Clayton & Gilmore, 2014; Clayton et al., 2015;
158 Gebuis & Reynvoet, 2012c). Therefore, the primary aim of the current study was to further examine
159 the nature of the visual cues in the stimuli generated by the Panamath protocol, with a particular
160 focus on the way in which convex hull varies. Secondly, the study investigated the extent to which
161 convex hull congruency affects ANS task performance for older and younger adults.

162

163 **Method**

164 *Participants*

165 Forty participants were recruited, 20 older adults aged 62-70 (14 females; $M_{age} = 65$, $SD = 2.9$) and 20
166 younger adults aged 18-24 (16 females; $M_{age} = 20$, $SD = 1.8$). Younger participants were recruited
167 through the Department of Psychology at the University of Hull and received course credit. Older
168 participants were recruited by the first author from the local community and participated
169 voluntarily. The study was approved by the Department of Psychology Ethics Committee at the
170 University of Hull. All participants were fully informed of the study aims and provided written
171 consent.

172 *Screening*

173 Participants were screened at recruitment for a history of psychiatric disorder, depression, or
174 abnormal vision. Older adults were administered the Mini Mental State Exam (MMSE: Folstein,
175 Folstein, & McHugh, 1975) with a score $<27/30$ providing a cut-off for exclusion. The Geriatric
176 Depression Scale (GDS: Yesavage et al., 1982) was administered to all participants, with a score >5
177 providing a cut-off point (as in Norris et al., 2015). No participants were excluded due to scores
178 beyond cut-off on the MMSE or GDS.

179 *Procedure*

180 Approximate Number System acuity was measured using the downloadable Panamath software
181 (Halberda et al., 2008). Two spatially separate arrays of between 5 and 21 yellow and blue dots were
182 presented simultaneously side-by-side on a grey background for 200ms (yellow on the left, blue on
183 the right), followed by a 200ms backward mask of randomly distributed yellow and blue pixels.
184 Participants initiated each trial using the space bar, and were asked to decide which array was more
185 numerous (yellow or blue). Participants responded as quickly as possible without sacrificing accuracy
186 using the 'A' (yellow) and 'L' (blue) keys, which were covered with correspondingly-coloured dots.
187 The dot stimuli were generated with two within-subject factors: visual cue control with 3 levels
188 (congruent [both cumulative surface area and average dot size are congruent to numerosity],
189 matched [cumulative surface area is matched and average dot size is incongruent], and
190 anticorrelated [both cumulative surface area and average dot size are incongruent]), and numerosity
191 ratio bin with 4 levels (1.1-1.19, 1.19-1.28, 1.32-1.43, and 2.28-2.47; ratio bins 1, 2, 3, and 4
192 respectively). There were 420 trials in total. Convex hull size and convex hull congruency were
193 calculated post hoc for each trial by using the Graham (1972) scan algorithm on screenshots of each

194 trial as generated by Panamath. This calculation also summed the total number of yellow and blue
195 pixels, providing a measure of the cumulative surface area of each array.

196 **Results**

197 We first report the visual characteristics of the stimuli generated by the Panamath protocol,
198 followed by an examination of the impact of these characteristics on young and older adults'
199 performance on the ANS task.

200 *Visual characteristics of the stimuli*

201 In order to control for the effect of visual cues, Panamath is designed to generate three types of
202 stimuli: congruent (cumulative surface area and average dot size positively correlate with
203 numerosity), matched (cumulative surface area is matched to numerosity, and average dot size
204 negatively correlates with numerosity) and anticorrelated (cumulative surface area and dot size
205 negatively correlate with numerosity). However, when we calculated the cumulative surface area of
206 the arrays by summing the number of blue and yellow pixels, we discovered that the matched trials
207 were not precisely matched in terms of cumulative surface area: the more numerous array always
208 had a greater cumulative surface area (mean pixel number difference = 150, range = 2-592).
209 Therefore in matched trials, cumulative surface area was actually congruent, even though in some
210 cases there was only a small pixel-number difference. Consequently, for our analyses we collapsed
211 the three Panamath conditions (congruent, matched, and anticorrelated) into two (cumulative
212 surface area congruent [congruent and matched] vs. cumulative surface area incongruent
213 [anticorrelated]).¹

214 With the convex hull of the arrays calculated for each trial, we sought to investigate to what extent
215 the Panamath protocol produced equally-weighted convex hull congruent and incongruent trials,
216 and how this was affected by other factors within the protocol (numerosity ratio and cumulative
217 surface area).

218 Figure 1 depicts the relationships between within-subjects factors on the Panamath protocol.
219 Numerosity ratio here refers to (left set/right set), rather than (larger set/smaller set) as defined by
220 Panamath, so that cumulative surface area ratio, convex hull ratio and numerosity ratio were
221 calculated in the same manner. The correlation between convex hull ratio and numerosity ratio ($r =$
222 $.720$, 95% CI $[.671, .763]$, $p < .001$) demonstrates a confound between within-subject factors on the
223 Panamath protocol: convex hull ratio increases with increasing numerosity ratio.

224

¹ Analysing the data with the original three levels of congruency as defined by the Panamath protocol did not affect the direction or the significance of the results.

225 (Figure 1 about here)

226 Moreover, an examination of the Panamath-defined numerosity ratio bins not only replicates the
227 finding that convex hull ratio (and therefore convex hull congruency) increases alongside numerosity
228 ratio, but also indicates that all trials in ratio bin 4 had a congruent convex hull (see Table 1). Indeed,
229 convex hull was congruent for the majority of all trials (335/420).

230 *Table 1: Number of trials per Panamath-defined numerosity ratio bin in the convex hull congruent*
231 *and incongruent conditions*

Numerosity Ratio (Bin)	Convex hull	
	Congruent	Incongruent
1.1-1.19 (1)	71	34
1.19-1.28 (2)	80	25
1.32-1.43 (3)	79	26
2.28-2.47 (4)	105	0

232

233 *ANS acuity*

234 The following analyses focus on accuracy as the dependent variable for ANS acuity, because
235 accuracy is thought to provide the most reliable and valid measure (Clayton et al., 2015; Guillaume,
236 Gevers, & Content, 2015; Inglis & Gilmore, 2014). As would be expected due to the ratio effect on
237 numerosity discrimination, there was a positive by-items correlation between accuracy and
238 numerosity ratio ($r = .563$, 95% CI [.494, .625], $p < .001$). Next, we investigated the effect of age
239 group, cumulative surface area congruency, and convex hull congruency on accuracy. A mixed
240 ANOVA was conducted with cumulative surface area congruency (congruent, incongruent) and
241 convex hull congruency (congruent, incongruent) as within-subjects factors, and age group (older,
242 younger) as a between-subjects factor. There were no main effects of cumulative surface area
243 congruency ($F(1, 38) = 3.185$, $p = .082$, $\eta_p^2 = .077$, $BF_{inclusion} = 1.07^2$) or age group ($F(1, 38) = .628$, $p =$

² The inclusion Bayes Factor compares the evidence in support of each effect by comparing across all possible models including the effect with all possible models without the effect. This was calculated in JASP.

244 .433, $\eta_p^2 = .016$, $BF_{inclusion} = 0.48$), and no interaction between age group and cumulative surface area
245 congruency ($F(1, 38) = 1.891$, $p = .177$, $\eta_p^2 = .047$, $BF_{inclusion} = 0.50$). However, accuracy was
246 significantly higher for convex hull congruent trials ($M = 83.25\%$, $SD = 37.34$) than for convex hull
247 incongruent trials ($M = 72.09\%$, $SD = 44.86$: $F(1, 38) = 258.190$, $p < .001$, $\eta_p^2 = .872$, $BF_{inclusion} >$
248 10^{305}). Moreover, convex hull congruency interacted with cumulative surface area congruency
249 ($F(1, 38) = 5.005$, $p = .031$, $\eta_p^2 = .116$, $BF_{inclusion} = 2.28$, as in Clayton et al., 2015): Figure 2
250 demonstrates that when convex hull was congruent, performance between cumulative surface area
251 congruent ($M = 83.40\%$, $SD = 37.21$) and incongruent trials ($M = 83.10\%$, $SD = 37.48$) was similar ($p =$
252 $.604$, Cohen's $d = .012$: LSD pairwise comparisons). However during convex hull incongruent trials,
253 participants tended to respond more accurately when cumulative surface area was incongruent
254 compared to when it was congruent ($p = .038$, Cohen's $d = -.064$).

255 (Figure 2 about here)

256 The interaction between convex hull congruency and age group was significant ($F(1, 38) = 4.328$, $p =$
257 $.044$, $\eta_p^2 = .102$, $BF_{inclusion} = 1.02$). Although accuracy on convex hull congruent trials was similar for
258 the younger ($M = 83.22\%$, $SD = 37.37$) and older groups ($M = 83.29\%$, $SD = 37.32$; $p = .968$, Cohen's d
259 $= .002$: LSD pairwise comparisons), younger adults outperformed older adults when convex hull was
260 incongruent (younger: $M = 73.35\%$, $SD = 44.22$; older: $M = 70.82\%$, $SD = 45.47$), although this
261 difference did not reach significance ($p = .200$, Cohen's $d = .056$: LSD pairwise comparisons, see
262 Figure 3).³

263 (Figure 3 about here)

264 Finally, the interaction between cumulative surface area congruency, convex hull congruency, and
265 age group did not reach significance ($F(1, 38) = 1.610$, $p = .212$, $\eta_p^2 = .041$, $BF_{inclusion} = 0.67$).

³ In light of the evidence for the influence of convex hull during numerosity discrimination in ageing, we reanalysed the findings from our previous study, where ANS acuity was similar for younger and older adults regardless of dot-size congruency (Norris, McGeown, Guerrini, & Castronovo, 2015). Convex hulls were calculated using the Graham (1972) scan algorithm. Responses to convex hull incongruent trials ($M = 78.00\%$, $SD = 41.44$) were found to be significantly less accurate than to convex hull congruent trials ($M = 91.16\%$, $SD = 28.40$: $F(1,48) = 231.077$, $p < .001$, $\eta_p^2 = .828$). Moreover, there was a marginal interaction between convex hull congruency and age group ($F(1,48) = 4.017$, $p = .051$, $\eta_p^2 = .077$), due to a tendency for poorer performance during convex hull incongruent trials for the older group compared to the younger group. Crucially, the impact of convex hull congruency on performance for all participants was significant, whereas the effect of dot size congruency was not.

266 **Discussion**

267 The current study investigated the impact of visual cue congruency on ANS acuity as measured by
268 the Panamath protocol in a group of younger and older adults. For the first time, we investigated
269 patterns of both cumulative surface area congruency and convex hull congruency on trials generated
270 by the Panamath protocol and their impact on older adults' ANS acuity. Although convex hull
271 congruency has been found to affect numerosity processing to a greater extent than the visual cues
272 that have been more frequently controlled in previous research (e.g. cumulative surface area and
273 dot size) (Clayton & Gilmore, 2014; Clayton et al., 2015; Gebuis & Reynvoet, 2012c; Gilmore et al.,
274 2016), to date, only the impact of dot-size congruency on ANS acuity in ageing has been examined.
275 In some studies, poorer inhibitory control has been proposed to account for declined performance
276 during dot size-incongruent trials in older age (Cappelletti et al., 2014, 2015). These studies had used
277 stimuli generated by the Panamath protocol, which does not control convex hull congruency. The
278 current study therefore explored the visual characteristics of stimuli generated by the Panamath
279 protocol and their impact on the measurement of ANS acuity, whilst directly investigating whether
280 older adults performed more poorly when convex hull was incongruent. The current findings
281 demonstrate that the Panamath protocol produces dot arrays that are confounded between convex
282 hull ratio and numerosity ratio. There was some evidence that older adults appeared to be more
283 susceptible to the influence of convex hull information when making numerosity judgements, but
284 the key test of this effect was only borderline significant ($p = .044$). Potential explanations for these
285 findings are discussed below.

286 *Visual characteristics of dot array stimuli*

287 Our analyses indicate that the Panamath protocol generates stimuli which favour a congruent over
288 an incongruent convex hull, an effect which becomes more pronounced as numerosity ratio
289 increases. These findings have clear implications for studies using the Panamath protocol. Because
290 convex hull is congruent on most trials, this may improve overall performance on the task. Crucially,
291 the current findings demonstrate that numerosity ratio and convex hull ratio can be confounded on
292 stimuli generated by the Panamath protocol. This affects the valid and reliable measurement of ANS
293 acuity because participants could perform at above-chance levels on dot comparison tasks purely by
294 responding on the basis of convex hull information and without the need to engage in numerosity
295 processing. Our findings also support the suggestion that participants integrate multiple visual cues
296 during numerosity discrimination (Clayton et al., 2015; Gebuis & Reynvoet, 2012b), resulting in
297 interactions when visual cues vary in their congruency with numerosity: when convex hull was
298 incongruent to numerosity, participants were more accurate when cumulative surface area was also

299 incongruent (as in Clayton et al., 2015; Gebuis & van der Smagt, 2011; Keller & Libertus, 2015). The
300 findings therefore emphasise the necessity of considering the impact of non-numerical visual cues
301 and the interactions between such cues during numerosity comparison.

302 *ANS acuity*

303 Reflecting previous findings, overall ANS acuity was similar between age groups, with no age-related
304 decline in performance on cumulative surface area incongruent trials (as in Norris et al., 2015).
305 Performance was poorer for all participants during convex hull incongruent trials. Although the
306 convex hull congruency effect appeared to be more pronounced for the older group compared to
307 the younger group, this effect was small and was not well-supported by the Bayesian analysis
308 compared to the overall influence of convex hull for all participants. Previous findings of stronger
309 dot-size congruency effects for older compared to younger adults may suggest that similar findings
310 should emerge for convex hull congruency. One possible explanation for the weak evidence in the
311 current study of a stronger convex hull congruency effect in older age may be that, due to the
312 confounded nature of the Panamath protocol, a relatively small number of convex hull incongruent
313 trials were generated. In previous studies, larger proportions of dot-size incongruent trials have
314 been used. Therefore, stronger evidence for the interaction between convex hull congruency and
315 age group, and indeed even a group difference for overall ANS acuity may have emerged had the
316 number of convex hull congruent and incongruent trials been equally-weighted. It is well established
317 from studies of inhibition tasks that the overall proportion of congruent and incongruent trials
318 impacts on the size of congruency effects (Logan & Zbrodoff, 1979). Crucially here, these effects are
319 not consistent in younger and older adults. West and Baylis (1998) found that the difference
320 between younger and older adults on a Stroop task was greater when the task consisted mostly of
321 incongruent trials compared with mostly congruent trials. It is possible therefore that the small
322 proportion of incongruent trials in the task used here may have masked differences between
323 younger and older adults that could be apparent in a more evenly-balanced version of the task.

324 Overall however, our results support the suggestion that convex hull affects numerosity
325 discrimination to a greater extent than dot-size or cumulative surface area when measuring ANS
326 acuity (Clayton & Gilmore, 2014; Clayton et al., 2015; Gebuis & Reynvoet, 2012c; Gilmore et al.,
327 2016).

328 *Methodological and Theoretical Conclusions*

329 The current study highlights that significant confounds may exist in the dot array stimuli produced by
330 the Panamath protocol, indicating that trials are overall more likely to be convex hull congruent,

331 possibly facilitating performance. A confound between convex hull ratio and numerosity ratio raises
332 concerns about the validity and reliability of non-symbolic numerosity comparison tasks conducted
333 with stimuli generated by the Panamath protocol. Moreover, in light of recent claims that the
334 Panamath protocol does not produce congruency effects (Keller & Libertus, 2015; Odic et al., 2014,
335 2013), our investigation indicates that researchers may have been focusing on the wrong visual cue:
336 convex hull appears to affect numerosity processing over and above dot-size and cumulative surface
337 area. Controlling only dot size is therefore insufficient (as in the Panamath protocol and other stimuli
338 generation methods used in the literature: e.g. Dehaene, Izard, & Piazza, 2005), as multiple visual
339 cues may be simultaneously extracted from dot arrays during numerosity discrimination (Clayton et
340 al., 2015; Gebuis & Reynvoet, 2012a, 2012c). Here we found that convex hull appears to exert more
341 of an influence on ANS task performance compared to the other visual cues present in a numerosity
342 display. Previous investigations, using different methods of generating dot stimuli, have found that
343 several visual cues (total circumference, convex hull, density, and cumulative surface area) influence
344 numerosity judgements over and above the influence of numerosity information (Leibovich & Henik,
345 2014).

346 The current findings also have implications for studies reporting a link between ANS acuity and
347 mathematical achievement. As researchers have used a range of visual cue-controls and numerosity
348 ratios in generating dot arrays, it is unclear to what extent relationships with mathematical
349 achievement may in fact reflect a link with the inhibitory control required to ignore convex hull
350 (Clayton & Gilmore, 2014; Clayton et al., 2015; Fuhs & McNeil, 2013; Gilmore et al., 2013, 2016;
351 Norris & Castronovo, 2016; Szűcs et al., 2013). Indeed, as studies assessing the ANS in children often
352 use easier (i.e. larger) numerosity ratios, convex hull may facilitate performance to an even greater
353 extent in these studies (as convex hull is more likely to be congruent). The current study therefore
354 highlights the necessity for researchers to seriously consider the influence of convex hull on
355 numerosity discrimination when exploring its relationship with maths achievement (Clayton et al.,
356 2015; Gilmore et al., 2016). Moreover, the results raise questions regarding previous conclusions of
357 declined ANS acuity in ageing (Halberda et al., 2012), and whether dot-size congruency can fully
358 account for these effects (Cappelletti et al., 2014; Norris et al., 2015). Future research must
359 investigate whether ANS acuity is declined in ageing when convex hull is systematically controlled:
360 should older adults' performances on more stringently-controlled paradigms (e.g. Gebuis &
361 Reynvoet, 2011) be declined compared to younger participants, this would support the suggestion
362 that ANS tasks involve inhibitory control (Clayton & Gilmore, 2014; Fuhs & McNeil, 2013; Gilmore et
363 al., 2013; Szűcs et al., 2013), and that convex hull affects numerosity discrimination to a greater

364 extent than dot size and cumulative surface area (Clayton & Gilmore, 2014; Clayton et al., 2015;
365 Gebuis & Reynvoet, 2012c; Gilmore et al., 2016).

366 Finally, the current findings contribute to the theoretical debate surrounding the extent to which
367 numerosity is the cue primarily extracted from dot arrays over and above the other approximate
368 quantities present (the non-numerical visual cues). On one hand, some argue that numerosity is the
369 primary cue extracted from non-symbolic arrays, and that numerosity therefore drives performance
370 on dot discrimination tasks, as opposed to the other non-symbolic quantities present in the array
371 (e.g. convex hull, dot size, cumulative surface area: Barth et al., 2006; Halberda et al., 2008). On the
372 other hand, the competing processes account (Gilmore et al., 2013) proposes that non-numerical
373 visual cues *are* extracted during numerosity discrimination, and that participants must inhibit their
374 influence in order to then respond to numerosity. It is therefore proposed that numerosity and the
375 other visual cues present in non-symbolic arrays must compete to be processed, with two possible
376 outcomes: firstly, participants may respond based on the salience of various visual cues (bigger dots,
377 larger convex hull). Secondly, if participants are able to inhibit a response to these salient visual
378 cues, then they can respond to numerosity. In the current study, convex hull congruency affected
379 numerosity discrimination performance: these findings therefore provide further evidence that
380 participants must inhibit convex hull when it is incongruent before being able to respond to
381 numerosity (Clayton & Gilmore, 2015; Gilmore et al., 2013), supporting the competing processes
382 hypothesis. In addition, some researchers have suggested that numerosity isn't primarily extracted
383 during ANS tasks, but rather a weighted combination of non-numerical visual cues is used by
384 participants to discriminate between dot arrays (Gebuis & Gevers, 2011; Gebuis & Reynvoet, 2012b,
385 2012c). The current study cannot rule out this suggestion. Although our data indicate a numerosity
386 ratio effect, in considering the confounded nature of the stimuli, this doesn't necessarily indicate
387 that numerosity is the primary cue being extracted. Because the majority of trials were convex hull
388 congruent, participants may not need to extract numerosity to make a correct discrimination on the
389 majority of trials. Indeed, participants could most often discriminate between the arrays using
390 convex hull and achieve above-chance overall accuracy for the current study. Our findings therefore
391 further highlight the need to directly investigate the influence of convex hull on numerosity
392 discrimination performance, particularly for protocols which do not manipulate nor measure convex
393 hull, and where such protocols have facilitated certain conclusions about the ANS (e.g. that it is
394 declined in ageing: Halberda et al., 2012; or that its acuity can predict formal mathematical
395 attainment: Halberda et al., 2008).

396 **References**

397 Abreu-Mendoza, R. A., Soto-Alba, E. E., & Arias-Trejo, N. (2013). Area vs. density: influence of visual
398 variables and cardinality knowledge in early number comparison. *Frontiers in Psychology, 4*.
399 <https://doi.org/10.3389/fpsyg.2013.00805>

400 Cappelletti, M., Didino, D., Stoianov, I., & Zorzi, M. (2014). Number skills are maintained in healthy
401 ageing. *Cognitive Psychology, 69*, 25–45. <https://doi.org/10.1016/j.cogpsych.2013.11.004>

402 Cappelletti, M., Pikkat, H., Upstill, E., Speekenbrink, M., & Walsh, V. (2015). Learning to Integrate
403 versus Inhibiting Information Is Modulated by Age. *Journal of Neuroscience, 35*(5), 2213–
404 2225. <https://doi.org/10.1523/JNEUROSCI.1018-14.2015>

405 Clayton, S., & Gilmore, C. (2014). Inhibition in dot comparison tasks. *ZDM, 1*–12.
406 <https://doi.org/10.1007/s11858-014-0655-2>

407 Clayton, S., Gilmore, C., & Inglis, M. (2015). Dot comparison stimuli are not all alike: The effect of
408 different visual controls on ANS measurement. *Acta Psychologica, 161*, 177–184.
409 <https://doi.org/10.1016/j.actpsy.2015.09.007>

410 Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., & Morgan, M. J. (2011). A common
411 visual metric for approximate number and density. *Proceedings of the National Academy of
412 Sciences, 108*(49), 19552–19557. <https://doi.org/10.1073/pnas.1113195108>

413 Dehaene, S. (1997). *The Number Sense: How the Mind Creates Mathematics*. New York: Oxford
414 University Press.

415 Dehaene, S., Izard, V., & Piazza, M. (2005). Control over non-numerical parameters in numerosity
416 experiments. *Unpublished Manuscript (available on Wwww.unicog.org)*.

417 DeWind, N. K., & Brannon, E. M. (2012). Malleability of the approximate number system: effects of
418 feedback and training. *Frontiers in Human Neuroscience, 6*.
419 <https://doi.org/10.3389/fnhum.2012.00068>

420 DeWind, N. K., & Brannon, E. M. (2016). Significant Inter-Test Reliability across Approximate Number
421 System Assessments. *Frontiers in Psychology, 7*. <https://doi.org/10.3389/fpsyg.2016.00310>

422 Dietrich, J. F., Huber, S., & Nuerk, H.-C. (2015). Methodological aspects to be considered when
423 measuring the approximate number system (ANS) – a research review. *Cognition*, 295.
424 <https://doi.org/10.3389/fpsyg.2015.00295>

425 Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: A practical method for
426 grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3),
427 189–198. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6)

428 Fuhs, M. W., & McNeil, N. M. (2013). ANS acuity and mathematics ability in preschoolers from low-
429 income homes: contributions of inhibitory control. *Developmental Science*, 16(1), 136–148.
430 <https://doi.org/10.1111/desc.12013>

431 Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: from reals to integers. *Trends in*
432 *Cognitive Sciences*, 4(2), 59–65. [https://doi.org/10.1016/S1364-6613\(99\)01424-2](https://doi.org/10.1016/S1364-6613(99)01424-2)

433 Gandini, D., Lemaire, P., & Dufau, S. (2008). Older and younger adults’ strategies in approximate
434 quantification. *Acta Psychologica*, 129(1), 175–189.
435 <https://doi.org/10.1016/j.actpsy.2008.05.009>

436 Gebuis, T., & Gevers, W. (2011). Numerosities and space; indeed a cognitive illusion! A reply to de
437 Hevia and Spelke (2009). *Cognition*, 121(2), 248–252.

438 Gebuis, T., & Reynvoet, B. (2011). Generating nonsymbolic number stimuli. *Behavior Research*
439 *Methods*, 43(4), 981–986. <https://doi.org/10.3758/s13428-011-0097-5>

440 Gebuis, T., & Reynvoet, B. (2012a). Continuous visual properties explain neural responses to
441 nonsymbolic number. *Psychophysiology*, 49(11), 1649–1659.

442 Gebuis, T., & Reynvoet, B. (2012b). The interplay between nonsymbolic number and its continuous
443 visual properties. *Journal of Experimental Psychology: General*, 141(4), 642–648.
444 <https://doi.org/10.1037/a0026218>

445 Gebuis, T., & Reynvoet, B. (2012c). The Role of Visual Information in Numerosity Estimation. *PLoS*
446 *ONE*, 7(5), e37426. <https://doi.org/10.1371/journal.pone.0037426>

447 Gebuis, T., & van der Smagt, M. J. (2011). False Approximations of the Approximate Number System?
448 *PLoS ONE*, 6(10), e25405. <https://doi.org/10.1371/journal.pone.0025405>

449 Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., ... Inglis, M. (2013). Individual
450 Differences in Inhibitory Control, Not Non-Verbal Number Acuity, Correlate with
451 Mathematics Achievement. *PLoS ONE*, 8(6). <https://doi.org/10.1371/journal.pone.0067374>

452 Gilmore, C., Cragg, L., Hogan, G., & Inglis, M. (2016). Congruency effects in dot comparison tasks:
453 convex hull is more important than dot area. *Journal of Cognitive Psychology*, 28(8), 923–
454 931. <https://doi.org/10.1080/20445911.2016.1221828>

455 Graham, R. L. (1972). An efficient algorithm for determining the convex hull of a finite planar set.
456 *Information Processing Letters*, 1(4), 132–133. [https://doi.org/10.1016/0020-](https://doi.org/10.1016/0020-0190(72)90045-2)
457 [0190\(72\)90045-2](https://doi.org/10.1016/0020-0190(72)90045-2)

458 Gray, S. A., & Reeve, R. A. (2014). Preschoolers' Dot Enumeration Abilities Are Markers of Their
459 Arithmetic Competence. *PLoS ONE*, 9(4), e94428.
460 <https://doi.org/10.1371/journal.pone.0094428>

461 Guillaume, M., Gevers, W., & Content, A. (2015). Assessing the Approximate Number System: no
462 relation between numerical comparison and estimation tasks. *Psychological Research*, 80(2),
463 248–258. <https://doi.org/10.1007/s00426-015-0657-x>

464 Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the
465 lifespan as revealed by a massive Internet-based sample. *Proceedings of the National*
466 *Academy of Sciences*, 109(28), 11116–11120. <https://doi.org/10.1073/pnas.1200196109>

467 Halberda, J., Mazocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal
468 number acuity correlate with maths achievement. *Nature*, 455(7213), 665–668.
469 <https://doi.org/10.1038/nature07246>

470 Hasher, L., & Zacks, R. T. (1988). Working Memory, Comprehension, and Aging: A Review and a New
471 View. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 22, pp. 193–225).

472 Academic Press. Retrieved from
473 <http://www.sciencedirect.com/science/article/pii/S0079742108600419>
474 Hurewitz, F., Gelman, R., & Schnitzer, B. (2006). Sometimes area counts more than number.
475 *Proceedings of the National Academy of Sciences*, *103*(51), 19599–19604.
476 <https://doi.org/10.1073/pnas.0609485103>
477 Inglis, M., & Gilmore, C. (2013). Sampling from the mental number line: How are approximate
478 number system representations formed? *Cognition*, *129*(1), 63–69.
479 <https://doi.org/10.1016/j.cognition.2013.06.003>
480 Inglis, M., & Gilmore, C. (2014). Indexing the approximate number system. *Acta Psychologica*, *145*,
481 147–155. <https://doi.org/10.1016/j.actpsy.2013.11.009>
482 Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers.
483 *Proceedings of the National Academy of Sciences*, *106*(25), 10382–10385.
484 <https://doi.org/10.1073/pnas.0812142106>
485 Keller, L., & Libertus, M. (2015). Inhibitory control may not explain the link between approximation
486 and math abilities in kindergarteners from middle class families. *Frontiers in Psychology*, *6*.
487 <https://doi.org/10.3389/fpsyg.2015.00685>
488 Kramer, A. F., Humphrey, D. G., Larish, J. F., & Logan, G. D. (1994). Aging and inhibition: Beyond a
489 unitary view of inhibitory processing in attention. *Psychology and Aging*, *9*(4), 491–512.
490 <https://doi.org/10.1037/0882-7974.9.4.491>
491 Leibovich, T., & Henik, A. (2013). Magnitude processing in non-symbolic stimuli. *Frontiers in*
492 *Psychology*, *4*. <https://doi.org/10.3389/fpsyg.2013.00375>
Leibovich, T., & Henik, A. (2014). Comparing performance in discrete and continuous comparison
tasks. *The Quarterly Journal of Experimental Psychology*, *67*(5), 899-917.
493 Lemaire, P., & Lecacheur, M. (2007). Aging and Numerosity Estimation. *The Journals of Gerontology*
494 *Series B: Psychological Sciences and Social Sciences*, *62*(6), P305–P312.

495 Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the
496 frequency of conflicting stimuli in a Stroop-like task. *Memory & cognition*, 7(3), 166-174.

497 Norris, J. E., & Castronovo, J. (2016). Dot Display Affects Approximate Number System Acuity and
498 Relationships with Mathematical Achievement and Inhibitory Control. *PLOS ONE*, 11(5),
499 e0155543. <https://doi.org/10.1371/journal.pone.0155543>

500 Norris, J. E., McGeown, W. J., Guerrini, C., & Castronovo, J. (2015). Aging and the number sense:
501 preserved basic non-symbolic numerical processing and enhanced basic symbolic processing.
502 *Frontiers in Psychology*, 6, 999. <https://doi.org/10.3389/fpsyg.2015.00999>

503 Odic, D., Hock, H., & Halberda, J. (2014). Hysteresis affects approximate number discrimination in
504 young children. *Journal of Experimental Psychology: General*, 143(1), 255–265.
505 <https://doi.org/10.1037/a0030825>

506 Odic, D., Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Developmental change in the acuity of
507 approximate number and area representations. *Developmental Psychology*, 49(6), 1103–
508 1112. <https://doi.org/10.1037/a0029472>

509 Piazza, M., & Izard, V. (2009). How Humans Count: Numerosity and the Parietal Cortex. *The*
510 *Neuroscientist*, 15(3), 261–273. <https://doi.org/10.1177/1073858409333073>

511 Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning Curves for Approximate
512 Numerosity in the Human Intraparietal Sulcus. *Neuron*, 44(3), 547–555.
513 <https://doi.org/10.1016/j.neuron.2004.10.014>

514 Rousselle, L., & Noël, M.-P. (2008). The development of automatic numerosity processing in
515 preschoolers: Evidence for numerosity-perceptual interference. *Developmental Psychology*,
516 44(2), 544–560. <https://doi.org/10.1037/0012-1649.44.2.544>

517 Spooner, J. W., Sakala, S. M., & Baloh, R. W. (1980). Effect of aging on eye tracking. *Archives of*
518 *Neurology*, 37(9), 575–576. <https://doi.org/10.1001/archneur.1980.00500580071012>

519 Szűcs, D., Nobes, A., Devine, A., Gabriel, F. C., & Gebuis, T. (2013). Visual stimulus parameters
520 seriously compromise the measurement of approximate number system acuity and

521 comparative effects between adults and children. *Frontiers in Psychology*, 4.
522 <https://doi.org/10.3389/fpsyg.2013.00444>

523 Warabi, T., Kase, M., & Kato, T. (1984). Effect of aging on the accuracy of visually guided saccadic eye
524 movement. *Annals of Neurology*, 16(4), 449–454. <https://doi.org/10.1002/ana.410160405>

525 West, R., & Baylis, G. C. (1998). Effects of increased response dominance and contextual
526 disintegration on the Stroop interference effect in older adults. *Psychology and aging*, 13(2),
527 206.

528 Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982).
529 Development and validation of a geriatric depression screening scale: A preliminary report.
530 *Journal of Psychiatric Research*, 17(1), 37–49. [https://doi.org/10.1016/0022-3956\(82\)90033-](https://doi.org/10.1016/0022-3956(82)90033-4)
531 4

532

533

534

535

536

537

538 **Figure headings**

539 *Figure 1: Scatter plots showing the relationship between trials' numerosity ratios, mean accuracies*
540 *(%), cumulative surface area ratios, and convex hull ratios. Ratios are plotted on logarithmic axes.*
541 *Ratios are calculated left array / right array so that all ratios vary below and above 1. Incongruent*
542 *trials are those for which the numerosity ratio is below 1 and the visual cue ratio(s) are above 1 or*
543 *vice-versa*

544 *Figure 2: Participants' accuracy on dot comparison trials showing an interaction between cumulative*
545 *surface area and convex hull congruency (error bars show standard error)*

546 *Figure 3: The impact of convex hull congruency on younger and older adults' performances (error*
547 *bars show standard error)*

548

549

550

551

552

553

554

555

556

557

558

559

560 **Disclosure of interest**

561 The authors report no conflicts of interest.