Human Newborn Color Vision: Measurement with Chromatic Stimuli Varying in Excitation Purity

Russell J. Adams

Memorial University of Newfoundland, St. John’s, Newfoundland, Canada A1B 3X9; and Clinical Associate of Neonatology, St. Clare’s Mercy Hospital, St. John’s, Newfoundland, Canada

and

Mary L. Courage

Memorial University of Newfoundland, St. John’s, Newfoundland, Canada A1B 3X9

Groups of newborn human infants (N = 180) were habituated to large 16° achromatic (“white”) lights of varying luminance (0.35 to 1.16 log cd/m²) and then tested for recovery of habituation to 16° green (dominant λ = 545 nm), yellow (dominant λ = 585 nm) or red (dominant λ = 650 nm) lights which varied in the level of excitation purity (range = 32 to 83%). Results showed that newborns discriminated the chromatic stimuli from white only when excitation purity values exceeded at least 41% for 545-nm green, 47% for 650-nm red, and 65% for 585-nm yellow, limits much higher than those for adults (<1%). Taken together with the results from previous experiments, these saturation discrimination data (with the exception of the yellow data), provide some support for an expanded MacAdam ellipse model of early color discrimination (Brown, 1993; Teller & Lindsey, 1993). This helps reinforce the current view that neonates’ vision is based on general rather than selective immaturities or inefficiencies within the requisite optical, photoreceptoral and neural mechanisms. © 1998 Academic Press

INTRODUCTION

Energy which radiates in sufficient quantity from the visible portion of the electromagnetic spectrum (400 to 700 nm) stimulates the human nervous system.
system and is perceived as ‘‘light’’. If that radiation is restricted to a fairly narrow band of wavelengths within the visible spectrum, the light will appear a specific color. For example, light radiating predominately at the shorter wavelengths (e.g., 450 nm) appears ‘‘blue’’ and light radiating at the longer wavelengths (e.g., 650 nm) appears ‘‘red.’’ In addition to wavelength, light can also vary on two other physical dimensions: its intensity, defined as the number of quanta/second that are delivered to the retina, and its colorimetric purity, defined as the proportion of ‘‘white’’ light that is mixed with light of a given wavelength. In nature, light varies along all three physical dimensions, thus producing an almost infinite number of perceived colors.

Of critical interest to vision scientists is to measure an organism’s ability to respond to differences in wavelength, intensity, or purity under a wide variety of stimulus conditions (see Boynton, 1979; Hurvich, 1981; or Jacobs, 1981 for extensive discussion). The psychophysical functions derived from such experiments (i.e., ‘‘hue,’’ ‘‘brightness,’’ and ‘‘saturation’’ discrimination functions) not only describe the limits of the organism’s perceptual abilities but provide essential information about the nature of the underlying anatomical and physiological mechanisms responsible for color vision and other critical visual functions.

Much has been learned about the early development of these neural mechanisms by examining various aspects of color vision in human infants. Several laboratories have conducted electrophysiological and behavioral experiments on young infants’ intensity/brightness discrimination (Adams, Maurer & Cashin, 1990; Allen Banks & Norcia, 1993; Clavadetscher, Brown, Ankrum, & Teller, 1988; Mercer, Courage, & Adams, 1991; Morrone, Burr, & Fiorentini, 1993; Peeples & Teller, 1975; Teller, Peeples, & Sekel, 1978), and wavelength/hue discrimination (Adams, 1989, 1995; Adams, Courage, & Mercer, 1994; Adams, Maurer, & Davis, 1986; Adams et al., 1990; Hamer, Alexander, & Teller, 1982; Mercer et al., 1991; Packer, Hartmann, & Teller, 1984; Varner, Cook, Schneck, McDonald, & Teller, 1985), but purity/saturation discrimination remains relatively unexplored. The only report of infants’ saturation discrimination (Bornstein, 1978) revealed that 4-month-olds discriminate chromatic stimuli of moderate excitation purity (30 to 50%) from those of low purity (5–11%). However, the usefulness of this early study is limited by methodological problems (Werner & Wooten, 1979), the lack of threshold estimates, and the fact that the infants were much older (and perhaps of less theoretical interest) than the infants tested in the more recent and better controlled color vision experiments cited above. To begin to address this significant gap in the infant color vision literature, we have attempted here to measure human newborns’ saturation discrimination in three spectral regions, namely those regions (the green, yellow, and red regions) in which newborns have already shown evidence of discriminating highly saturated chromatic stimuli from achromatic stimuli.

Data from experiments on newborns’ saturation discrimination are also
useful for addressing the issue of whether early visual development is limited by a critical immaturity within a specific optical, photoreceptoral, or neural mechanism, or is limited by more universal immaturities or inefficiencies across all mechanisms (Banks & Bennett, 1988; Brown, 1990). To date, tests of neonatal chromatic–achromatic discrimination (Adams, 1995; Adams et al., 1994) provide some support for the general inefficiency hypothesis. These experiments, which have employed highly saturated chromatic stimuli, show that the pattern of newborns’ successful and unsuccessful chromatic-achromatic discriminations is generally similar to adults’ but is decreased by several orders of magnitude. In other words, newborns’ color vision abilities appear to be a predictably scaled-down version of adults’. However, because all of the stimuli used in these experiments were highly saturated, the support for a general immaturity/inefficiency model is still tentative. Thus, the logical next step would be to select those chromatic stimuli that newborns have already shown evidence of discriminating from white (545-nm green, 585-nm yellow, and 650-nm red) and perform further experiments in which the saturation of those hues is reduced. In addition to providing a more accurate test of the general immaturity/inefficiency model, estimation of saturation discrimination in different spectral regions would better define the nature of the human newborn color space and the chromaticity characteristics of relative neutral zones (Adams, 1995).

METHOD

Subjects. Each of nine groups of infants (N = 173) was tested with a separate chromatic stimulus. The babies (90 females and 83 males) were 1 to 7 days of age (M = 3.1 days; SD = 1.2 days), at least 38 weeks gestation, free of any detectable visual and neurological abnormalities, and had no familial history of color vision anomalies. An additional 95 newborns were tested but not included in the sample: 88 because they were too sleepy or fussy to complete the procedure, 5 because of low interobserver reliability (Spearman rho < 0.80), and 2 because of apparent ocular problems that were noted during testing.

Apparatus and stimuli. The chromatic and achromatic stimuli were generated with the apparatus described in detail by Adams (1995) and Adams et al. (1990, 1991). Briefly, light is produced by a Beseler Dichro 45 MX II color computer, a microprocessor-controlled color head used in photographic enlarging. The system operates by passing broadband 3250°K white light emitted by a 250W lamp through an adjustable pack of cyan, magenta, and yellow glass dichroic interference filters, then diffusing and focusing the light into a mixing chamber. Within the chamber are three gallium arsenide photodetectors which continually monitor the chromatic characteristics of the light. The digital output of the detectors is displayed on the outside of the apparatus. The light exits the mixing chamber through a small 3 × 3 cm opening composed of heat-absorbing glass and is then directed and focussed
by a lens onto a vertical rear-projection screen. The luminance of the light on the screen is controlled by a variable aperture located between the lens and the glass opening of the color computer’s mixing chamber.

In our previous experiments, newborns showed evidence of discriminating from $16 \times 16$ white squares, $16 \times 16^\circ$ 545-nm green, 585-nm yellow, and 650-nm red squares, all of which possessed high levels of excitation purity (77% to 80%) and which appeared to adults to be highly saturated (Adams, Maurer, & Cashin, 1990). In the present experiment, the same Beseler Dichro 45 MX II optical system was used to recreate as closely as was possible, the $16^\circ$ red, green, and yellow stimuli, but this time each of the three hues was tested (with separate groups of newborns) at three lower excitation purity values, namely 55, 41, and 32% for the 545-nm green; 83, 65, and 46% for the 585-nm yellow; and 55, 47, and 39% for the 650-nm red. Although we intended the purity values of the red and green sets to be about the same, limitations in the apparatus prevented precise duplication. We tested with slightly higher purity values of the yellow stimulus because previous experiments revealed that newborns’ performance with this hue was relatively poorer than that with the red and green (Adams et al., 1990).

Despite the slight differences in purity value between the red and green sets, it is more important that each set of stimuli falls close to the same respective axes within the C.I.E. (1931) color space as those from the previous experiments. Figure 1 illustrates this by showing that each set of stimuli radiates out from the reference white which is located near the centre of the color space. All of the stimuli in the present experiment have the same respective dominant wavelength as those in the previous experiments (i.e., 545, 585, or 650 nm). Specifically, the C.I.E. (1931) $x$ and $y$ coordinates (measured in situ with a Minolta Chroma Meter II) were 0.35 and 0.57 for 55% green, 0.37 and 0.53 for 41% green, 0.38 and 0.50 for 32% green, 0.51 and 0.47 for 83% yellow, 0.49 and 0.45 for 65% yellow, 0.47 and 0.44 for 46% yellow, 0.59 and 0.33 for the 55% red, 0.56 and 0.34 for 47% red, and 0.54 and 0.35 for 39% red. The coordinates for all luminances of the (reference) white square were 0.42 and 0.41 and its correlated color temperature was estimated to be about 3200°K.

It is important to note that the stimuli produced with the optical system used in this and the previous studies are broad-band chromatic lights. For ease of description, the stimuli are designated by their dominant wavelength and color appearance (e.g., 650-nm red). However, the reader should be careful to note that the use of dominant wavelength does not imply that the stimuli are monochromatic nor that the color name implies that the hue appears the same to infants as it does to adults.

**Procedure.** The methods to control brightness cues and the details of the habituation procedure were identical to those described in detail in several previous studies (e.g., Adams, 1995; Adams et al., 1990, 1991). Briefly, each newborn was reclined in an infant seat that was adjusted to an angle of 45°.
FIG. 1. Location of the three sets of chromatic stimuli of varying excitation purity along their respective axes within the C.I.E. (1931) color space. Each axis radiates from the coordinates representing the reference white light used in the experiments (x = 0.42, y = 0.41) with dominant wavelengths of 545 nm (green), 585 nm (yellow), or 650 nm (red). Note that the 77% green, 80% yellow, and 80% red were from a previous report (Adams et al., 1990).

A pair of foam bumper pads supported the infant’s head in an upright position. The seat was placed so that the infant’s eyes were 38 cm from the center of a rear-projection screen (41 × 28 cm) mounted in a black board. Small 1-cm peepholes on each side of the screen permitted observers behind the screen to see the infant’s eyes. The two observers were equipped with switches that operated a shutter system mounted in front of the projector lens, and with timers to record the infant’s fixation times.

During testing, the overhead lights were extinguished and, using an infant-controlled procedure (Horowitz, Paden, Bhana, & Self, 1972), we habituated infants to a series of 16 × 16° white lights that, from trial to trial, appeared at one of five luminance values (range = 0.35 to 1.16 log cd/m²). The order of
these luminances was counterbalanced across infants. During each habituation trial, the two observers, one of whom was unaware of the spectral characteristics of the stimuli, independently recorded the length of the infant’s fixation on the stimulus by timing the period from when the reflection of the stimulus first fell over the center of the baby’s pupil until it fell off the pupil. After both observers judged the infant to have looked away from the stimulus, a trial ended and the shutter in front of the lens was closed, removing the stimulus. The habituation trials continued in this manner until the infant reached the criterion for habituation (i.e., a 50% reduction of looking time on three consecutive trials from that on the first three trials). The intertrial interval was approximately 10 s. Following habituation, the infant was given four test trials, two with a fifth white square (range = 0.35 to 1.16 log cd/m²; C.I.E. x and y coordinates = 0.42 and 0.41) not shown during habituation, and two with the appropriate chromatic stimulus. The ABBA or BAAB order of the white square and the chromatic test square are counterbalanced across babies. The luminance of the chromatic squares shown during the test approximated the middle of the range of the whites shown during habituation (range = 0.72 to 0.86 log cd/m²).

RESULTS

Sixty-five percent of the newborns successfully completed the procedure. The average interobserver reliability coefficient among subjects in the final sample was 0.91 (range = 0.80 to 1.00). During the habituation phase, newborns fixated the white squares for an average time of 251 s (range = 70 to 740 s), and they required an average of 8.8 trials (range = 6 to 20 trials) to reach the habituation criterion. The average length of fixation during the first 3 criterion trials was 171 s (range = 35 to 391 s).

For each infant, three summary statistics were obtained. These were (1) the mean fixation time to the chromatic square during the two test trials, (2) the mean fixation time to the novel achromatic square during the two test trials, and (3) to measure whether infants ‘‘treat’’ the achromatic test squares similarly to the achromatic squares presented at the end of the habituation phase, the mean of the last two habituation trials was also obtained. Several steps were taken to analyse these data. First, to normalize the distribution of fixation times, the data were transformed to natural logarithms (loge). Next, to address the primary research question concerning the points in the newborn infants’ color space at which they can make (and fail to make) discriminations among achromatic and chromatic stimuli that vary in excitation purity, we conducted three sets of single-factor within-subject analyses of variance (ANOVAs), one set for each of the three hues. This was considered to be the most direct test of whether infants are capable of discriminating chromatic from achromatic stimuli at specific points in the different spectral regions. Also, $\omega^2$ values were calculated for each group to indicate the magnitude of treatment effects. However, because there were three separate sets of statistical
comparisons, an alpha level of 0.017 was adopted to test for significance among these three sets of comparisons (see Kepple, 1991). If the ANOVA was significant, post-hoc Neuman–Kuels tests were conducted to analyze for differences among the fixation time measures. And finally, in conditions in which significance was achieved, $F$ ratios were calculated to compare the variance between fixation time to the chromatic square vs the test white square. The purpose of this last analysis was to evaluate whether significance could be accounted for by only a subset of infants in the group who showed increases in fixation time to the test chromatic stimulus, or whether the effect was due to a general increase in the group data.

Figure 2 summarizes the data by showing for each group of infants, the mean log fixation times for each of the three measures in each of the three sets of analyses. The single-factor ANOVA and $\omega^2$ values for each of the three levels of excitation purity in the 545-nm green condition revealed significant differences among the three measures of fixation time for the 55% green $[F(2,38) = 4.94, p < .012; \omega^2 = 0.12]$ and for the 41% green $[F(2,38) = \ldots$
NEWBORN COLOR VISION

The analyses of the fixation times for infants tested with the three levels of excitation purity in the 585-nm yellow condition revealed significant differences for the 83% yellow \(F(2,30) = 10.02, p < .001; \omega^2 = 0.27\) and the 65% yellow \(F(2,38) = 15.90, p < .001; \omega^2 = 0.33\]. Finally, for the 650-nm red condition, analyses revealed significant differences in fixation times for the 55% red \(F(2,38) = 4.92, p < .013; \omega^2 = 0.12\] and the 47% red \(F(2,38) = 6.63, p < .003; \omega^2 = .16\]. Post-hoc Neuman–Keuls tests showed that in each case, significance was accounted for by newborns’ longer fixation to the chromatic squares than to the white squares presented either during the test or during the last two habituation trials (all \(p < .025\)). F ratios calculated between the variance of the fixation times to the chromatic square vs. that to the test white square revealed no significant differences within any of the above groups (all \(p > .05\)). As discussed earlier, this suggests that significant differences were accounted for by a general increase in newborns’ fixation time to the chromatic squares and not to the performance of a subset of infants in the group.

However, among the fixation time measures in the groups tested with the 32% green, 46% yellow, and 39% red, ANOVA revealed no significant differences among the three measures of fixation time (all \(p > .05\)). As expected for conditions in which significance was not achieved, \(\omega^2\) values were relatively low, ranging from 0.04 to 0.07.

DISCUSSION

The present experiment was designed to probe further the limits of human newborn color vision by retesting at lower levels of excitation purity, the 16° chromatic stimuli that newborns have previously shown evidence of discriminating from white, namely 545-nm green, 585-nm yellow, and 650-nm red. Results indicated that newborns were capable of making these chromatic-acromatic discriminations when the purity levels were at least 41% for 545-nm green, 65% for 585-nm yellow, and 47% for 650-nm red. For comparison, adults tested with these chromatic stimuli discriminate purity differences of less than 1% (Boynton, 1979; Hurvich, 1981).

The specific design of the present experiments and the controls for brightness should ensure that any discrimination of the chromatic from the white lights are made on the basis of differences in spectral information. Alternatively, one could argue that the discrimination was based on newborns detecting a brightness difference between the white squares and the chromatic square (see Brown, 1990). However, as explained in previously published reports (Adams, 1995; Adams et al., 1986, 1990, 1991), there are several reasons why this explanation is unlikely: (1) the relatively wide range of luminances employed during the habituation phase is likely to enclose each baby’s luminance match, (2) newborns have demonstrated that they are extremely insensitive to achromatic brightness cues after luminance is “jittered” during the habituation phase (see Adams et al., 1990, Experiment 1), and (3)
the procedure has been effective in revealing chromatic-achromatic discrimination failures. For example, in the present experiments, newborns show evidence of discriminating the white from the 650-nm red at one purity level (47%), but fail with the red at a slightly lower level (39%). Note that except for differences in excitation purity, both stimuli have the same wavelength and luminance characteristics. This implies that luminance is not a significant cue and that the habituation paradigm is effective in revealing situations of both successful and unsuccessful discrimination, presumably on the basis of infants’ ability to detect chromatic information.

The findings reported here support the growing body of evidence indicating that human neonatal color vision is extremely poor. Experiments to date (see Adams, 1995 for a review) reveal that newborns fail to discriminate from white, a wide variety of large 16° chromatic stimuli (namely 450- and 470-nm blue, 493-nm blue-green, 565- and 572-nm yellow-green and purple (complementary = 571 nm). Moreover, with stimulus sizes smaller that 8° newborns have yet to show evidence of making any chromatic–achromatic or chromatic–chromatic discriminations (Adams, 1989; 1995). The present findings show that the few chromatic stimuli that neonates can discriminate from white not only need to be large, but also (compared to adults) need to be highly saturated.

Perhaps most importantly, the investigation of newborns’ ability to discriminate stimuli of varying purity helps to shed light on the debate about whether early human vision is limited by an immaturity within a particular visual system mechanism, such as is found in a color-deficient adult who lacks a specific photoreceptor/photopigment or whether the deficits are due to more broad-based immaturities across all mechanisms (Banks & Bennett, 1988; Brown, 1990). It was suggested by one of the anonymous reviewers of a recent paper that the existing pattern of newborns’ successful and unsuccessful chromatic–achromatic discriminations is fit nicely by a simple quantitative expansion (120×) of the adult MacAdam ellipse which corresponds approximately to the point in the C.I.E. color space occupied by the reference white light (see Adams, 1995, Fig. 5). In his now classic experiment, MacAdam (1942) found that if one plots the chromaticity coordinates of the standard deviations of the color matches that surround any reference stimulus within the color space, these points will form an ellipsoid shape around the referent. Thus, the ellipse is a proportional estimate of all of the stimuli that are just discriminable from the referent. The size of these ellipses varies substantially across different regions within the color space and reflects presumably, the nature and relative efficiency of the underlying color vision mechanisms. The importance of MacAdam’s findings is underscored by the fact that other researchers have now adopted an ellipse model to interpret the existing infant color and contrast data and to make quantitative predictions about reductions in the shape and size of infants’ ellipses that should occur during visual development (Brown, 1993; Teller & Lindsey, 1993).
FIG. 3. Expansion (by 100×) of the approximate adult MacAdam ellipse which corresponds to the C.I.E. coordinates of the white light (W) used in all of the newborn experiments conducted to date (Adams, 1995; Adams et al., 1986, 1990, 1991; and the present study). The asterisks indicate those 16° chromatic stimuli that newborns discriminate from white, and the open circles represent those failed. Note that the ellipse encloses (except for three of the 585-nm yellows) all of those chromatic stimuli failed but does not enclose those which newborns show evidence of discriminating from white.

Figure 3 shows the proposed newborn ellipse from the previous paper (Adams, 1995) adjusted slightly (100× adult size rather than 120×) to best fit the addition of the three new sets of chromatic stimuli tested in the present experiment. On the positive side, note that the edge of the ellipse intersects the axes of the 545-nm green and the 650-nm red between the stimuli successfully discriminated from white (41% green and 47% red), and those not discriminated (32% green and 39% red). Thus, for both the 545-nm green and 650-nm red sets of stimuli as well as most of those tested in the previous experiments (450- and 470-nm blue, 493-nm blue-green, 565- and 572-nm yellow-
green, 595-nm orange, and 571-nm (or purple), the expanded ellipse appears to enclose those not discriminated from white, and exclude those that newborns do show evidence of discriminating.

However, the ellipse explanation is diminished substantially by the results with 585-nm yellow. Newborns show evidence of discriminating white from 585-nm yellow at three excitation purity values: 65, 83 (present study), and 80% (Adams et al., 1990). Note that Figure 3 shows that the ellipse encloses all chromatic stimuli from about 563 nm to about 593 nm, even those at maximum purity levels. Thus, newborns should not discriminate from white, any chromatic stimuli in this spectral region. Although this prediction holds for the other stimuli tested within this region (i.e., 565- and 572-nm yellow-greens), the results for some of the 585-nm yellows are inconsistent. What is also puzzling is that in experiments employing slightly different methods and slightly smaller stimuli, newborns (Adams et al., 1994) and even older 2-month-old infants (Mercer et al., 1991; Teller et al., 1978) fail to discriminate 561-, 580-, and 585-nm yellows from white, a result that is consistent with the expanded ellipse model. Even if the expanded ellipse is reduced and/or the angle of its central axes shifted (i.e., it is tilted in a different direction), the present results with yellow cannot be reconciled with the data from the other spectral regions.

Unfortunately, we cannot offer any definitive explanation for this paradoxical result except to propose that there was some artifact within the conditions of the yellow experiments or within the yellow light itself. However, observation and careful remeasurement of the light failed to reveal anything obvious. On the other hand, perhaps many of the infants within the groups tested with yellow differed systematically in some way from the other groups. One observation that may support this explanation was that the percentage of infants who failed to complete testing was, for some inexplicable reason, consistently higher in the groups tested with 585-nm yellow ($M = 48\%$; range = 43 to 59%) compared to the many groups tested with the other 16° chromatic stimuli ($M = 27\%$; range = 17 to 41%) both in this and in our previous experiments (Adams, 1995; Adams et al., 1986, 1990, 1991). This difference increases the possibility that because of the greater attrition rate, the yellow groups may have been composed of more exceptional (i.e., more alert, attentive, or mature) infants. Therefore, given that the statistical analyses are designed to detect group trends, a sample which includes a relatively greater number of alert/mature babies may result in a data set skewed significantly toward positive evidence of discrimination.

Alternatively, perhaps the yellow data are genuine and newborns possess a unique combination of both specific and general immaturities which result in an “ellipse” which is initially asymmetrical (i.e., it excludes a greater proportion of the yellow region of the color space than that predicted in Fig. 3). In the case of the present results, the shape that fits the data would be one that runs narrowly through the yellow-green region and widens as it
enters the blue region of the color space. Perhaps with development, as the shape shrinks in size, it may become more ellipse-like because of more rapid reductions in particular spectral regions. The symmetrical ellipse shape would in turn indicate that infants’ color vision mechanisms have achieved a relatively “balanced” or general state of immaturity. Although speculative, this suggestion is supported by data from other experiments (see Adams et al., 1994) which reveals that, over the first 2 months, chromatic-achromatic discrimination appears to develop relatively faster within the short-wavelength region, a result which is consistent with the above argument. Future experiments designed to track more specifically, early developmental changes in the pattern of chromatic-achromatic or chromatic-chromatic discriminations within specific portions of the color space should help test this hypothesis.

In summary, except for the paradoxical 585-nm yellow data, much of the evidence obtained from color vision experiments in newborn and in older human infants (Banks & Bennett, 1988; Brown, 1993; Teller & Lindsey, 1993) appear to support the notion that unlike color-deficient adults, neonates relatively deficient chromatic vision is based primarily on general immaturities or inefficiencies within many of the requisite optical, photoreceptoral, and neural mechanisms. It will be interesting to examine whether the results from future developmental studies of color and spatial vision conform to a “shrinking ellipse” or to some other quantitative model. If so, it will provide a very solid footing for more accurate prediction of the normal development of many important visual system functions, an outcome which has great significance for the interpretation of individual clinical cases.

REFERENCES


Received: May 3, 1996; revised: July 21, 1997