Convergence and accommodation development is pre-programmed in premature infants

Anna M Horwood\textsuperscript{1,2} PhD
Sonia S Toor\textsuperscript{1} PhD
Patricia M Riddell\textsuperscript{1} DPhil

University of Reading, UK

1. Infant Vision Laboratory, School of Psychology & Clinical Language Sciences, University of Reading, UK
2. Orthoptic Department, Royal Berkshire Hospital, Reading, UK

Word Count manuscript 5232 (excl title, abstract, refs, figs & tables)

Key Words Convergence Accommodation Development Prematurity Strabismus

Address for correspondence & reprints:-
Dr Anna Horwood, PhD, DBO(T)
School of Psychology & Clinical Language Sciences
University of Reading
Earley Gate
Reading
RG6 6AL
UK
a.m.horwood@reading.ac.uk Tel. (+44) 118 378 5553 Fax (+44) 118 378 6715

AH and ST were supported by a UK Medical Research Council Clinician Scientist Award G0802809 to AH
Abstract

Purpose This study investigated whether vergence and accommodation development in pre-term infants is pre-programmed or is driven by experience.

Methods 32 healthy infants, born at mean 34 weeks gestation (range 31.2-36 weeks) were compared with 45 healthy full-term infants (mean 40.0 weeks) over a 6 month period, starting at 4-6 weeks post-natally. Simultaneous accommodation and convergence to a detailed target were measured using a Plusoptix PowerRefII infra-red photorefractor as a target moved between 0.33m and 2m. Stimulus/response gains and responses at 0.33m and 2m were compared by both corrected (gestational) age and chronological (post-natal) age.

Results When compared by their corrected age, pre-term and full-term infants showed few significant differences in vergence and accommodation responses after 6-7 weeks of age. However, when compared by chronological age, pre-term infants’ responses were more variable, with significantly reduced vergence gains, reduced vergence response at 0.33m, reduced accommodation gain, and increased accommodation at 2m, compared to full-term infants between 8-13 weeks after birth.

Conclusions When matched by corrected age, vergence and accommodation in pre-term infants show few differences from full-term infants’ responses. Maturation appears pre-programmed and is not advanced by visual experience. Longer periods of immature visual responses might leave pre-term infants more at risk of development of oculomotor deficits such as strabismus.
**Introduction**

Bifoveal fixation is maintained by the precise coordination of vergence, versions and accommodation to maintain ocular alignment and image clarity. During post natal development, sensory fusion, motor fusion and accommodation become more closely coordinated\(^1-^5\) as visual experience acts on a basic genetic structure. It is unclear, however, whether these systems and relationships are initially pre-programmed and dependent on physical maturation, or influenced by visual experience from the outset. Comparing performance between pre-term and full-term infants provides an opportunity to explore these developmental processes. Figure 1 illustrates the two alternative possibilities\(^6\). If responses are mainly pre-programmed then both full-term and pre-term infants will reach maturity at the same corrected (post-conceptual / gestational) age but the pre-term infants will be older when compared by chronological (post-natal) age. If responses are more experience-dependent then both groups will reach maturity at similar chronological ages, but the pre-term infants will have reached this at an earlier stage of physical maturation (younger corrected age).
Experience Dependent

Figure 1.

Illustration of differences in hypothetical development of mature responses (vergence and accommodation in this case) between full-term and pre-term infants in pre-programmed and experience-dependent scenarios (based on the illustration in Jandó et al. – with publisher’s permission). The maturational hypothesis predicts that full- and pre-term infants’ responses should develop at the same rate when matched by the corrected age (top left), but pre-term infants will be chronologically older when they mature (top right). The experience dependent hypothesis predicts that pre-term infants should develop mature responses before full-term infants when matched by the corrected age (lower left), but at the same chronological age (lower right).

Using this paradigm, previous research suggests that most sensory visual development is mainly pre-programmed and the earlier visual experience resulting from prematurity does not advance most aspects of visual development (for reviews see 7, 8). The effect of prematurity on development of convergence and accommodation during early infancy, has only been described in studies of very small groups, but these also suggest a maturational time course for convergence 9 and accommodation 10.

Importantly for this paper, however, a recent study by Jandó et al. 6, found that the development of the binocular response to dynamic random dot correlograms (DRDCs) in pre-term infants depended on visual experience, not physical maturation. DRDCs are binocular stimuli that only elicit a characteristic visual evoked potential (VEP) in mature binocular systems 11 and are therefore a marker for cortical binocularity in developing infants 12, 13. The same study, however,
found that pattern reversal VEP latency, which is a measure of integrity of the visual pathway, was not advanced by premature birth, so demonstrating that despite an immature visual pathway, the visual cortex can accept environmental stimulation from birth. These results provided a rationale for more detailed exploration of whether the development of convergence and accommodation is maturational or experiential: but there is also clinical relevance.

Children born pre-term are known to have a higher prevalence of accommodative \(^{14, 15}\) and non-accommodative \(^{16-18}\) strabismus. However, what causes this increased prevalence is unclear \(^{19, 20}\). We know that full-term neonates can have periods of ocular misalignment\(^{21}\), inaccurate vergence and accommodation\(^{1, 3}\) and even clinically diagnosed eye muscle palsies\(^{22}\) without any apparent long term harm, but if misalignment persists or increases into the critical period for binocularity, the risk of strabismus, suppression and amblyopia is known to be severe. Tychsen has suggested that decorrelated sensory input between the eyes in the critical period for binocular vision is “a sufficient cause for infantile esotropia”\(^{23}\).

We hypothesized that a mismatch in developmental timing between the sensory and motor components of binocularity could increase the risk of strabismus. If vergence development relates to the corrected age, it would develop later post-delivery in pre-term infants and so these infants would have longer with imprecise vergence and frequent misalignments. If experience-dependent sensory binocularity\(^{5}\), which normally only emerges once vergence is more stable, emerges relatively earlier, immature vergence, which is normally of little consequence, would become a sufficient cause of decorrelated sensory input and be an additional risk factor for the development of strabismus.

This paper describes the development of vergence and accommodation in groups of low-risk pre-term and full-term infants in order to test the experience-dependent vs. maturational hypotheses.
Methods

The study adhered to the tenets of the Declaration of Helsinki and was approved and scrutinised by institutional and UK National Health Service Ethics Committees. Informed consent was obtained from the parents of all infants.

Participants

We defined the corrected age and the chronological age as recommended by the American Academy of Pediatrics Committee on Fetus and Newborn\textsuperscript{24}. The chronological age was defined as the time elapsed from birth, while the corrected age was the chronological age reduced by the number of weeks born before 40 weeks of gestation. The corrected age was calculated from the expected delivery date calculated from the first day of the last menstrual period. 36 pre-term infants born between 31 weeks + 2 days and 36 weeks of gestational age (mean 34.09, SD 1.35 weeks) were recruited from a local maternity hospital. Of these, 32 infants were able to be tested at least once. We chose not to study more premature infants where high rates of retinopathy of prematurity, general health complications, later developmental and perceptual difficulties\textsuperscript{25} might have confounded the data. Three infants were also defined as “small for dates” (low birth weight for their gestational age) and two weighed less than 1500g (1465g and 1361g). None had suffered any perinatal or post-natal neurological complications, all were healthy when tested and none has subsequently developed strabismus and at the time of writing all are at least 2.5yrs old (corrected age).

Reasons for pre-term delivery were mainly twin pregnancy (53%) and pre-eclampsia (15%). We were unable to analyse the twin data separately. Of the many twins, we only collected data from both twins in six pairs, and rarely from both twins at the same visit. Only one set of monozygotic twins were tested.
Pre-term infants were compared with 45 typically developing full-term infants (born between at gestational age 37wks+2days – 42wks+1day: mean 40.0 weeks ±1.6 days), recruited from our departmental Infant Database. Data from these infants contributed to a previous publication, which reported data for the infants on visits when they showed no or minimal (less than +2.0D ) hyperopia. This paper reports some additional from 44 testing sessions in 19 infants (out of a total of 300 sessions) when these infants showed mild hyperopia (up to +3.0D at 16 weeks of age).

All infants were recruited soon after birth. We booked the first test at between 6 weeks corrected age for both groups (because younger infants are rarely testable), although three younger infants were tested in the full-term group, then every two weeks until 20 weeks of age, and finally at 26 weeks of age. Since most aspects of binocular vision develop between 6 and 16 weeks we were not expecting that attempting to collect earlier data would help answer our research question.

**Laboratory testing**

A brief history was taken to confirm normal development and an orthoptic assessment excluded strabismus.

All infants were tested with a remote haploscopic photorefractor described previously (see Supplementary file). It incorporates a Plusoptix SO4 photorefractor in PowerRefII mode, which continuously and simultaneously records refraction and eye position at 25Hz, which allows us to calculate accommodation in diopters (D) and vergence in meter angles (MA). The photorefractor is set in a target presentation apparatus consisting of two concave mirrors and a moving monitor. The target appears to move backwards and forwards in front of the observer between distances of 0.25m and 2m (presented in a pseudo-random order of 0.33m (3D and 3MA demand), 2m (0.5D and MA), 0.25m (4D and MA), 1m (1D and MA), 0.5m (2D and MA). Meter
Angles are a preferable measure of vergence as they are a constant measure of response in relation to demand in populations where IPD varies between participants, and over the course of development. Thus for example, our 0.5m target presented to an infant with an IPD of 45mm would demand 2MA, 13.5 prism diopters or 7.68 degrees of convergence, while for an adult with an IPD of 60mm the same target would still demand 2MA, but 18 prism diopters or 10.2 degrees of convergence. MAs also provide an easy comparison between the appropriateness of vergence and accommodation for target demand at each distance. Data from the 0.25m target were not analysed for three reasons. Most commonly and importantly we find an unacceptable loss of data resulting from small pupils at this distance. There is also a small astigmatic error due to the mirror offsets (of subjectively approximately 0.5D at 25cm) but which reduces below 0.25D and is therefore not problematic at the other distances. Thirdly, the fusional stimulus is slightly different at 25cm because the far edges of the target screen fall slightly beyond the binocular fusional overlap of the lower mirror which is seen in physiological diplopia. We retain the target in the testing order so that a farther target always precedes a nearer one and vice versa.

Vergence and accommodation responses were measured while the infant watched a binocular, cartoon clown target containing a range of spatial frequencies as it moved backwards and forwards. Some target details were only separated by one pixel (visual angle of approximately 1 min arc at 0.33m) but it also contained large elements, high contrast edges, bright colours, alternating elements, eyes and a hairline to be maximally interesting to neonates with poorer visual acuity. The target subtended 3.15° at 2m and 18.3° at 0.33m. If possible each child was tested twice in each session and the data were averaged. The Plusoptix monitor allowed the tester to watch the infant in real time to assess attention and fixation and also to follow recording traces even when the accommodation responses exceeded the operating range of the photorefractor. We only report data collected when the infant was observed to have fixated the
target steadily for at least 2 seconds at each fixation distance. The Plusoptix SO4 has a linear operating range of -7.0/+5.0D (i.e. up to 7D of accommodation and 5D of hyperopia). Beyond this, our unpublished calibrations and those of others\textsuperscript{29} demonstrate that although the photorefractor continues to calculate a figure for refraction, this is an underestimation of the true value. This varies between individuals, so without individual calibration is not precisely quantifiable. Data from infants who demonstrated hyperopic refractive error over +5.0D estimated using maximum hyperopic refraction found during testing (MHR) were excluded before quantitative analysis. We have reported that MHR correlates closely with cycloplegic refraction in other child and infant groups\textsuperscript{30}.

Raw data were processed offline\textsuperscript{28}. Vergence in MA was calculated from the horizontal eye position of each eye, correcting for individually calculated angle lambda and inter-pupillary distance. Individual refraction calibrations and repeatability calculations were not possible for such young infants, but for group comparison studies such as this, averaged data is acceptable\textsuperscript{29}. We calculated accommodation in diopters, using the increasingly myopic photorefraction which occurs on accommodation, with a correction for a slight systematic error (the photorefractor underestimates accommodative response by approximately 0.5D) using a formula derived from group calibration studies\textsuperscript{28} using young adults. Calculations of response gain in relation to target demand (the slope of the stimulus response functions) used at least three data points (four if possible) at the different fixation distances. Where we report responses to particular targets, we have limited them to the nearest (0.33m, 3 MA & D) and the furthest (2 m, 0.5 MA & D).

Statistical Analysis and Data Presentation

We present our results in two ways. Firstly we provide descriptive figures to indicate the spread of responses. Since accommodation responses beyond the linear operating range of the
photorefractor are likely to underestimate the degree of refraction to an unknown extent, this full
dataset was not analysed statistically. If we had excluded these data completely, however, we
felt we would have misrepresented the spread of infant behaviour.

We then calculated group means and 95% confidence intervals (CI) of all data within range.
These data were analysed using two-way between-groups ANOVA (with age group and pre-
term/full-term as factors), to investigate between-group differences in vergence and
accommodation responses and gains at intervals of two weeks. A main effect of age indicates
that vergence and/or accommodation change with age and a main effect of group indicates
overall differences between pre-term and full-term infants. Most importantly, any age x group
interaction would suggest that the two groups differ only at certain ages. If more between-group
differences in responses are found when groups are compared by their corrected age, this
would indicate that development of vergence and/or accommodation is experience-dependent.
More group differences when groups are compared by their chronological age would suggest
development is more maturational.

Post hoc testing used Bonferroni correction for multiple comparisons where appropriate.

Results

Testability and Repeatability

Numbers testable at each age point for both the corrected age and chronological age are
illustrated in Table 1. While most infants provided usable data on most visits, only 4 pre-term
and 13 full-term infants provided such data at every visit, so data were treated as cross-
sectional. Of the maximum potential number of testing sessions over the study period, 55% of
the pre-term infants and 18% of in the full-term infants either were unable to attend or were not
able to be tested at all due to being asleep or fretful on a booked session.
<table>
<thead>
<tr>
<th>Age at testing</th>
<th>4-5 wks</th>
<th>6-7 wks</th>
<th>8-9 wks</th>
<th>10-11 wks</th>
<th>12-13 wks</th>
<th>14-15 wks</th>
<th>16-17 wks</th>
<th>18-19 wks</th>
<th>24-27 wks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FULL-TERM</strong></td>
<td>1*</td>
<td>31</td>
<td>36</td>
<td>37</td>
<td>33</td>
<td>31</td>
<td>29</td>
<td>31</td>
<td>36</td>
</tr>
<tr>
<td>Total tested</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(of 45 in study)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hyperopic</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>session</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>excluded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unrecordable</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>e.g. pupils/lids</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>point excluded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accom out of</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>range (&gt;7D)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>point excluded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% datapoints</td>
<td>0%</td>
<td>4.0%</td>
<td>5.5%</td>
<td>2.0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0.6%</td>
<td>0%</td>
</tr>
<tr>
<td>excluded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PRE-TERM</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(of 32 in study)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total tested</td>
<td>16</td>
<td>24</td>
<td>22</td>
<td>19</td>
<td>22</td>
<td>16</td>
<td>4</td>
<td>7</td>
<td>24</td>
</tr>
<tr>
<td>Hyperopic</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>session</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>excluded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unrecordable</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>e.g. pupils/lids</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>point excluded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accom out of</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>range (&gt;7D)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>point excluded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% datapoints</td>
<td>7.8%</td>
<td>5.2%</td>
<td>4.5%</td>
<td>1.3%</td>
<td>3.4%</td>
<td>3.1%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>excluded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Chronological</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total tested</td>
<td>3</td>
<td>17</td>
<td>24</td>
<td>16</td>
<td>23</td>
<td>16</td>
<td>27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hyperopic</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>session</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>excluded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unrecordable</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e.g. pupils/lids</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>point excluded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accom out of</td>
<td>0</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>range (&gt;7D)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>point excluded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% datapoints</td>
<td>8.3%</td>
<td>8.8%</td>
<td>6.2%</td>
<td>6.2%</td>
<td>2.1%</td>
<td>1.5%</td>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>excluded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Numbers testable at each age point. Pre-term infants were delivered on average six weeks early. At 8-9 weeks chronological age a pre-term infant would be equivalent developmentally to a 2-3 week full-term.
Premature infants, particularly the large number of twins, were especially difficult to test regularly. These factors added to the normal difficulties of testing infants. But if an infant attended and was attentive, complete runs of targets at the different fixation distances were always recorded. Repeated measurements within a single visit were more often possible for older infants, whether full term or pre-term (e.g. 23% repeatable at 6-7 weeks and 58% at 12-13 weeks of corrected age for the pre-term infants). Repeated measurements were averaged where available. Variability in repeated measurements within individuals was similar to that between different infants at each corrected age point (95% confident intervals were not significantly different), but younger infants were much more variable overall (95%CI for vergence gain at 6-7 weeks: between individuals = +0.12; within an individual = +0.09; while at 12-13 weeks: between individuals = +0.045; within an individual = +0.04).

**Exclusions and Refraction**

Myopia did not exceed -0.5D for any infant tested. Some of the youngest infants behaved myopically (over accommodated) for distance fixation. However, their accommodation relaxed at least once during testing to an emmetropic or hyperopic refraction, confirming that they were not genuinely myopic.

One pre-term infant appeared consistently significantly more than 5.0D hyperopic on multiple visits and their data were excluded completely from further analysis. 2 (6.2%) premature infants, and 4 (8.8%) full-term infants showed >5.0D hyperopia (beyond the linear operating range of the photorefractor) fleetingly (i.e. for a single data point) at some time, all in the first 12 weeks of life and the data from that single session were excluded (Table 1). No refraction from these infants ever exceeded a photorefractor calculation of +7.0D hyperopia. No infant whose session...
data were excluded showed evidence of manifest refraction >+3.00D by 16 weeks of age, so all had emmetropized to within normal limits.

The proportion of infants with hyperopia greater than +2.0D in each group were similar across time when compared by their corrected age e.g. 39% vs 33% respectively at 10-11 weeks and 29% vs 25% at 14-15 weeks. At 24-27 weeks of corrected age the infants’ mean refraction estimated by the MHR measured during the testing session was +0.18D (95%CI -0.25D / +0.66D) in the full-term infants and +0.28D (95%CI -0.43 / +0.99D) in the pre-term infants (t(55)=1.36, p=0.178, n.s.).

Full Dataset

Figure 2 illustrates the ranges of vergence and accommodation responses at two time points, 6-7 weeks of corrected age (which was on average 12-13 weeks of chronological age for the pre-term group), and again at 12-13 weeks of corrected age (18-19 weeks of chronological age for the pre-term infants). We chose these two time points as 6-7 weeks is before mature binocular responses develop in full-term infants, while 12-13 weeks is when vergence and accommodation are not significantly different from adults\(^3\), and sensory binocularity is typically emerging\(^4\).

Figure 2 illustrates the whole dataset including out-of-range accommodation estimates (gray shaded areas). 42 individual datapoints (2.3% of the total tested) exceeded the linear operating range of the phororefractor (>7D accommodation). 24 infants (evenly distributed between pre-term and full-term) provided these datapoints fleetingly for the nearest targets in their first 12 weeks (corrected age if pre-term) and for all except one infant in each group these were between approximately 7.0D and 10.0D. The other two infants contributed six datapoints between approximately 10.0D and 12.0D).
There are two important comparisons in Figure 2. The first is a corrected age match comparison (full-term (top charts) vs pre-term infants (bottom charts)), where performances are similar. Many of the youngest full-term and corrected age pre-term infants (left charts in figure) showed highly erratic accommodation. What we have previously termed “all or nothing” patterns were common, where accommodation response to an approaching target was flat for the more distant targets, but then was either appropriate or excessive (and sometimes out-of-range) for the nearest target, despite concurrent linear vergence. 11 (5.5%) of the 198 individual data points collected at 0.33m in the pre-term infants, and 19 (6.5%) of the 291 points collected in the
full-term infants were greater than 7.0D. Before 12 weeks of age, over-accommodation for the nearest target exceeded 4.5D at 0.33m in 28.5% of full-term infants and 38.5% of the corrected age pre-term infants.

The second comparison is between full-term infants with pre-term infants matched by chronological age. It was not possible to compare full term with pre-term infants at 6-7 weeks since insufficient data was collected from the pre-term infants, but the comparison at 12-13 weeks is illustrated in the top right and bottom left of the figure. This shows that full-term infants’ vergence and accommodation is more linear than chronologically age-matched pre-term infants.

**Analysis of Data in Range**

For statistical analysis we compared infants matched by both their corrected age and chronological age, considering response gain as well as responses for near (0.33m) and distance (2m). Vergence measurements were all within the linear range of the photorefractor across the range tested, so all infants’ vergence gains were calculated using responses at 4 distances. For accommodation, out-of-range points were excluded and gains were calculated from the responses to the three remaining distances. Gains thus calculated are likely to be a slight underestimate of the true gain. Such exclusions occurred most frequently at 8-9 weeks corrected age. Here the median accommodation response for the 0.33m target of the full data set (using out-of-range point which we know are inaccurate) was 0.34D more than the mean of the more selected data. If the median from the full dataset had been used to calculate the gain, it would have increased the gain by 0.12. At other ages differences were less. Four accommodation data points were available for 93% of the target runs for the full-term infants and 90% of those from the pre-term infants.
<table>
<thead>
<tr>
<th></th>
<th>Corrected Age</th>
<th></th>
<th>Chronological Age</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>p</td>
<td>η^2</td>
<td>F</td>
</tr>
<tr>
<td>Vergence Gain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age in weeks</td>
<td>11.68</td>
<td>.000</td>
<td>.207</td>
<td>20.625</td>
</tr>
<tr>
<td>Prem /Term</td>
<td>1.32</td>
<td>.251</td>
<td>.003</td>
<td>5.299</td>
</tr>
<tr>
<td>Age x Prem/Term</td>
<td>4.46</td>
<td>.000</td>
<td>.091</td>
<td>4.819</td>
</tr>
<tr>
<td>Vergence at 2m</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age in weeks</td>
<td>3.36</td>
<td>.000</td>
<td>.070</td>
<td>3.919</td>
</tr>
<tr>
<td>Prem /Term</td>
<td>0.01</td>
<td>.934</td>
<td>.000</td>
<td>3.053</td>
</tr>
<tr>
<td>Age x Prem/Term</td>
<td>1.02</td>
<td>.428</td>
<td>.022</td>
<td>2.108</td>
</tr>
<tr>
<td>Vergence at 0.33m</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age in weeks</td>
<td>14.31</td>
<td>.000</td>
<td>.249</td>
<td>12.785</td>
</tr>
<tr>
<td>Prem /Term</td>
<td>0.39</td>
<td>.533</td>
<td>.001</td>
<td>7.383</td>
</tr>
<tr>
<td>Age x Prem/Term</td>
<td>4.18</td>
<td>.000</td>
<td>.088</td>
<td>5.733</td>
</tr>
<tr>
<td>Accom Gain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age in weeks</td>
<td>2.31</td>
<td>.012</td>
<td>.049</td>
<td>.039</td>
</tr>
<tr>
<td>Prem /Term</td>
<td>2.29</td>
<td>.131</td>
<td>.005</td>
<td>2.397</td>
</tr>
<tr>
<td>Age x Prem/Term</td>
<td>2.73</td>
<td>.003</td>
<td>.057</td>
<td>3.819</td>
</tr>
<tr>
<td>Accom at 2m</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age in weeks</td>
<td>2.33</td>
<td>.011</td>
<td>.050</td>
<td>11.885</td>
</tr>
<tr>
<td>Prem /Term</td>
<td>14.94</td>
<td>.000</td>
<td>.033</td>
<td>1.135</td>
</tr>
<tr>
<td>Age x Prem/Term</td>
<td>1.98</td>
<td>.033</td>
<td>.043</td>
<td>3.933</td>
</tr>
<tr>
<td>Accom at 0.33m</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age in weeks</td>
<td>1.97</td>
<td>.035</td>
<td>.045</td>
<td>11.583</td>
</tr>
<tr>
<td>Prem /Term</td>
<td>29.46</td>
<td>.000</td>
<td>.065</td>
<td>3.105</td>
</tr>
<tr>
<td>Age x Prem/Term</td>
<td>1.67</td>
<td>.086</td>
<td>.038</td>
<td>1.429</td>
</tr>
</tbody>
</table>

Table 2

Results of ANOVA of vergence and accommodation gains and responses at 2m and 0.33m. Significant differences are shaded.
Vergence gain (top), vergence responses to target at 2 meters (center) and vergence responses to target at 0.33m (lower). Left column: responses matched by corrected age. Right column: responses matched by chronological age. Statistically significant differences on post-hoc testing indicated by asterisks. Error bars indicate 95% confidence intervals. * indicates p<0.05; ** indicates p<0.01

Figure 3

Full-Term vs Premature

Full-Term vs Premature
Figure 4

Accommodation gain (top) calculated from at least three fixation distances, and actual responses at 2 meters (center) and 0.33m (lower). Left column: responses matched by corrected age. Right column: responses matched by chronological age. Statistically significant differences on post-hoc testing indicated by asterisks. Error bars indicate 95% confidence intervals. * indicates p<0.05; ** indicates p<0.01
Results of the ANOVAs comparing response gains and responses at 2m and 0.33m between groups are shown in Table 2 and post hoc significant differences are indicated in Figures 3 (vergence) and 4 (accommodation).

Again, we compared groups matched by both corrected and chronological age. When matched by their corrected age there were the expected significant developmental improvements in all infants. Pre-term infants relaxed their accommodation significantly less at 2m than the full-term infants, but there were no other overall group differences. There were significant age x group interactions in four of the six comparisons but post-hoc testing showed that differences were only significant at 6-7 weeks of age (Figures 3 and 4), where the pre-term infants under-converged for near, and over-accommodated for distance targets. Subsequently, up to 24-27 weeks, there were no differences in accommodation and vergence responses between full-term and pre-term infants matched by their corrected age.

When infants were matched by chronological age there were significant pre-term/full-term group differences for all comparisons except accommodation at 2m. Full-term infants showed more appropriate responses than the chronologically age matched pre-term infants (gain closer to 1, responses closer to the target demand). There was also a significant age x group interaction for all comparisons except accommodation at 0.33m. Post hoc testing showed that the majority of significant differences were found between infants aged between 10-16 weeks and were particularly clear at 10-11 weeks of age. While the full-term infants’ responses appeared to have matured (were similar to responses at the oldest age tested), those of the pre-term infants were still immature.

To test the linearity of vergence and accommodative responses for each group we calculated correlation coefficients ($r^2$) for individual stimulus response slopes where four data points (at 0.33m, 0.5m, 1m and 2m) were available. Infants matched by their corrected age demonstrated
similar linearity of response e.g. for vergence at 12-13 weeks mean $r^2$ were 0.94 and 0.91 respectively for full-term and the corrected age pre-term infants. However, when matched by chronological age 12-13 week pre-term infants demonstrated less linear vergence ($r^2 = 0.77$ for pre-term infants and 0.94 for full-term infants) ($t=2.57, p=0.019$), not significantly different from full-term infants at 6-7 weeks. Similar analysis for accommodation showed that mean $r^2$ for the full-term and the corrected age pre-term infants did not differ significantly (0.74 and 0.77 respectively), but pre-term infants of the same chronological age had a lower mean $r^2$ of only 0.53 ($t(39)=2.4, p=0.02$), again not-significantly different from full-term infants at 6-7 weeks.

**Discussion**

This study investigated the developmental time course for vergence and accommodation responses in full-term and pre-term infants matched by both chronological and corrected age. Our results suggest that vergence and accommodation in pre-term infants follow a maturational developmental trajectory and that responses are not accelerated by the additional visual experience of earlier birth. Full-term infants show more adult-like vergence and accommodation responses when compared to chronologically age-matched pre-term infants.

These results contrast with those of Jandó et al\textsuperscript{6} who showed an experience-dependent development of sensory binocularity, where the additional visual experience in preterm infants resulted in earlier development. 50% of Jandó et al’s\textsuperscript{6} pre-term infants responded to DRDCs by 1.92 months post-natally (approximately 8 weeks). If sensory binocularity develops earlier in pre-term infants, but accommodation and vergence responses do not, then early development of sensory binocularity is unlikely to be the cause of maturation of vergence and accommodation. Instead, it is possible that the oculomotor system supports or reinforces the development of sensory binocularity.
Vergence accuracy and a gain close to one characterize adult-like responses. More recent research has demonstrated that, in full-term infants, vergence is adult-like by 8-9 weeks\(^1,3\), earlier than suggested by older literature where such young infants were not assessed\(^3\) or good vergence responses less commonly found\(^4\). The early large neonatal misalignments found in infants younger than 2 months of age are also reducing dramatically\(^21,4,31\). Thus good alignment for targets at all fixation distances is typically in place before the onset of stereopsis and sensory binocularity (Wong A et al. IOVS 2008;49:e-abstract 3748)\(^8,26,32-34\). In contrast, our pre-term infants still showed immature vergence until about 15 weeks of age.

If sensory and oculomotor visual systems had been found to mature in parallel, then the effects of prematurity on visual development would be insignificant as the onset of critical periods for vergence control and sensory binocularity would be similarly delayed. However, if any aspect of sensory binocularity (with concurrent susceptibility to suppression and amblyopia) can be advanced by experience, while oculomotor control is not, a mismatch of developmental trajectories might result in decorrelated input from each eye to the visual cortex at a time when cortical binocularity is entering a critical period that has been advanced through early visual experience.

Additional infant studies have demonstrated that development of stereopsis does not depend on the development of vergence\(^35,4\). Thorn at al\(^4\) suggest that good alignment is not necessary for development of the neural mechanisms underlying binocular vision, but is necessary for maintenance of these mechanisms. Tychsen argues that "binocular decorrelation is a sufficient cause of infantile esotropia when imposed during a critical period of visuomotor development"\(^23\). Immature biases to esodeviation such as asymmetrical monocular OKN\(^27\) and better convergence than divergence\(^26\) may be retained in premature infants, resulting in an increased
risk of infantile esotropia. Our findings therefore suggest a mechanism that might account for increased prevalence of strabismus in pre-term infants.

**Accommodation**

Immature accommodation is more erratic and less linear than vergence at the same age. In pre-term infants, this variability is extended for longer after birth. Lower gain was often the result of over-accommodation in the distance, but excessive accommodation for near was also common, often after almost flat responses to the three farther targets, as has been found in previous studies\(^3,37\). Accommodation development in pre-term infants also related to their corrected age rather than their chronological age, with the same gradual increase in accommodation gains over the first weeks that Banks found for two younger full-term infants using dynamic retinoscopy\(^{10}\). Banks' research also suggested a similar pre-programmed course of development. We did not detect, however, the same clear developmental trajectory of accommodation development in full-term infants as reported by Banks\(^{10}\) because most of our full-term infants were already showing response gains of well over 1.0 (and which related to their refraction) by 6-7 weeks.

Our results suggest that not only are vergence inaccuracies occurring when cortical binocularity could be emerging, but the linkages between vergence and accommodation will be less consistent during this extended period of mismatched retinal input and imprecise accommodation. Although we have reported that mean full-term infant AC/A ratios are not significantly different from those of adults\(^5\), the variability of response in preterm infants would result in a weaker linkage between vergence and accommodation responses for a greater developmental period. Thus, increased risk of strabismus in preterm infants might also be driven by lack of reinforcement of AC/A and CA/C ratio linkages.
Finally, good accommodation is also implicated in emmetropization. Previous studies have shown that binocular input dramatically enhances not only vergence but also accommodation in full-term infants, older children and adults. As well as inaccurate vergence (and so inter-ocular decorrelation) being a “sufficient” cause of esotropia, any damage to cortical binocularity might then also damage accommodation, and thus be implicated in the defective emmetropization that is more common in those born both pre-term and with strabismus.

Thus, prematurity may not only cause infantile esotropia, but might also be implicated in strabismus with an accommodative element.

Study Limitations

While comparisons of these data with those of Jandó et al support our arguments above, there are differences in testing paradigm between the two studies which might explain apparent differences between developmental time courses between the groups for other reasons. Jandó et al measured cortical activity which required no behavioural response. VEP is easier to test successfully in very young infants and VEP testing is a less demanding task than our paradigm.

Our task involves a longer processing time, requires a motor response to a sensory signal, and is more likely to be susceptible to attentional variation. It is therefore possible that the attentional system in premature infants needs to have reached a sufficient level of maturity for them to perform the tests used here. In this case, the difference in timing between full term and preterm infants might be the result of differences in maturation of higher order behavioural mechanisms rather than maturation of vergence and accommodation per se.

All infants, especially pre-term twins, present a significant challenge in testing, so a complete set of longitudinal data was rare, and many testing sessions were abandoned or cancelled for reasons unrelated to the study. However, this is only likely to affect the quantity, not the quality
of the results. Despite small numbers in the youngest infant groups, statistical significance was still reached.

We could not definitively differentiate attentional and physical immaturity, but either means that pre-term infants will have inaccurate vergence and accommodation for longer after birth. Immature responses could be due to immaturity of the control mechanisms, so despite sensory detection of the change of target distance, rapid, co-ordinated physical responses cannot yet occur. Alternatively, acuity, attention or interest in detailed targets may be insufficiently developed to drive appropriate responses. Accommodation is certainly active in very early infancy, as evidenced by the difference between cycloplegic (generally hyperopic) and non-cycloplegic (generally myopic) refraction of neonates (for review see Thorn et al 42), and convergence is also clearly possible during frequent large neonatal misalignments 21, but seems poorly controlled. We also accept that the reduction in variability of responses from the older infants could also partly be due to averaging of more infants’ data, but even the averaged data became less variable with time.

A major limitation of the Plusoptix photorefractor is its relatively small operating range. Although out-of-range accommodation responses were still collected, we could not measure them accurately because calculations from the Plusoptix become non-linear, so a reading of 8D might be the given from an accommodative response of between 7D and 9D, and this error may vary between individuals. By excluding these points our statistical testing used a slightly smaller dataset (and probably under-estimated mean over-accommodation), but the type and proportions of excluded data were similar in each group. We continue to use the Plusoptix photorefractor because it is one of the few instruments able to refract and assess eye position binocularly, naturalistically, simultaneously and continuously.
We considered excluding the very non-linear responses, where a pattern of flat or low gain
responses was found to targets at 0.5m or beyond, with a sudden large over-accommodation
response to the 0.33m target. These responses are different from largely linear adult responses
and were sometimes out of the linear range of the photorefractor. By excluding them, however,
we would miss-describe neonatal responses, of which they are a feature. We accept that when
the excessive near response is out-of-linear-range they are difficult to quantify using our
equipment, but they are of interest for two reasons. Flat accommodation responses for more
distant targets, followed by appropriate or excessive accommodation for near suggest that while
vergence seems generally well controlled over the linear range of target distances,
accommodation can be driven independently once a level of blur (or disparity) reaches a
threshold. These responses also have implications for the development of the AC/A ratio
because they suggest that the relationship between accommodation and vergence is different at
different target demands, suggesting that in infancy A/C linkages are unstable.

We could also not perform the individual calibrations for accommodation that would have been
ideal for such studies\textsuperscript{29}, although group comparisons are often used in studies such as this. The
Plusoptix photorefractor accuracy compares well with refraction derived from retinoscopy
(around +/- 0.75D)\textsuperscript{28, 43}, while our measure of vergence change is more precise because we
correct for variables such as IPD and angle lambda\textsuperscript{28}. There may therefore have been some
individual between-participant differences in accuracy of refraction within the operating range of
the photorefractor, but there should be no optical reasons why calculation of refraction of
younger or premature infants \textit{per se} should be less accurate (once data is captured). The fact
that more linear vergence was demonstrated simultaneously with erratic accommodation shows
the infants were attending to the target and refraction was on-axis, but frequently well outside
ranges which could be attributed to measurement error.
We had too few significantly hyperopic infants to investigate early hyperopia as a separate issue. We had similar proportions of apparently hyperopic infants in each of our groups when matched by their corrected age, so this is unlikely to have affected our results.

In conclusion, vergence and accommodation follow a pre-programmed developmental trajectory so pre-term infants appear to have longer visual experience of immature responses. This may extend into the period when experience-dependent cortical binocularity emerges. A mismatch in the time course between the development of oculomotor and sensory binocularity might contribute to the increased risk of strabismus in children born pre-term.

**Acknowledgements**

Financial conflicts of interest – none. AH and ST were supported by UK Medical Research Council Clinician Scientist Fellowship G0802809. We are very grateful for help and advice in planning and carrying out recruitment of the pre-term infants given by Dr Greg Boden, Consultant Paediatrician, and Research Nurses Sue Hallett and Morag Zelisko at the Royal Berkshire Hospital, Reading, UK
References


Supplemental File

Convergence and accommodation development is pre-programmed in premature infants

Horwood, Toor, Riddell

Supplemental Materials

Details and Validation of Method

Apparatus

The remote haploscopic videorefractor (RHV) presented targets at five fixation distances while collecting continuous recordings of eye position and accommodative response. The method and calibration procedures have been described in detail previously (Horwood & Riddell, 2008).

There are two optical pathways designed so that data collection and target presentation can be separated. The eye position and refraction data are collected continuously from each eye at 25Hz via an infra-red “hot” mirror (E in Figure 1) using a PlusoptiXSO4 PowerRefl photorefractor (Plusoptix GmbH, Nurnberg, Germany). The mirror transmits visible light so that the participant has an unimpeded view of the target, but reflects infra-red so that the camera sensors can be placed in the same optical plane as the target but without obscuring it. The participants view a monitor screen via two concave mirrors arranged so that the virtual image of the monitor is seen to move backwards and forwards directly in front of the participant (Figure 1). The advantage of using these mirrors for the target presentation pathway is that one eye’s view of the target can be occluded remotely by covering half the upper concave mirror in the stimulus pathway (F). The participants can then only see the target with one eye, but photorefraction of both eyes can still take place via the other optical pathway. Having the occluder remote from the participants’ face makes it particularly suitable for use with infants. Typically, approximately one third of older participants in our lab are aware they have been occluded, one third are aware “something has happened to the image” but cannot define what it was, while a further third are completely unaware of the occlusion. During testing, older infants frequently tried to touch the nearest target images, and adults, when asked to try to touch the nearest images at the end of testing, pointed to appropriate points in space, confirming it is a realistic target.

The target monitor moves such that the image is placed optically at 0.25m, 0.33m, 0.5m, 1m and 2m from the participant’s eyes, representing response demand of 4, 3, 2, 1 and 0.5 diopters (D) and meter angles (MA). By using meter angles as a the unit of measurement of vergence we were able to compare vergence and accommodation responses in relation to target demand more accurately between participants with widely different inter-pupillary distances and also plot both on the same scales e.g. a 0.5m target demands 2D of accommodation and 2MA of vergence. A pseudo-random testing order of 3, 0.5, 4, 1, 2 D and MA demand was used so that near and farther targets alternated. Data from the 4D & MA demand target was discarded because excessive pupillary constriction prevents collection of many readings at 25cm, but this target was retained in the testing sequence to maintain the near/distance alternation. Even if data is collectable, there are two additional possible causes of unacceptable imprecision at 0.25m that are negligible for the farther targets. Firstly, there is a slight induced astigmatic error of subjectively around 0.5DC induced by the vertical offset of the concave
mirrors which reduces to subjectively less than 0.25DC at 0.33m. Secondly, the participant sees the clown face within the (masked as much as possible by graded dark acetate filters) screen edges as it approaches.

The lower concave mirror is actually seen in physiological diplopia, with the screen visible within the overlap of the two diplopic images. For the 0.25 target, this physiological diplopia overlap is smaller and just excludes the very far edges of the image of the whole screen (although the clown face is still well within the overlap) This might also slightly degrade the fusional stimulus at 0.2

We also excluded this target because of the possibility that off-axis differences in peripheral refraction might induce inaccuracy, although even at 25cm and even if only one eye was doing most of the ocular rotation due to a head turn, the vergence angle would only be just over approximately 10°, well within the limits within which peripheral refractive errors are insignificant (Calver, Radhakrishnan, Osuobeni, & O’Leary, 2007).

The monitor was moved by a belt powered by a motor outside the apparatus and beyond the farthest target distance at 2.75m from the participants. While the motor could be heard during target motion, so alerting the participants that movement was occurring, it gave no clues to the target position or direction of movement. The target screen moved at 0.4 meters per second.

**Target Validation**

The clown target to stimulate accommodation as much as possible, containing high contrast, coloured edges of a wide range of spatial frequencies, down to one pixel in size, and in particular to include facial features with eyes, mouth and a “hairline” so as to be most interesting to infants, with two versions with different details alternating at 1Hz.

Gabor targets have been used to open the accommodative loop in adult studies (Tsuetaki & Schor, 1987) but to our knowledge there are no guidelines as to which type of Gabor in terms of contrast gradient maximally opens the accommodation loop best (particularly in infants compared to adults).

We made a pragmatic choice of Gabor target to use from those available in high enough resolution from the online literature. We needed a clear contrast gradient and high enough spatial frequency to allow accurate motor fusion of the image. The target chosen was downloaded from Figure 4 of Allard et al, *Journal of Vision* February 22, 2006 vol. 6 no. 4 article 3 [http://www.journalofvision.org/content/6/4/3.full](http://www.journalofvision.org/content/6/4/3.full), with a formula of

\[ S(x, y) = \cos(fx + p) \exp(-\frac{x^2 + y^2}{2\sigma^2}), \]

and this image seemed clear and of good resolution. In our paradigm, although disparity cues were completely eliminated, our scaled Gabor target contained higher spatial frequencies at 2m (1.58) than the <0.5cpd recommended by Tsuetaki& Schor (Tsuetaki & Schor, 1987). We chose this slightly higher spatial frequency since we were keen to keep an adequate stimulus for accurate fusion. Also, mindful of failure by others when testing infants with a completely diffuse target (Currie & Manny, 1997), we aimed to retain some cognitive interest within the target itself to maintain attention and retain disparity cues while minimizing detail.
We had to adapt this Gabor image by merging it with a 2D Gaussian blob on a black background using the Adobe Photoshop “Merge Layer” (“normal” setting) feature to further retain fusion, direct attention, and ensure the two targets were of similar subjective size (although this was of course difficult with the Gabor where there are no edges). We needed to have a black surround so that the screen edge within the black shuttering was a minimally visible as possible.

As with the clown target we needed two Gabor versions, and the overall size of the stimulus needed to be equivalent, so Photoshop was used to further manipulate the image. The image was coloured either green or yellow (so that duochrome effects of using colours at more extremes of the visible spectrum would be minimized). Different sizes of this blob were trialled subjectively on adults at the pilot stages so that the Gabor section of the target merged invisibly with the black surround with no apparent edges and appeared subjectively the same size as the clown, and both the green and yellow versions appeared the same size. We also checked subjectively that inducing refractive blur with lenses up to +/- 4D (broadly the levels of optical blur likely to be induced by the stimuli and levels of refractive error we were interested in) made little difference to the subjective clarity of the image at 33cm. The yellow version of the Gabor initially subjectively appeared slightly bigger despite identical processing, so when alternating with the green version there was a jumping backwards and forwards illusion, so we reduced the size of the blob surround slightly so that this illusion disappeared, although the spatial frequency of the central grating portion remained the same.

The paper shows that the o condition still drove some residual responses. These could not have been due to disparity cues because these are entirely excluded by the remote occlusion. It has been suggested that our DoG target may not sufficiently minimize blur cues, so we have carried out a subsequent study of 29 young naive adult participants, comparing accommodation and vergence responses to three alternative low detail targets. The first was the target used for this study (A). The second (B) was an image with a grating with resolution of 0.99 cycles/deg at 2m and 0.16 cycles/deg at 33cm when unscaled, and 0.99 cycles/deg at all distances when scaled. The third (C) was the diffuse spot target (see Figure) we had used for the Gabor target processing. They were matched for luminance across the target and because the target edge was diffuse, also for subjective impression of size and brightness.

There were no significant differences between the two Gabor targets in terms of vergence and accommodation responses and response gains in any of the eight target conditions. In comparison to
the diffuse spot, both Gabor targets produced significantly greater accommodation responses gain than the spot in two of the eight conditions (the DiPr and Di targets). Examination of the data, however, also showed highly significant differences in vergence between these targets ($F = 51.32, p < .0001$ and $F = 20.74, p < .0001$) but only due to less vergence to the spot target at 3MA demand (DiPr: $A = 2.81\text{MA}$, $B = 2.68\text{MA}$, $C = 2.14\text{MA}$; Di: $A = 2.71$, $B = 2.73$, $C = 2.63\text{MA}$). If vergence response gain was used as a covariate, differences in accommodation gain between the targets were not significant ($F(2, 67) = 1.94$, $p = 0.15$ and $F(2, 67) = 2.4$, $p = 0.1$). We typically find disparity a stronger drive to accommodation than is blur in naive young adults, so it is likely that the poorer accommodation to the spot target is as much, or more, due to the target being an insufficient target to drive vergence (and so lead to poor accommodation indirectly via the CA/C linkage), rather than being a more impoverished a blur cue. We accept that the choice of such targets represents a compromise between minimizing blur while retaining adequate disparity and looming cues, but we feel this confirms the superiority of the Gabor targets as the optimal stimuli to minimize blur while retaining fusional potential. Reducing the spatial frequency of the Gabor target did not reduce responses significantly so we feel that a target with lower spatial frequency would have been not opened the accommodation loop further without also compromising fusion.

The other possible source of residual cues in the Min condition could have come from the size cues of the screen edges. The whole apparatus, including the black cloth screen is contained in black shuttering so that the target is seen against a dark background, but the minimal background screen luminance of the black target surround on the screen is still very dimly visible against the physical black screen edges. We have masked this with a diffuse gradient printed on an acetate overlay, but some minimal residual size cues remain, even though dynamic looming is eliminated by screening all target movement in the proximity-free conditions. It seems likely that these residual proximal cues are the source of the minimal-cue residual responses. This is supported by the data in the paper which finds that responses to the Min target reduce with age, in line with the decline in the influence of proximal cues. Other sources of residual responses could have been experience of repeated testing, or unquantifiable voluntary influences.

**Data Collection**

The tester watched the traces during testing and the target was only moved to the next position in the sequence when traces of both vergence and accommodation could be seen by the tester to have been stable for at least two seconds. Off-line, data were converted to vergence (in degrees) and accommodation (in D), and responses were charted against time for the whole run of all five target positions and visually inspected (Figure 5). The macro searched for spikes of data caused by blinks and removed data points immediately before and after them. Representative vignettes of the most stable 25 continuous data points were selected for each target position. Vignettes were only chosen from sections of the data where the response had settled and flattened out for at least 0.5 sec (Tondel & Candy, 2007), but before any tonic changes would be expected to have occurred, so although there may have been a dynamic cue to the target position, we did not assess a dynamic response in this study. These responses were averaged, and the accommodative and vergence planes were calculated using a macro developed in our laboratory which uses raw data corrected for individual angle lambda and interpupillary distance (IPD), and a systematic error in increasing underestimation of accommodation in comparison to dynamic retinoscopy found during earlier calibration studies (using the formula
1.2385x+0.799, where \( x \) equals the PlusoptiXSO4 accommodation measure (Horwood & Riddell, 2008).

We obtained the best estimate of angle lambda when fixing at infinity by plotting the \( y \)-intercept of the nasal displacement from the pupil center averaged across both eyes at all four fixation distances in the all cue (\( bdp \)) condition. True IPD was calculated from the \( y \)-intercept of PlusoptixSO4 IPD plotted against target distance at all four fixation distances in the \( bdp \) condition. This was used to calculate responses in MA for each participant. Individual accommodation calibration was not carried out in view of the long testing session for very young infants, but Blade & Candy (2006), using similar apparatus have shown that group means in infants and adults are similar. Inter- scorer reliability on masked scoring where each scorer was free to choose the vignette was excellent. For both vergence and accommodation, this analysis showed a high agreement: for vergence: \( r=0.99 \), mean inter-scorer difference = 0.037±95%CI 0.37MA; for accommodation \( r=0.99 \), mean inter-scorer difference 0.0095±95%CI 0.175D.

The testing order is described in the main article text.

Attention

Infant fixation was monitored in real time on the PlusoptiX video-monitor at the time of testing. Attention during each run was scored immediately after each run on an ordinal scale between 1 and 5, (1= totally calm and attentive throughout, 5= totally inattentive to the target). Only runs scoring 3 or less (3 = the infant was observed to be looking steadily at the target for at least two seconds despite mild fussiness) were analyzed. Infants were less engaged by the more impoverished targets and attention frequently waned towards the end of repeated testing. However, 82% of infants whose attention score reduced during testing in single cue conditions achieved a better attention score when the \( bdp \) target was re-presented immediately afterwards at the end of testing, demonstrating that fatigue was not the main reason for attention loss. Wilcoxon Signed Ranks tests showed no significant difference between the first and repeated \( bdp \) conditions (\( z=-1.732 \), \( p=0.08 \)). Attention was significantly better with 3-cue vs 1-cue (1.1 vs 1.36; \( z=-4.421 \), \( p<0.0001 \)), 2-cue vs 1-cue (\( z=-4.13 \), \( p<0.0001 \)) and 1-cue vs “zero”-cue (\( z=-4.06 \), \( p<0.0001 \)) conditions although there were no attention differences between targets within the 1 or 2 cue blocks. Infants became more distractible with the impoverished cues, so fewer runs were collectable but infants were just as likely to be distracted by being rendered monocular in the disparity-free conditions as they were by waiting for the screen to be removed in the proximity-free conditions or by being given the less salient DoG target in the blur-minimized conditions.

Refraction estimates

We needed to estimate refractive error as accurately as possible in the infants, since many would be expected to be significantly hyperopic, thus resulting in variability in accommodation demand for any given stimulus. Cycloplegic refraction on each visit would have provided a gold standard measure of maximum hyperopia, but was not attempted in order to maximize participant recruitment and retention during a longitudinal study dependent on repeated attendance, and in response to ethical constraints. We have reported that a reliable estimate of true hyperopic error when cycloplegia is not available is given by the maximally hyperopic refraction (MHR) found at any point during the whole testing session in our laboratory (Horwood & Riddell, 2009). In that study MHR correlated closely with cycloplegic retinoscopy (\( r=0.93 \)) and was a better estimate than Mohindra retinoscopy (Mohindra, 1977). Results of cycloplegic retinoscopy were available for 17 of the infants in this study between three
and six months of age and agreement was very good between these results and MHR ($r = 0.83, 95\% CI \pm 0.50D$). In addition, while Mohindra retinoscopy was not possible with some fussier infants, a measure of MHR was obtainable at every visit. We therefore used MHR as the best estimate of refractive error for each infant each visit. Infants with anisometropia >1.00D or MHR of >+2.00 were excluded in this study. A continuous measure of astigmatism was not recorded by the PlusoptiX SO4 in PowerRefl mode, so was not considered here. Within the angular change in fixation (<10°) demanded by even the nearest target, off-axis errors or differences in peripheral refraction (Charman & Radhakrishnan, 2010) are not likely to have induced refraction artifacts.

References


