

1 The impact of involving young children in improving 2 the understanding of a new eye test

3 Running title

4 Impact of children's involvement in research

5 Authors

6 Therese Casanova¹, Carla Black, Sheima Rafiq², Jessica Hugill³, Jenny C A Read, Kathleen Vancleef⁴

7 Corresponding author: Dr Kathleen Vancleef, Phone: +44 (0)1865 281388, Email:

8 Kathleen.vancleef@psy.ox.ac.uk

9 Affiliations at time of work

10 Newcastle University, Institute of Neuroscience, Henry Wellcome Building, Framlington Place, NE2
11 4HH, Newcastle-upon-Tyne, United Kingdom

12 Acknowledgements

13 We would like to thank Adam O'Neill for administrative support, the orthoptists from Newcastle Eye
14 Clinic for their feedback and assistance in data collection, all children and parents for taking part, our
15 patient panel for their advice at various stages of the research, and all staff of schools, nurseries, and
16 museums in Newcastle-upon-Tyne for hosting and supporting us.

17 Conflicts of interests

18 Jenny CA Read is a consultant for Magic Leap. The ASTEROID vision test has recently been licensed to
19 a company. The other authors declare no other conflicts of interest.

20 Funding details

21 This manuscript presents independent research commissioned by the Health Innovation Challenge Fund
22 (HICF-R8-442, WT102565/z/13/z), a parallel funding partnership between the Wellcome Trust and the
23 Department of Health. The views expressed in this manuscript are those of the authors and not
24 necessarily those of the Wellcome Trust or the Department of Health.

25 Keywords

26 patient and public involvement, stereotest, children, test development, PPI, med tech, stereopsis,
27 engagement, co-production

28

29

¹ NHS Business Services Authority, Stella House, Goldcrest Way, Newcastle, NE5 8NY

² Pennine Care NHS Foundation Trust, Children's Acute and Ongoing Needs Service, Callaghan House, Cross Street, OL10 2DY, Heywood, United Kingdom

³ Durham University, School of Education, DH1 1TA, Durham, United Kingdom

⁴ NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
University of Oxford, Department of Experimental Psychology, Anna Watts Building, Radcliffe Observatory Quarter, Woodstock Road, University of Oxford, OX2 6GG, United Kingdom

30 **Abstract**

31

32 **Background.** Although considered important, the involvement of young children in research is rare
33 and its impact has never been measured. We aim to provide a good practice example of involvement
34 of young children in the development of a tablet-based 3D eye test or stereotest, ASTEROID. We
35 focus on improving understanding of test instructions.

36 **Methods.** After a pre-measure of understanding was taken, we explored issues with the test
37 instructions in patient and public involvement (PPI) sessions. Feedback was collected via
38 observations, rating scales and children's comments. An interdisciplinary panel reviewed the
39 feedback and decided on the implementation. Subsequently, a post-measure was collected. In Study
40 1, 650 children (2-11.8 years old) took part in the pre-measure, 111 children (1-12 years old) in the
41 PPI sessions, and 52 children (4-6 years old) in the post-measure which also served as a pre-measure
42 for Study 2. PPI sessions for Study 2 attracted 122 children (1-12 years old) and adults. 53 people (2-
43 39 years old) participated in the post-measure of Study 2.

44 **Results.** Following feedback in Study 1, we added a frame cue and included a shuffle animation. This
45 increased the percentage of correct practice trials from 76% to 97% ($t(231)=14.29$, $p<.001$). After
46 adding a cardboard demo in Study 2, the percentage of correct trials remained stable but the
47 number of additional instructions given decreased ($t(103)=3.72$, $p<.001$).

48 **Conclusions.** Our study demonstrates impact of involvement of very young children in research
49 through accessible activities. It can encourage researchers to involve young children and contribute
50 to developing guidelines.

51

52 Introduction

53 The UK National Advisory Group promoting public involvement in health and social care research,
54 INVOLVE, defines Patient and Public Involvement (PPI) as research which is ..."carried out 'with' or
55 'by' members of the public rather than 'to', 'about' or 'for' them¹. This does not include people in
56 their role as research subjects or participants in a study, but rather as advisers or collaborators in the
57 research². PPI continues to be of growing importance for research across disciplines. Involvement of
58 children in research is encouraged in the guiding principles of the United Nation's convention on the
59 Right of the Child. Article 12 acknowledges that children and young people have the right to be
60 involved in decisions that affect their life, express their views and have their views listened to³.
61 Involving children in research from design to dissemination is therefore no longer a preferred
62 approach but a requirement⁴.

63 Traditionally, children's perspectives have been filtered through interpretations of parents and
64 carers rather than children being involved themselves with their unique insights into their own
65 reality⁵. Despite the change in policy and vision, examples of involvement of children in research are
66 very limited, especially across the range of healthcare provision^{4,6,7}.

67 For instance, on 28th April 2019, the INVOLVE evidence library⁸ included 321 works of involvement
68 of people in designing research in the health sector, 19 of which relate to children's research and
69 PPI. Of these 19, only seven describe original research that involves children, and none outline
70 involvement of children under 6 years old. Dunn and colleagues⁴ concluded that despite many
71 "efforts to include children's voices, translation into research and pedagogical practice is still
72 evolving".

73 Whilst securing involvement from adults may be more straightforward, involving children in research
74 is challenging. There is the intrinsic issue of generational divide and power differences between
75 researchers and children they endeavour to work with⁴. Specifically, younger children are arguably
76 at more risk of difficulty with understanding questions that ask for their feedback. The difficulty of
77 choosing age-appropriate involvement methods for young children has been seen as a barrier to

78 involvement⁴ and good practice methodology on how PPI with children should be achieved is
79 lacking². Some guidelines are available^{5,9} but they only advise on running discussion groups with
80 young people, generally above 12 years old. In sum, while toolkits and case studies from national
81 bodies provide starting points from which research groups could *promote* PPI practice, a large
82 evidence base containing research that implements this does not yet exist.
83 Furthermore, no evidence is available on the impact of PPI with young children upon research to
84 drive the development of standards of good practice⁶. Taken together, there is an urgent need for
85 evidence on *how* young children can be involved in research and *what* impact can be achieved.

86 Aims

87 The current paper describes how children were involved in improving children's understanding of
88 ASTEROID, a new 3D eye test, after initial prototype development. The purpose is to highlight the
89 importance and impact of involving young children in the development of medical technology. Most
90 crucially, this paper aims to contribute to the slowly growing pool of research that involves children
91 and young people and provide evidence of its impact on research. Such an evidence base is the all-
92 important next step to ensure that Public and Patient Involvement becomes a fundamental part of
93 all kinds of research.

94 General methods

95 The ASTEROID study

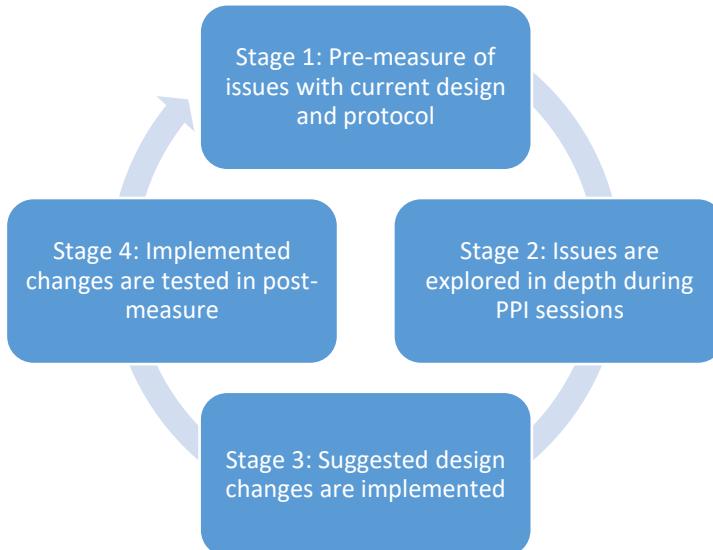
96 Stereopsis, stereoscopic or simply stereo vision all refer to the perception of depth via binocular
97 vision, or vision using both eyes. Normal stereoscopic vision is associated with correct development
98 of visual functions and alignment of the eyes^{10–12}. Measuring stereopsis is therefore common in
99 children with suspected amblyopia or strabismus. However, current ways of measuring stereopsis
100 can be unengaging and under-sensitive (e.g. Huynh, Ojaimi, Robaei, Rose, & Mitchell, 2005). To
101 address this, a new test, known as “Accurate STERotest On a mobile Device” or ASTEROID, was
102 developed on a glasses-free 3D tablet (see Figure 1 and Study 1 Methods for a detailed description,
103 see also¹⁴). The 3D tablet eradicates the need for children to wear 3D glasses during testing,

104 something that has proven to be an issue with younger children with the currently available
105 stereotests¹⁵. To engage children and young people, ASTEROID took the form of an animated game.
106 ASTEROID was validated against a gold standard stereotest and normative data were collected.
107 Anecdotal evidence and feedback gathered during data collection highlighted some issues. Key to
108 this is ensuring that each participant understood what “seeing the 3D image” would be like to
109 ensure that any failures were genuinely due to problems with stereo vision rather than
110 understanding test instructions. These initial outcomes triggered an iterative PPI approach.

111 **Design of the studies**

112 An iterative approach was undertaken to improve ASTEROID (see Figure 1). Each Study started with
113 a data collection stage in which children completed ASTEROID in the context of our larger validation
114 study and we observed an issue related to poor understanding of ASTEROID (pre-measure). In the
115 second stage, this issue was then explored in depth during patient and public involvement (PPI) with
116 a different group of children or adults. By collecting feedback from participants with no previous
117 experience of ASTEROID, we were able to avoid learning effects. PPI outcome triggered changes to
118 ASTEROID (Stage 3). Subsequently, the new version of ASTEROID was evaluated in a separate sample
119 (post-measure, Stage 4). This enabled us to collect quantitative evidence of the impact of our PPI.
120 Verbal feedback and observations during Stage 4 highlighted other issues and subsequently
121 triggered another cycle. In this paper we describe two such cycles that aimed to improve
122 understanding of ASTEROID (Study 1 and 2). Our secondary aim was to monitor how engagement
123 levels changed following improvement of ASTEROID (Study 3).

124



125

126 Figure 1. Schematic representation of how Patient and Public Involvement was embedded in the development
127 of ASTEROID.

128 Study 1: Practice trials and trial-to-trial transition

129 Stage 1: Pre-measure

130 *Participants.* 650 children between 2.1 and 11.8 years old (median age = 6.2, IQR = 3.1, 318 boys and
131 332 girls) participated in Stage 1. They completed the ASTEROID stereotest in the context of a larger
132 validation study with other vision tests. All parents received an information leaflet about the study
133 and an opt-out consent form. If requested by the school or nursery an opt-in consent procedure was
134 used. Children were always asked for oral or non-verbal assent at the time of testing. The study
135 protocol was compliant with the Declaration of Helsinki and was approved by the Ethics Committee
136 of the Newcastle University Faculty of Medical Sciences (approval number 01078). Data were
137 collected in local schools and nurseries.

138 *Instruments.* ASTEROID (version 0.932 and 0.933) is a stereotest that runs on a 3D tablet computer
139 with an autostereoscopic display, meaning no 3D glasses were required for use. The stereotest is
140 embedded in an enjoyable game designed to keep children engaged and responsive. A Bayesian
141 staircase was implemented to adjust the difficulty level of each trial, depending on the user's ability.
142 The test takes the form of four dynamic random-dot stereograms (Figure 2). A disparate target

143 appears randomly in one of these stereograms. The child is verbally instructed to look at each of the
144 four “squares” individually to see which one sticks out. The viewing distance is monitored by the
145 front camera. To ensure that the child understands the task, ASTEROID starts with non-stereo
146 practice trials. In these trials, a colour is added on top of the stereo cue, to make the target clearly
147 visible. This cue is also visible even to colourblind individuals, due to the change in luminance. If the
148 child solves the practice trials correctly, the colour cue gradually fades away over the next few trials
149 and will eventually disappear, leaving only the stereo cue. There are a minimum of 25 stereo trials –
150 the number of non-stereo practice trials increases with the number of mistakes made in the practice
151 trials. Technical details of ASTEROID are described elsewhere ¹⁴.



152

153 Figure 2: Screenshot of a practice trial of ASTEROID (version 0.933). The tests show four dynamic random-dot
154 stereograms. One of the four stereograms has a square with a different disparity and appears to float above
155 the display. In the practice trials this square also has different coloured dots. The colour cue is removed in the
156 disparity-only trials.

157 *Outcome measures.* Our primary outcome measures were (1) the proportion of children who solved
158 less than 80% of the practice trials, (2) the proportion of practice trials that were solved correctly by
159 each child, and (3) the number of additional verbal instructions needed for each child. Performance
160 in the practice trials gives us a good indication of the understanding of the task because even
161 children with vision problems like stereoblindness or colourblindness should be able to solve the
162 practice trials correctly. Secondary, observations of children playing ASTEROID were valuable in
163 generating hypotheses for poor understanding of the task that were further explored in Stage 2.

164 *Results.* 39% of the children showed a poor understanding of the ASTEROID task by solving less than
165 80% of the practice trials correctly. The average proportion of correct practice trials was 0.76 and on
166 average 10 additional prompts were needed per child to explain the task. We particularly observed
167 difficulties in the transition between the practice trials and the stereo trials.

168 Stage 2: PPI sessions

169 *People involved.* 111 children between 2 and 12 years old (median age = 4, IQR = 3, 53 girls, 55 boys,
170 3 gender not recorded) were involved in the PPI sessions. They attended local science and history
171 museums in Newcastle-upon-Tyne (United Kingdom) between October 2016 and January 2017 and
172 were invited to join a PPI session for 5-10 minutes. No ethical approval was required for PPI¹⁶.
173 Children were thanked for their feedback with a 'Junior Scientist' certificate and a sticker.

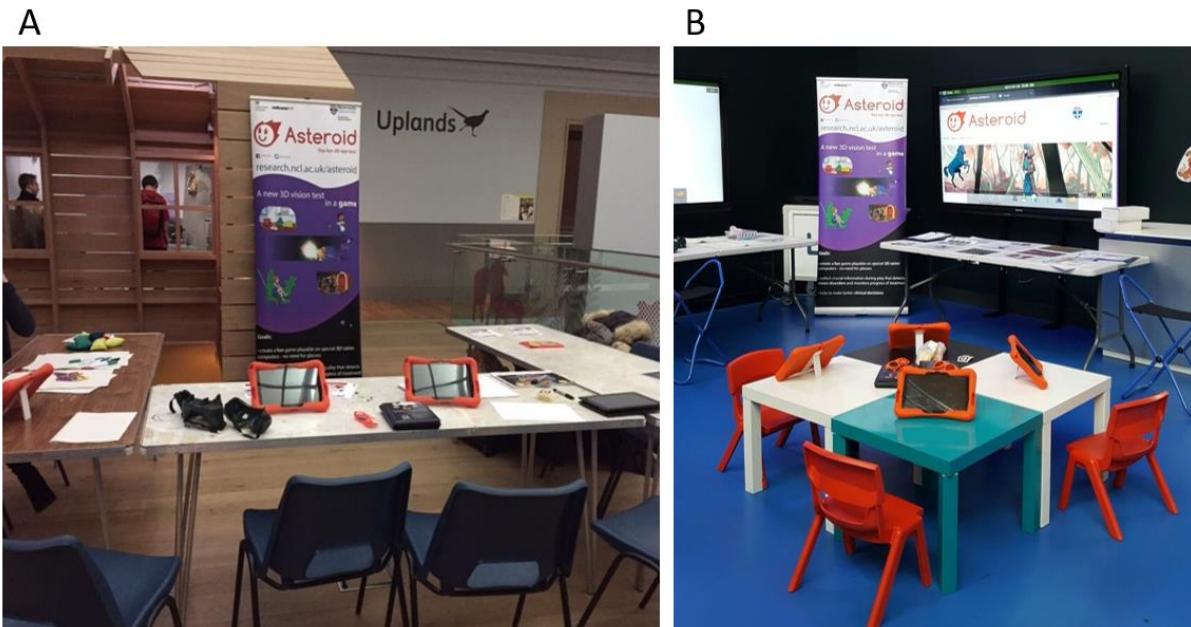
174 *Level and nature of involvement.* We organised three informal drop-in PPI sessions (Table 1). We set-
175 up a stall with engagement activities around vision testing and neuroscience in an area of the
176 museum with high traffic (Figure 3). The stall contained games including visual distortion goggles,
177 bean-bags and examples of visual illusions. The competitive nature of some of the activities engaged
178 children and their siblings, whilst certificates and other small prizes were available as a reward.
179 Crucially, children were invited to try out ASTEROID (version 0.932 and 0.933) and comment on any
180 aspect of the design. Opinions were gathered by research assistants: CB, TC, JH, and SR. CB and SR
181 had experience in engaging children for vision testing in their previous roles as an optometrist and

182 orthoptist respectively. TC has a postgraduate degree in Health Psychology and had worked with in a
183 primary school as a Teacher Training Programme Coordinator. JH worked as a primary school
184 teacher and completed an MSc in Psychology before joining the research team.

185 Table 1. Overview of PPI sessions for Study 1

Session number	Date	Venue	Nature of session	Age range	N	ASTEROID version number	Individual feedback provided?	Engagement rating provided?
1	28/10/2016	Discovery Museum ¹	Drop-in session	3-12	26	0.932	N	Y
2	07/01/2017	Great North Museum: Hancock ¹	Drop-in session	2-11	38	0.933	N	Y
3	24/01/2017	Centre for Life ²	Drop-in session	1-6	47	0.933	Y	Y

186 ¹Local museum with free entrance; ² Local museum with entrance fee; N = number of people involved



187
188 Figure 3. Our set-up for the drop-in PPI activities at Discovery museum (A) and Centre for Life (B)

189 *Outcome measures.*

- 190 1) Observations of the participant's progress (e.g. focus, hesitation, difficulties etc.) were
191 collated on each trial.
192 2) Individual verbal feedback from participants: We asked questions like what they found
193 difficult, what was unclear to them, whether they enjoyed ASTEROID, etc. People were free
194 to give any feedback. This information was collected in Session 3 only.

195 *Outcomes of PPI*. Observations of some children indicated a lack of understanding of the gameplay.
196 For example, for a 2-year old boy it was noted that he “only played a few trials and showed no
197 understanding of what to do”. A similar observation was made for another 2-year old girl: “Happy to
198 play but didn’t understand”, and for a 1.5 year old girl: “Didn’t understand, mum modelled a lot”. In
199 many children we observed hesitation on how to proceed at the end of the non-stereo practice
200 trials, indicating a gap in understanding on what to do during the non-stereo and stereo stages of
201 the game. In addition, comments made during sessions indicated that the colour cue in combination
202 with the colourful dots in the non-stereo trials primed children to look for a local colour change in a
203 few dots instead of an overall depth change in the stereo trials.
204 Second, some children had difficulties understanding that on each trial the locations of the target
205 was randomly determined. In subsequent trials, these children would tap all four stereograms in
206 alternation. For instance, if they incorrectly tapped the top left stereogram in one trial, in the next
207 trial they would not revisit this location but instead try for instance the top right location. If the
208 target was not found there, they would move to yet another position in the third trial. This seemed
209 to indicate that children did not understand that the placement was random, but imagined that the
210 target location remained the same until it was found.

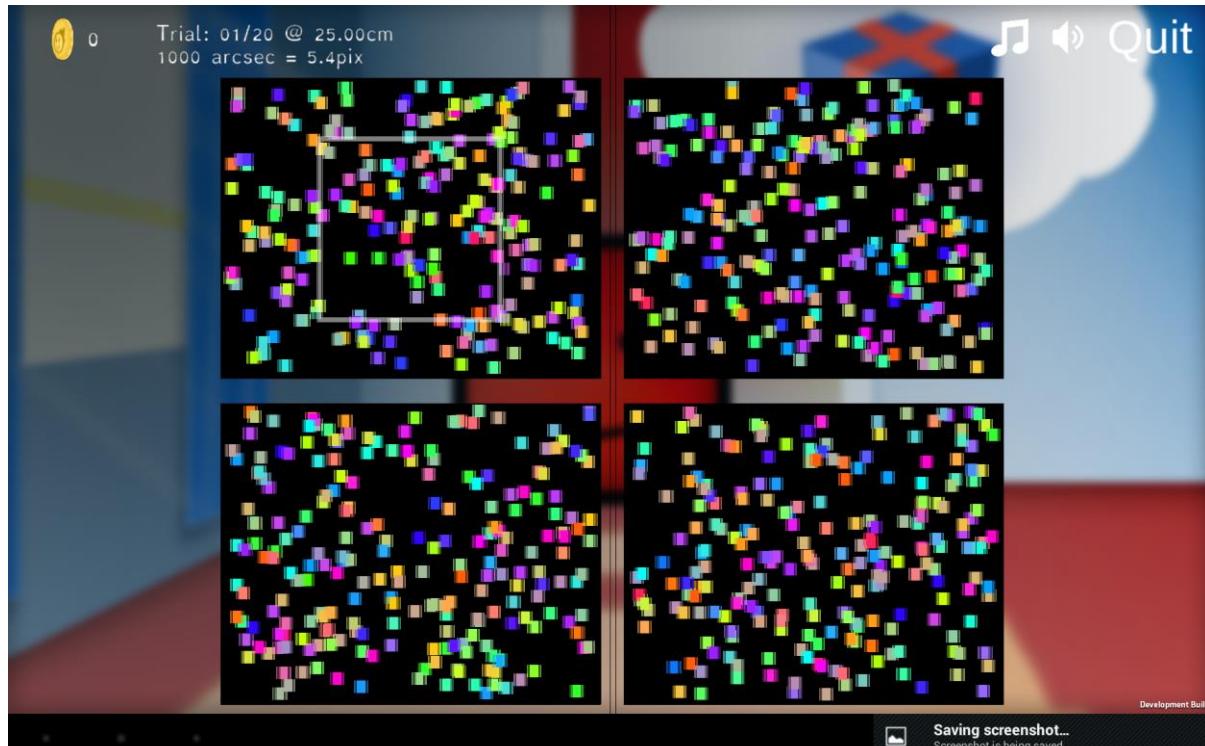
211 Stage 3: Implementing changes

212 *Methods*. Feedback from Stage 2 was summarised by research assistants and discussed at a cross-
213 disciplinary meeting attended by vision scientists, computer scientists and game developers. Once
214 consensus was reached, the proposed changes were implemented in a new version of ASTEROID.

215 *Results*. To solve the first problem, we replaced the colour cue by a frame (Figure 4). The second
216 problem was solved by adding a card shuffle animation. At the end of the trial, the stereograms
217 would flip around, move to the centre of the screen, mix up, move out again to the four corners of
218 the display and flip back face up (see video clip at <https://youtu.be/w8q-4uejwdk>). The animations

219 mimicked playing cards being shuffled and dealt out. We believed this would explain to young
220 children that in the next trial the target could appear in any of the four locations.

221



222



223

224 Figure 4. Screenshots (version 0.937) of the changes made after the first Public and Patient Involvement
225 round: a frame cue (top) and a shuffle animation (bottom, see also video at <https://youtu.be/w8q-4uejwdk>).

226 Stage 4: Post-measure

227 *Participants.* 52 children between 4.5 and 6 years old (median age = 5, IQR = 0.4, 27 boys and 25
228 girls) participated in Stage 4. They completed the ASTEROID stereotest (version 0.938) and other
229 vision tests in the context of a larger validation study.

230 *Procedures.* Ethical procedures, data collection procedures and outcome measures were the same as
231 in Stage 1.

232 *Impact of PPI.* 4% of the children showed a poor understanding of the ASTEROID task (defined as less
233 than 80% correct on the practice trials) after we implemented the shuffle animation and the frame
234 cue in the practice trials, compared to 39% before the changes were made. The proportion of
235 correct practice trials significantly increased from an average of 0.76 in the pre-measure to 0.97 in
236 the post-measure (Welch two-sample t-test: $t(231)=14.29$, $p<.001$). However, we observed a similar
237 average number of verbal instructions that needed to be given (mean in both pre and post-measures
238 = 10; Welch two-sample t-test: $t(77)=0.81$, $p = .42$).

239 Study 2: Instructions

240 Stage 1: Pre-measure

241 The post-measure data of Study 1 served as the pre-measure data of Study 2.

242 Stage 2: PPI sessions

243 *People involved.* 122 children and adults were involved in the PPI sessions of Study 2. Children were
244 between 1 and 12 years old (median age = 4, IQR = 4). 37 of them were girls and 24 of them were
245 boys. Age and gender was not recorded for the adults. The children attended local science and
246 history museums in Newcastle-upon-Tyne (United Kingdom) in February 2017 and were invited to
247 join a PPI session for 5-10 minutes. Adults were either recruited via our research volunteer pool and

248 registered to attend a session at Newcastle University or they attended a drop-in session at a public
249 venue (Newcastle University Medical School foyer or local museum). Children were thanked for their
250 feedback with a 'Junior Scientist' certificate and a sticker.

251 *Level and nature of involvement*. The public was involved through eight PPI sessions (Table 2). All
252 people involved were given the opportunity to try out ASTEROID (version 0.938-0.94) in drop-in
253 sessions similar to those described in Study 1. Children under 12 were asked for individual feedback
254 and engagement ratings. Sessions with adults were set up to obtain verbal reports on understanding
255 the ASTEROID stereotest.

256 Table 2. Overview of PPI sessions for Study 2
257

Session number	Date	Venue	Nature of session	Age range	N	ASTEROID version number	Individual feedback provided?	Engagement rating provided?
4	24/02/2017	Discovery Museum ¹	Drop-in session	2-11	27	0.938	Y	Y
5	28/02/2017	Centre for Life ²	Drop-in session	1-4	27	0.937	Y	Y
6	05/05/2017	Newcastle University Medical School Foyer	Drop-in session	Adults	14	0.94	Y	N
7	26/05/2017	Newcastle University, Institute of Neuroscience	Registration required	Adults	8	0.94	Y	N
8	09/06/2017	Centre for Life ²	Drop-in session	Adults	9	0.94	Y	N
9	16/06/2017	Lit&Phil ¹	Drop-in session	Adults	10	0.94	Y	N
10	23/06/2017	Newcastle University, Institute of Neuroscience	Registration required	Adults	10	0.94	Y	N
11	30/06/2017	Newcastle University, Institute of Neuroscience	Registration required	Adults	17	0.94	Y	N

258 ¹Local museum with free entrance; ² Local museum with entrance fee; N = number of people involved
259

Outcome measures.

260 The outcome measures were the same as in Study 1.

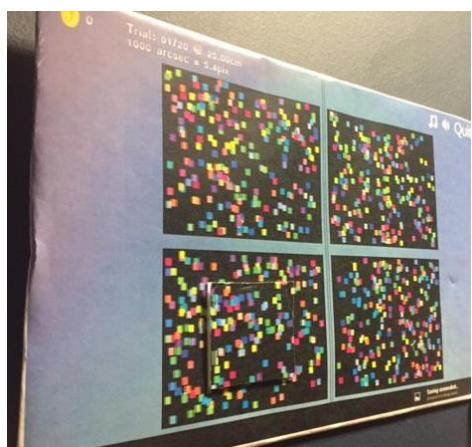
261 *Outcomes of PPI*. Adults described the target as 'a square', 'centre that sort of sticks out', 'popping
262 out and behind'. The outline of the square seems to become more difficult to see at lower disparities
263 when adults described it as 'a circle appears', 'something looks off', 'I can see 3D without square'

264 and even 'going with the gut'. We used these variety of ways to describe the 3D target to the
265 children but observed that children find it difficult to understand any verbal description of a 3D
266 target. However, once they have seen the target in 3D for the first time, they understand what they
267 need to look out for in the subsequent trials.

268 **Stage 3: Implementing changes**

269 *Methods*. Just as in Study 1, feedback from Stage 2 was discussed at a cross-disciplinary meeting and
270 consensus changes were implemented in a new version of ASTEROID.

271 *Results*. We decided to provide the children with a visual and tactile aid to explain them what our 3D
272 target looks like. We therefore made a cardboard demo of an ASTEROID trial. The demo show a print
273 screen of a trial as the background. In the bottom right location, a square with the same dot pattern
274 as the background was glued on top of the background with a 2mm cardboard layer in between
275 (Figure 5).



276

277 Figure 5. Cardboard demo added after the second Public and Patient Involvement round.

278 **Stage 4: Post-measure**

279 *Participants*. 53 participants between 2 and 40 years old (median age = 10.8, IQR = 5.1, 17 boys and
280 36 girls) participated in Stage 4. They completed the ASTEROID stereotest (version 0.940) with
281 additional cardboard instructions in the context of a larger validation study.

282 *Procedures*. Ethical procedures, data collection procedures and outcome measures were the same as
283 in Stage 1.

284 *Impact of PPI*. 6% of the participants showed a poor understanding of the ASTEROID task (defined as
285 less than 80% correct on the practice trials) after we implemented the cardboard instructions
286 compared to 4% without the cardboard instructions. There was no significant increase in average
287 proportion of correct practice trials between pre- and post-measure (mean proportion correct in
288 pre- and post-measure = 0.97; Welch two-sample t-test: $t(92)=-0.02$, $p=.98$). However, we found a
289 significant decrease in the average number of additional verbal instructions given from 10 to 8
290 (Welch two-sample t-test: $t(103)=3.72$, $p<.001$). An age difference between the pre- and post-
291 measure might possibly confound this effect, we therefore rerun our analyses including only subjects
292 younger than 9 years old ($n=22$, median age = 6.3, IQR = 3.4, 8 boys and 14 girls). This indeed
293 resulted in an insignificant difference in the number of additional verbal instructions given (Welch
294 Two sample t-test: $t(31)=0.83$, $p=.41$).

295 Study 3: Engagement with ASTEROID

296 Our primary aim was to increase children's understanding of ASTEROID, however ASTEROID score
297 can only reflect stereo ability if thresholds are not inflated by poor motivation. ASTEROID is
298 therefore embedded in a game and different game themes are available to engage children with
299 different interests. We monitored engagement levels during the development of ASTEROID.

300 Methods

301 Five groups of children and adults were involved in evaluating engagement (Table 3). The groups are
302 described in detail above. Group 4 was a subset of the children and adults involved in Stage 2 of
303 Study 2 (session 4-5). Different outcome measures were collected from different groups (see Table
304 3). We monitored the number of encouragements given by the researchers during game play, such
305 as 'Keep going', 'Well done', 'Just have a go', etc. In other groups, children rated their level of

306 enjoyment on a smiley face rating scale with five levels (Figure 6). Their ratings were converted to
307 scores between 1 (saddest face) and 5 (happiest face).

308 Table 3 Groups of people involved in evaluating engagement with ASTEROID

People involved	N	Described in	ASTEROID version	Outcome measure
Group 1	650	Study 1, Stage 1	0.932 or 0.933	Number of encouragements
Group 2	112	Study 1, Stage 2	0.932 or 0.933	Engagement rating
Group 3	52	Study 1, Stage 4	0.938	Number of encouragements
Group 4	53	Study 2, Stage 2	0.937 or 0.938	Engagement rating
Group 5	53	Study 2, Stage 4	0.940 + cardboard demo	Number of encouragements

309 N = number of people involved



310 Figure 6. Rating scale for engagement with ASTEROID. The scale was accompanied by the question: Did you
311 enjoy playing the Asteroid game?

312 Results

313 ASTEROID was rated as a fun game by the children with an average rating of 4.8 in Study 1 and 4.7 in
314 Study 2. The implementation of the frame and the shuffle animation did not change the rating
315 (Welch two-sample t-test: $t(68) = 0.91$, $p = 0.36$). Although engagement was not the primary aim of
316 the PPI sessions, occasionally positive engagement with ASTEROID was spontaneously noted by the
317 research assistants, for instance ‘Really enjoyed it, laughed at triangle/prize coming up’, ‘Happy to
318 play’, ‘Liked simple [game] but preferred chicken story better’. We found a significant increase of
319 average number of encouragements from 9.6 to 13.8 after we implemented the shuffle animation
320 and the frame cue (Welch two-sample t-test: $t(64)=8.25$, $p<.001$). This indicated that children
321 needed more encouragement and ASTEROID was less engaging after Study 1, possibly because of
322 the additional time occupied by the shuffle animation. However, the average number of
323 encouragements significantly reduced to 8.0 after we added the cardboard demo (Welch two
324 sample t-test: $t(103)=8.32$, $p<.001$).

325 Discussion

326 Our aim was to improve children's understanding of the instructions of a new stereotest that
327 measures 3D vision, ASTEROID. Through two PPI studies, we consulted 233 children and adults, an
328 unprecedented number in studies with children. The insights based on their feedback were
329 extremely valuable for the development team and we made three major changes in subsequent
330 versions of ASTEROID: (1) we replaced the colour cue by a frame cue; (2) we added a shuffle
331 animation to explain that the target could appear in any location on the next trial; (3) we made a
332 visual and tactile cardboard demo to explain to the children what target they were looking for in
333 ASTEROID.

344 The involvement of the public and the changes we made to ASTEROID following their feedback had
345 major impact on the level of understanding of ASTEROID. We noticed a substantial increase in the
346 average proportion of practice trials answered correctly after implementing the frame cue and the
347 shuffle animation. In addition, the percentage of children that showed a poor understanding of
348 ASTEROID (< 80% correct on practice trials) reduced. The impact on the number of additional verbal
349 instructions given to the child during ASTEROID is less straightforward. We observed a slight increase
350 after making the first set of changes, but this decreased again after we included the cardboard
351 demo. This seems to indicate that the cardboard demo was able to replace some of the verbal
352 instructions, while the frame cue and shuffle animation were not. Given the nature of the problem
353 that these changes were trying to solve and the type of verbal instructions given, this is not that
354 surprising. Most verbal instructions were variations on what target to look out for: 'Tap the one that
355 is different', 'Which one is sticking out', 'Do any look like they are popping out'? The cardboard
356 demo was included because children had difficulties understanding any verbal descriptions of the 3D
357 target. So, including a visual and tactile aid removed the need for additional verbal instructions. On
358 the other hand, the frame cue and the shuffle animation were solving rather different problems, not
359 directly related to the verbal instructions given.

350 Our study provides a major contribution to an evidence pool of involvement of children and young
351 people in research. With typical sample sizes of 5-20 children or young people¹⁷⁻¹⁹, our sample of
352 233 people (of which 165 were under 12) is a major step forward in hearing a wide range of
353 opinions. In addition, our study is the first to involve children as young as 1 year old. 30% of our
354 sample are children of 3 years old or younger. Instead of asking parents' opinions as is commonly
355 done with children that age^{2,5}, we describe methods to successfully gain opinions directly from very
356 young children: using smiley face rating scales for engagement level and using observations and
357 simple questions to investigate understanding of the task. For more in-depth knowledge on the
358 reasoning of solving an ASTEROID trial, we relied on verbal reports from adults unfamiliar with 3D
359 computer tablets. A third strength of our study is the quantitative way in which we measured the
360 impact of our PPI through pre- and post-measures, demonstrating a positive and measurable effect
361 of PPI on research.

362 The impact of our PPI was possibly limited by the type of people we involved. By running most of our
363 PPI session in public places and through drop-in sessions, we aimed to lower the barriers for
364 involvement. However, most children involved were visiting a local science or history museum with
365 their parents. Museum visitors do not necessarily reflect general demographics. In addition, one of
366 the museum (Centre for Life) has an entrance fee of £11 for adults and £6.50 for children between 5
367 and 17. This probably caused an underrepresentation of children from lower socio-economic
368 backgrounds. A second limitation of our study is that the median age of participants in our post-
369 measure of Study 2 (median age = 10.8) was higher than in our pre-measure (median age = 5). Our
370 analyses on a subsample of the younger children indeed suggests that increased understanding was
371 mediated by increased age in our sample rather than the cardboard change we have implemented. A
372 larger sample of younger children is needed to confirm this.
373 We learned the following lessons from our PPI sessions that might be useful for other researchers.

374 • The drop-in nature of our sessions required very little commitment of the children and the
375 parents, which lowered the barrier for involvement.
376 • Besides ASTEROID we brought a range of activities and games related to vision and the brain
377 to engage the public, parents and siblings. This certainly helped to attract and engage
378 people.
379 • We set up our stall/table with activities in a central place in the venue with high volume of
380 passage rather than opting for a quiet separate room as is usually preferred for research in
381 public places.
382 • The experience of our team in working with children – both clinically and in schools – was
383 extremely valuable. They were able to make children feel at ease, bring over their
384 enthusiasm and be sensitive to a child's individual needs.

385 In conclusion, children's contributions have measurably impacted on the development of ASTEROID,
386 a new stereotest for children. By increasing accessibility and through creative methods we gathered
387 feedback from 165 children between 1 and 12 years old and 68 older children and adults. The
388 changes implemented following their feedback significantly improved our stereotest. Our approach
389 can be inspirational for future researchers and contribute to an evidence pool of good-practice in
390 involvement of children and young people in research.

391 References

- 392 1. INVOLVE. What is public involvement in research? [https://www.invo.org.uk/find-out-](https://www.invo.org.uk/find-out-more/what-is-public-involvement-in-research-2/)
393 more/what-is-public-involvement-in-research-2/. Published 2019. Accessed April 28, 2019.
- 394 2. Bird D, Culley L, Lakhanpaul M. Why collaborate with children in health research: An analysis
395 of the risks and benefits of collaboration with children. *Arch Dis Child Educ Pract Ed*.
396 2013;98(2):42-48. doi:10.1136/archdischild-2012-303470
- 397 3. United Nations. *The United Nations Convention on the Rights of the Child.*; 1989.
398 https://downloads.unicef.org.uk/wp-content/uploads/2010/05/UNCRC_united_nations_convention_on_the_rights_of_the_child.pdf?_ga=2.204693527.682166437.1557765856-481335225.1557765856. Accessed May 13,
401 2019.
- 402 4. Dunn J. Insiders' perspectives: a children's rights approach to involving children in advising on

- 403 adult-initiated research. *Int J Early Years Educ.* 2015;23(4):394-408.
404 doi:10.1080/09669760.2015.1074558
- 405 5. Shaw C, Brady L, Davey C. *Guidelines for Research with Children and Young People*. London,
406 United Kingdom; 2011.
- 407 6. Dovey-Pearce G, Walker S, Fairgrieve S, Parker M, Rapley T. The burden of proof: The process
408 of involving young people in research. *Heal Expect.* 2019;(January):1-10.
409 doi:10.1111/hex.12870
- 410 7. Collier S. Research essentials. *Nurs Child Young People.* 2014;26(6):12.
411 doi:10.7748/ncyp.26.8.12.s14
- 412 8. INVOLVE. Evidence library. <https://www.invo.org.uk/resource-centre/libraries/evidence-library/>. Published 2019. Accessed April 28, 2019.
- 414 9. Wallace E, Eustace A. Evaluation of Consumer Involvement in the NIHR Clinical Research
415 Network : Children 2013-2014. Overview report. 2014;(258825).
416 <http://www.crn.nihr.ac.uk/resources/evaluation-of-consumer-involvement-in-the-nihr-clinical-research-network-children-2013-2014/?h=9>.
- 418 10. Ciner EB, Ying G-S, Kulp MT, et al. Stereoacuity of preschool children with and without vision
419 disorders. *Optom Vis Sci.* 2014;91(3):351-358. doi:10.1097/OPX.0000000000000165
- 420 11. Elliott S, Shafiq A. Interventions for infantile esotropia. *Cochrane Database Syst Rev.*
421 2013;(1):CD004917. doi:10.1002/14651858.CD004917.pub3
- 422 12. Fricke TR, Siderov J. Stereopsis, stereotests, and their relation to vision screening and clinical
423 practice. *Clin Exp Optom.* 1997;80(5):165-172. doi:10.1111/j.1444-0938.1997.tb04876.x
- 424 13. Huynh SC, Ojaimi E, Robaei D, Rose K, Mitchell P. Accuracy of the Lang II Stereotest in
425 Screening for Binocular Disorders in 6-year-old Children. *Am J Ophthalmol.* 2005;140(6):1130-
426 1132. doi:10.1016/j.ajo.2005.06.032
- 427 14. Vancleef K, Serrano-Pedraza I, Sharp C, et al. ASTEROID: A New Clinical Stereotest on an
428 Autostereo 3D Tablet. *Transl Vis Sci Technol.* 2019;8(1):25. doi:10.1167/tvst.8.1.25
- 429 15. Broadbent H, Westall C. An evaluation of techniques for measuring stereopsis in infants and
430 young children. *Ophthalmic Physiol Opt.* 1990;10(1):3-7.
- 431 16. INVOLVE, The National Research Ethics Service. *Patient and Public Involvement in Research and Research Ethics Committee Review.*; 2009. www.nres.npsa.nhs.uk. Accessed May 15, 2019.
- 434 17. Rankin D, Harden J, Barnard KD, Stephen J, Kumar S, Lawton J. Pre-adolescent children's
435 experiences of receiving diabetes-related support from friends and peers: A qualitative study.
436 *Heal Expect.* 2018;21(5):870-877. doi:10.1111/hex.12802
- 437 18. Gill FJ, Leslie GD, Marshall AP. Barriers and facilitators to implementing a process to enable
438 parent escalation of care for the deteriorating child in hospital. *Heal Expect.* 2018;21(6):1095-
439 1103. doi:10.1111/hex.12806
- 440 19. Hackett CL, Mulvale G, Miatello A. Co-designing for quality: Creating a user-driven tool to
441 improve quality in youth mental health services. *Heal Expect.* 2018;21(6):1013-1023.
442 doi:10.1111/hex.12694
- 443