Current Biology, Vol. 14, 247–251, February 3, 2004, ©2004 Elsevier Science Ltd. All rights reserved. DOI 10.1016/j.cub.2004.01.031

Stabilized Structure from Motion without Disparity Induces Disparity Adaptation

Fang Fang and Sheng He*
Department of Psychology
University of Minnesota
75 East River Road
Minneapolis, Minnesota 55455

Summary

3D structures can be perceived based on the patterns of 2D motion signals [1, 2]. With orthographic projection of a 3D stimulus onto a 2D plane, the kinetic information can give a vivid impression of depth, but the depth order is intrinsically ambiguous, resulting in bistable or even multistable interpretations [3]. For example, an orthographic projection of dots on the surface of a rotating cylinder is perceived as a rotating cylinder with ambiguous direction of rotation [4]. We show that the bistable rotation can be stabilized by adding information, not to the dots themselves, but to their spatial context. More interestingly, the stabilized bistable motion can generate consistent rotation aftereffects. The rotation aftereffect can only be observed when the adapting and test stimuli are presented at the same stereo depth and the same retinal location, and it is not due to attentional tracking. The observed rotation aftereffect is likely due to direction-contingent disparity adaptation, implying that stimuli with kinetic depth may have activated neurons sensitive to different disparities, even though the stimuli have zero relative disparity. Stereo depth and kinetic depth may be supported by a common neural mechanism at an early stage in the visual system.

Results and Discussion

Spatial Context Can Disambiguate the Ambiguous Rotating Cylinder

Ambiguity of structure from motion generated from orthographic projection of 3D moving objects can be disambiguated by information (e.g., disparity, speed, contrast, etc.) that specifies the depth order of the moving elements [5–8]. Multistable ambiguous stimuli tend to covary [9–11], suggesting the possibility that the perception of an ambiguous stimulus could be influenced by its spatial context. Sereno and Sereno (1999) demonstrated that motion of the 2D surround of an ambiguously rotating stimulus can bias the oppositely moving dots to be perceived as the front surface of a 3D kinetic sphere as stimulus could almost completely stabilize the ambiguous stimuli.

The stimulus used in our study is a vertical rotating cylinder generated from an orthographic projection of dots on a rotating 3D cylinder and its similar stimuli used in previous psychophysical [3, 7] and physiological [4, 15, 16] studies. The ambiguous stimulus, perceived as a rotating cylinder with its rotation direction switching every few seconds, was presented to only one eye (Figure 1A). (The percepts of two concave or convex sheets, moving across each other, are also possible [3] but were rarely seen by our observers; hence, they are not discussed in this paper and not depicted in figures.) When disparity information was added to the two ends of this bistable cylinder (i.e., a whole cylinder was presented to one eye, and only two ends of the cylinder were presented to the other eye), the whole cylinder was perceived to rotate in the direction specified by the disparity in the two ends, although the middle section contained no information to specify the depth order (Figure 1B). For the four observers tested, all perceived the cylinder as rotating unambiguously, 100% of the time, over multiple 1 minute periods. The spatial contextual cue was very effective in disambiguating the ambiguous motion.

Our observation differs from earlier reports of contextual biases on ambiguous rotation. The contextual bias due to simple 2D motion contrast simply enhances the opposite direction of motion in the central region and thus biases dots moving in such a direction to be perceived as being in front [12]. In the case of linkage between multiple bistable stimuli, the coupling tends to break down between unambiguous and ambiguous stimuli [11]. The key reason that the ambiguous and unambiguous sections in our stimulus remain strongly linked is that monocular presentation of the ambiguous section of the stimulus reduced the disparity contrast between nonzero relative disparity in the unambiguous section and zero relative disparity in the ambiguous section. Additionally, unlike in earlier studies in which the ambiguous and unambiguous stimuli appeared as separate and distinct objects, we made the ambiguous and unambiguous sections of the stimulus appear to be parts of the same object and thus enhanced the effectiveness of the disambiguation.

Occlusion in general is a strong cue to depth relationships. The occlusion cue has been shown to be somewhat effective in disambiguating ambiguous kinetic depth perception [17, 18]. We also tested if an occlusion cue can disambiguate the surface assignment of the

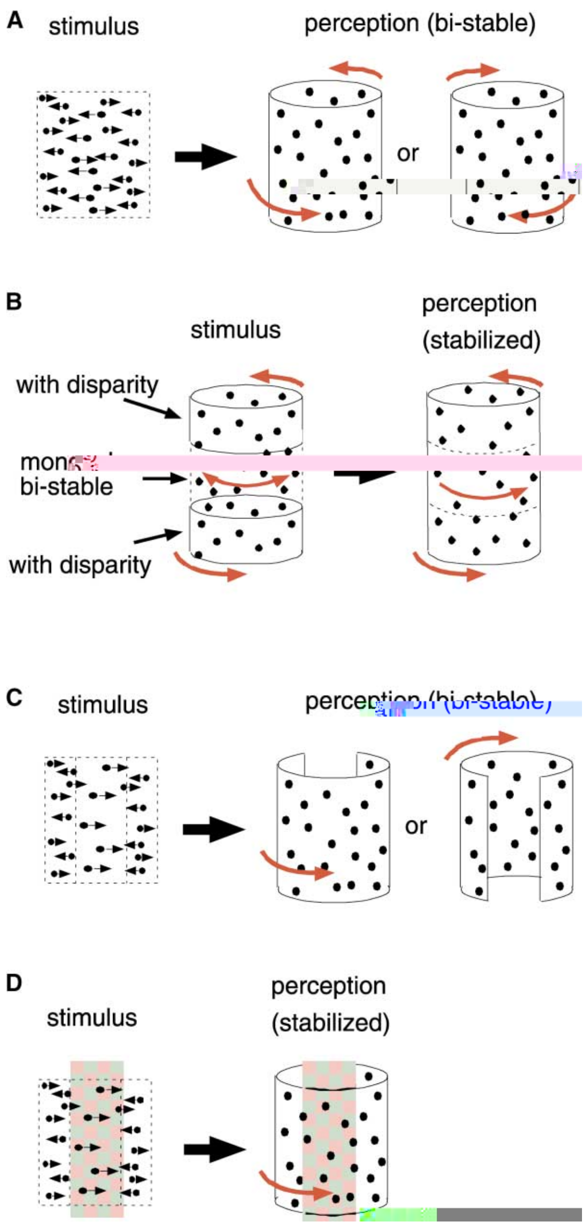

Figure 1. Ambiguous Stimuli and Their Stabilization from Contextual Cues

(A) Bistable rotating cylinders. The 2D motion signal is consistent with either of two 3D interpretations.

(B) When the bistable cylinder is placed between two unambiguously rotating cylinders (from disparity), the physically bistable middle section is disambiguated by the two ends.

(C) A section of dots moving in one direction is removed, creating apparent objects like occluders, but the percept remains bistable.

(D) A visible checkered occluder is placed behind the front surface, blocking dots of the back surface. Perception is completely stabilized.

to the back surface. We then sought to enhance the occluder by making it explicit. A checkered rectangle was placed behind the front surface and blocked parts of the back surface. This manipulation was effective in eliminating the ambiguity of surface assignment (Figure 1D). The perceived rotation became completely unambiguous for three of the four observers (see Experimental Procedures) over multiple 2 min test periods and became almost completely unambiguous for the observer S.H., who occasionally (less than 10% of the time) saw the dots as rolling behind a semitransparent occluder.

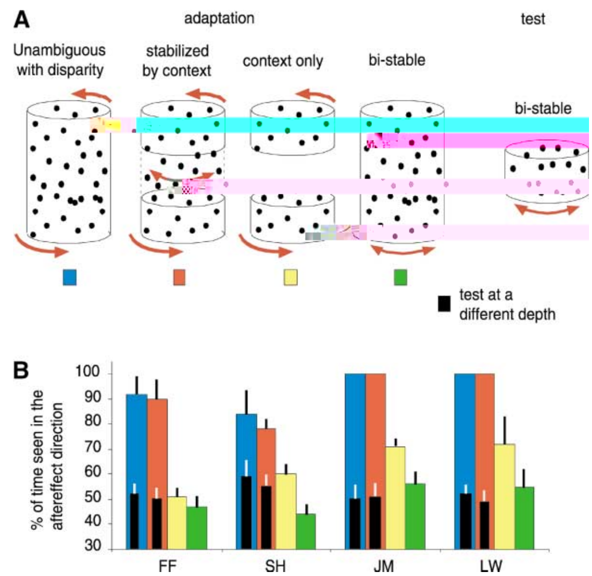

Figure 2. Effects of Adaptation to the Rotating Cylinders, including the Context-Stabilized Ambiguous Stimulus

(A) Four different adaptation stimuli were used. The test stimulus was an ambiguous cylinder. For the first two adaptation conditions, the test stimulus was placed at the same, as well as a different, stereo depth from the adaptation stimuli.

(B) The adaptation effects, as measured by the proportion of time observers perceived the rotation direction opposite to the adapted direction. When the adapting stimuli were either disambiguated with full disparity or contextual disparity, the aftereffect was significantly larger than the other conditions ($p < 0.01$). The aftereffect was substantially reduced when the test stimulus was placed at a different depth than the adapting stimuli (black bars). Error bars are 1 standard deviation. See the text for details.

Disambiguated Motion Can Generate an Aftereffect

Prolonged exposure to unambiguously rotating stimuli [7, 19], but not to an ambiguously rotating stimulus [20], can lead to rotation aftereffects. Can we observe an aftereffect from a stimulus that is perceptually stabilized by its context? Note that in the current study, the adapting portions, direction of rotation of the test dots, that are in front, are not specified in the local adapting stimulus but are perceptually stabilized by context.

Immediately after 1 min of adaptation to one of the four adapting stimuli, observers were presented with a bistable test cylinder for 15 s (Figure 2A). As shown in Figure 2B, consistent with earlier studies [7, 20], adapting to the cylinder that was disambiguated by full disparity resulted in a strong aftereffect. However, adapting to the context-stabilized ambiguous rotating cylinder also resulted in a strong aftereffect. All four observers perceived the test stimulus rotating in the direction opposite to the adapting direction for most of the 15 s

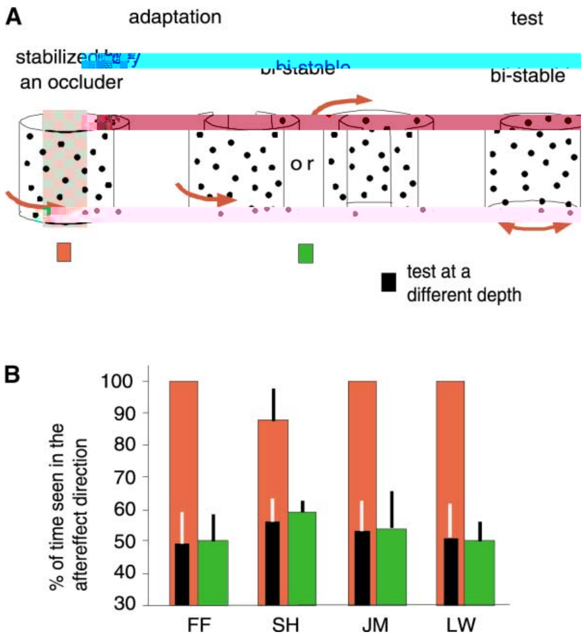

Figure 3. Effects of Adaptation to the Rotating Cylinder Stabilized by the Occlusion Cue

(A) The two adaptation stimuli had the same 2D motion signal. The stimulus with the explicit occluder was stabilized, whereas the one with the implicit occluder remained bistable, which served as a nice control condition. For the stabilized adaptation condition, the test stimulus was placed at the same, as well as a different, stereo depth from the adaptation stimulus.

(B) The aftereffect in the physical-occluder condition is significantly larger than that in the control condition, in which the 2D motion was the same but the 3D interpretation was bistable ($p < 0.01$). The aftereffect also required that the adapting and test patterns be placed on the same depth plane (black bars). Error bars denote 1 standard deviation.

testing period. In addition to the two stabilized rotation stimuli included as adaptors (fully disparity unambiguous, cone-stabilized, ambiguous), two control conditions were also included. In one control (cone only), observers adapted to the two end units alone, without the middle ambiguous section. This was to test whether the aftereffect could simply be a spreading of adaptation from adjacent regions as a result of, for example, large receptive fields of the underlying neurons. Another control condition (bistable) was simply the extended bistable cylinders. This was to test whether merely being exposed to a bistable rotating cylinder for 1 min would lead to some stabilization during the test phase. After adaptation in both control conditions, observers perceived the test cylinder as a bistable one, alternatively rotating in either direction with close to 50% chance (Figure 2B). When adapted to the two end units alone, the two naive observers (J.M. and L.W.) showed a weak aftereffect, likely due to less stable fixation during adaptation. However, the small aftereffect is much weaker than that generated by the stabilized, ambiguous adaptor.

When the ambiguous cylinder was stabilized with an occluder, the adaptation effect was also very strong (Figure 3). Three of the four observers always perceived the test stimulus to be rotating in the direction opposite to the adapted direction. Observer S.H. was the only one who had occasional reversals in rotation direction during adaptation and, consequently, he showed a slightly weaker adaptation effect (test stimulus rotating in the aftereffect direction 88% instead of 100% of the time). For a control condition, we took advantage of the observation that when the occluder was not explicitly depicted (subjective occluders), perception was not stable, but alternated between two interpretations of depth (see Figure 1C). The 2D motion in the control condition was the same as motion with the explicit occluders. However, after adaptation to the control stimulus for 2 min, none of the observers showed any evidence of an aftereffect (Figure 3B). Note that, in both the test and the control condition, there was only one direction of motion signal in the middle section, which could and did lead to a simple 2D motion aftereffect. However, the simple 2D motion aftereffect could not influence the assignment of depth order to the front or the back surface of the ambiguous test cylinder, as demonstrated by the absence of a rotation aftereffect in the control condition (Figure 3).

The Aftereffect Is Retinotopic and Disparity Specific

The adaptation effect found here is retinotopically specific. In other words, the test pattern must be presented at the same retinal location as the adapting pattern [21, 22]. This retinotopic specificity is evident after adaptation to a rotating cylinder that has been disambiguated by disparity or stabilized by cones or occluders. For example, in Figure 2, the cone-only condition did not generate the adaptation effect. In further tests, the aftereffect was not observed as long as there was no spatial overlap between the adapting and testing stimuli. Moreover, surprisingly, this adaptation effect also requires that the test pattern be placed at the same stereo depth plane as the adapting pattern. The aftereffect disappeared if the adapting and test stimuli were presented with different absolute disparities (Figure 4A). Under such conditions, all observers perceived that the test pattern alternated direction of rotation, with each direction being observed for nearly the same amount of time (black bars in Figures 2 and 3). The retinotopic and disparity specificity of this aftereffect implies that this adaptation occurs relatively early in the visual system when one considers that rotation-sensitive neurons have quite large receptive fields [23]. It is interesting to note that the stabilization of rotation direction, over intermittent presentation [13, 14], seems to be somewhat retinotopic specific but not disparity specific [24].

The aftereffect could originate in mechanisms encoding depth together with translational motion. Alternatively, the aftereffect could be a rotation aftereffect [19]. In the latter case, because the aftereffect was observed only when the test stimuli and adapting stimuli were presented at the same disparity and location, our data suggest that, at the same retinal location, there are separate rotation-sensitive neurons of different disparities. This requirement makes the rotation adaptation model less parsimonious, although theoretically possible. How-

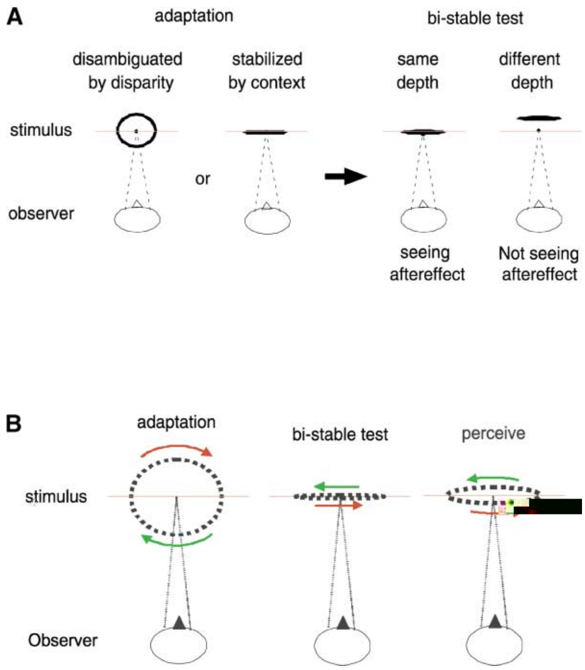

Figure 4. Adaptation Is Depth (Disparity) Specific

(A) The aftereffect was only observed when the test pattern was placed at the same depth plane as the adapting pattern. This was true for both the unambiguous adapting stimuli with disparity and the context-stabilized adapting stimuli.

(B) Illustration of motion direction contingent disparity aftereffects. During adaptation to a cylinder that is rotating clockwise, the dots moving to the left and to the right have different disparities (near and far, or crossed and uncrossed). When test includes moving dots with zero relative disparity (bistable), the leftward-moving dots are pushed away from the observer (green arrow), whereas the rightward-moving dots are pushed closer to the observer (red arrow). As a result, the test pattern is seen as rotating counterclockwise. Note that this aftereffect depends on the existence of different disparities associated with the two motion directions during adaptation.

ever, additional considerations argue against this model. First, an opponent mechanism tuned to rotation would predict that after prolonged adaptation to an unambiguous rotation, one would perceive a static cylinder to rotate in the opposite direction. However, this is not the case [7]. We failed to observe a rotation aftereffect with a static test pattern. Second, neurons responsible for complex motion perception show a large degree of position- and scale invariance [23, 25], but, here, the aftereffect observed was position specific in location and size. Third, the aftereffect is not tied to the structure of the adapting stimulus. We observed that, after adaptation to the stabilized rotating cylinder, a flat sheet of oppositely moving dots with zero relative disparity showed a depth order contingent with the prediction of the disparity adaptation contingent on motion direction.

We favor the interpretation that the aftereffect is a motion direction-contingent disparity aftereffect, similar to that reported by Nawrot and Blake [7] (see Figure 4B). However, the key difference between our results and the results of Nawrot and Blake is that Nawrot and Blake found nonzero relative disparity between the two sets of dots moving in opposite directions, whereas in our experiment the two sets of dots had zero relative disparity. In other words, we believe that the kinetic depth adapted disparity-tuned neurons as if they had nonzero relative disparities. This interpretation implies that, within certain limits, kinetic depth is indeed equivalent to the disparity depth in the sense that the disparity-tuned neurons are selective to depth signals defined by motion. Nawrot and Blake (1993) showed that disparity and kinetic depth could be perceptually metameric [22]. Here, our experiments suggest that the two mechanisms can cross-adapt, which is a stronger indication that they share the same neural mechanisms.

In 2D motion, attentional tracking can induce a motion aftereffect when tested with a dynamic flicker stimulus [26]. Attention was also shown to modulate the adaptation to 3D rotation [27]. Can attentional tracking account for our observation? We tested this possibility by reducing the number of dots in the disparity-defined, unambiguous rotating cylinder while preserving the perception of a rotating cylinder. The logic is that the attention system tracks the direction of rotation, whether there are 600 or 30 dots, but a system that depends on the energy of the motion and disparity signal would be much less stimulated by the 30 dots than the 600 dots. If the aftereffect were due to attentional tracking, then we would expect that tracking 30 dots should also generate an aftereffect. However, we failed to observe an aftereffect when we reduced the number of dots, suggesting that the aftereffect was not due to attentional tracking.

Conclusions

Contextual and pictorial information can disambiguate and stabilize an ambiguous kinetic stimulus. The stabilized ambiguous motion can generate a consistent aftereffect. The aftereffect observed is likely to be a motion direction-contingent disparity aftereffect, originated from the neural equivalence between disparity and motion parallax.

Experimental Procedures

Observers

Two experienced observers (F.F. and S.H.) and two naïve observers (W.L. and J.M.), with normal or corrected-to-normal vision, participated in these experiments. No formal stereovision tests were given to the observers, but all observers could perceive random dot stereograms.

Apparatus and Stimuli

The stimuli were presented stereoscopically with liquid-crystal (LCD) shutter glasses (StereoGraphics Corporation, San Rafael, CA). The moving dots were generated on a PC and presented on a SONY Trinitron Multiscan G420 19 inch monitor, with a spatial resolution of 1280 × 1024 pixels and a refresh rate of 100 Hz. During the experiments, observers wore the LCD glasses with their viewing distance set at 57 cm. The basic stimulus used in the experiments was a rotating cylinder defined with 600 small, randomly placed dots (0.08° × 0.08°). The speed of each dot followed a sine function. The 2D projection of the cylinder subtended 5 deg vertically and 4 deg horizontally. The dots were white (82.1 cm/m^2) against a black background. For conditions in which the cylinder motion was disambiguated by the disparity, disparity was varied smoothly (within the limit of pixel size) from zero disparity at the edge +0.1 (−0.1) deg ee a c di a i a he cen e . The c linde a ed a 0.231 e i n / .

In he fi ada a i n e e imen (Fig e 2), f kind ada ing im li e e ed. The e e (1) a a ing c linde i h c m le e, nambig di a i inf ma i n; (2) a a ing c linde i h n-ambig di a i inf ma i n a i end (i.e., he middle ec i n he e e' im l a em ed f m c ndi i n 1 gene a e c ndi i n 2. The end e e each 1.5° all, and he middle ec i n a 2° all); (3) he end a a ing c linde i h nambig di a i inf ma i n (i.e., he middle ec i n b h e e ' im li e e em ed f m c ndi i n 1 gene a e c ndi i n 3; (4) a bi able a ing c linde . The e e ' im li e e iden ical in hi c ndi i n. The e im l a a bi able, a ing c linde e ending nl 2° e icall ; h , he e im l a nl e en ed in he l ca i n he middle ec i n he ada ing im li. Unde c ndi i n 1 and 2, he bi able e im l a al laced ei he a he ame diffe en de h lane (0.2 deg di a i f all d) a he ada ing im li.

In he ec nd ada a i n e e imen (Fig e 3), he e e e kind ada ing im li. (1) A a ing c linde (i a ame e e he ame a ha in he fi e e imen) i h a checke ed ed/g een ec angle laced behind he f n face and bl cking a e ical ec i n he back face. The ec angle b ended 6.2° e icall and 2.8 deg ee h i n all . P ible af e image e e a ided b he checke c , i ching e e 6 . (2) A e ical ec i n he d m ing in he di ec i n a em ed (i.e., he ec angle in c ndi i n 1 e e changed he backg nd c). The e im l a a bi able c linde e ending 5° e icall . Unde c ndi i n 1, he e im l a e en ed in ei he he ame de h lane a he ada ing im l a a diffe en de h lane (0.2° di a i f all d).

D ing he ada a i n and e e i d, a fi a i n n a laced in b h he cen e he ada ing im l and he cen e he e ing im l , b h a he cen e he m ni .

6. L ng e -Higgin , H.C., and P a dn , K. (1980). The in e e a-i n a m ing e inal image. P c. R. S c. L nd. B. Bi l. Sci. *208*, 385 397.
7. Na , M., and Blake, R. (1989). Ne al in eg a i n inf ma-i n ecif ing c e f m e e i and m i n. Science *244*, 716 718.
8. Sch a , B., and S e ling, G. (1983). L minance c n he e cei ed 3-D c e d namic 2-D di la . B ll P ch n S c *21*, 456 458.
9. Eb , D.W., L mi , J.M., and S n n, E.M. (1989). Pe ce al linkage m l i le bjec a ing in de h. Pe ce i n *18*, 427 444.
10. Gillam, B. (1976). G ing m l i le ambig c n : a d an nde anding face e ce i n. Pe ce i n *5*, 203 209.
11. G mann, J.K., and D bbin , A. (2003). Diffe en ial ambig i ed ce g ing me a able bjec . Vi i n Re . *43*, 359 369.
12. Se en , M.E., and Se en , M.I. (1999). 2-D cen e - nd ef-fec n 3-D c e-f m-m i n. J. E . P ch . H m. Pe -ce . Pe f m. *25*, 1834 1854.
13. Le d, D.A., Wilke, M., Maie , A., and L g he i , N.K. (2002). S able e ce i n i all ambig a e n . Na . Ne ci. *5*, 605 609.
14. Maie , A., Wilke, M., L g he i , N.K., and Le d, D.A. (2003). Pe ce i n em all in e lea ed ambig a e n . C . Bi l. *13*, 1076 1085.
15. B adle , D., Chang, G., and Ande en, R. (1998). Enc ding h ee-dimen i nal c e-f m-m i n b ima e a ea MT ne n . Na e *392*, 714 717.
16. D dd, J., K g, K., C mming, B., and Pa ke , A. (2001). Pe ce all bi able h ee-dimen i nal fig e e ke high ch ce abili ie in c ical a ea MT. J. Ne ci. *21*, 4809 4821.
17. P ffi , D.R., Be en hal, B.I., and R be , R.J., J . (1984). The le ccl i n in ed cing m i abili in m ing n -ligh