Vision Research 98 (2014) 89–98

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Aging and visual length discrimination: Sequential dependencies, biases, and the effects of multiple implicit standards

J. Farley Norman [a,*], Jacob R. Cheeseman [b], Michael W. Baxter [b], Kelsey E. Thomason [b], Olivia C. Adkins [b], Connor E. Rogers [b]

[a] Department of Psychological Sciences & Center for the Study of Lifespan Development, Western Kentucky University, Bowling Green, KY 42101-1030, USA
[b] Department of Psychological Sciences, Western Kentucky University, Bowling Green, KY 42101-1030, USA

ARTICLE INFO

Article history:
Received 13 November 2013
Received in revised form 11 February 2014
Available online 3 April 2014

Keywords:
Length discrimination
Visual memory
Aging
Method of single stimuli

ABSTRACT

Younger (20–25 years of age) and older (61–79 years) adults were evaluated for their ability to visually discriminate length. Almost all experiments that have utilized the method of single stimuli to date have required participants to judge test stimuli relative to a single implicit standard (for a rare exception, see Morgan, On the scaling of size judgements by orientational cues, Vision Research, 1992, 32, 1433–1445). In the current experiments, we not only asked participants to judge lengths relative to a single implicit standard, but they also compared test stimuli to two different implicit standards within the same blocks of trials. We analyzed our participants' judgments to evaluate whether significant sequential dependencies occurred. We found that while individual younger and older adults possessed similar length difference thresholds and exhibited similar overall biases, the judgments of older adults within individual blocks of trials were more strongly biased (than younger adults) by preceding responses (i.e., their judgments on any given trial were more strongly affected by responses to previously viewed stimuli). In addition, the judgments of both younger and older adults were more strongly biased by preceding responses in the blocks of trials with multiple implicit standards. Overall, our results are consistent with the operation of the tracking mechanism described by Criterion-setting theory (Lages and Treisman, Spatial frequency discrimination: Visual long-term memory or criterion setting? Vision Research, 1998, 38, 557–572).

1. Introduction

The method of single stimuli (MSS) has a venerable history. For over a hundred years, researchers have found that participants can make precise discriminations of kinesthetic, tactile, visual, auditory, olfactory, and gustatory stimuli even when no explicit standard is presented (e.g., Fernberger, 1931; Fry, Haupt, & Wartena, 1933; Harris, 1948; Martin & Müller, 1899; McKee, 1981; Morgan, Watamaniuk, & McKee, 2000; Nachmias, 2006; Norman, Holmin, & Bartholomew, 2011; Norman & Todd, 1998; Norman et al., 2008; Pfaffmann, 1935; Treisman & Lages, 2010; Wenzel, 1949; Wever & Zener, 1928). In this method, a participant makes judgments about the magnitude (i.e., weight, length, pitch) of a stimulus – for example, whether an object is heavier or lighter than a standard weight that is never explicitly presented. Difference thresholds obtained from such judgments can be just as low (or lower) as those obtained when using the conventional method of constant stimuli (e.g., Morgan, Watamaniuk, & McKee, 2000; Nachmias, 2006; Norman & Todd, 1998; Norman, Holmin, & Bartholomew, 2011).

As Morgan, Watamaniuk, and McKee (2000, p. 2342) have pointed out "to make a judgment in the MSS, observers must use some representation or memory of the stimuli they have been shown before". According to the results of their model, human observers derive their knowledge of the implicit standard in the MSS from the average of the magnitudes of the 10–20 most recently presented test stimuli (also see Dyjas, Bausenhart, & Ulrich, 2012). For each test stimulus, the participant can then judge whether its magnitude is greater or less than that running average. A second possibility (Criterion-setting theory, CST) has been developed by Treisman, Lages, and colleagues (e.g., see Lages & Paul, 2006; Lages & Treisman, 1998, 2010; Treisman & Williams, 1984). In this view, what is stored in memory (and used for categorization or discrimination) is a response or decision criterion (e.g., see Macmillan, 2002; Vogels & Orban, 1986). According to CST, a

* Corresponding author. Address: Department of Psychological Sciences, 1906 College Heights Blvd., #21030, Western Kentucky University, Bowling Green, KY 42101-1030, USA.
E-mail address: Farley.Norman@wku.edu (J.F. Norman).

http://dx.doi.org/10.1016/j.visres.2014.03.011
0042-6989/

participant's response criterion varies across trials within a session under the influence of stabilization and tracking mechanisms that serve to optimize performance. The stabilization mechanism adjusts the criterion in such a manner that responses on any given trial are negatively affected by preceding *stimuli* (e.g., if a participant is judging length, a prior "longer" stimulus would reduce the probability of responding "longer" on a subsequent trial), while the tracking mechanism adjusts the criterion such that responses are positively affected by preceding *responses* (e.g., a prior response of "longer" would increase the probability of responding "longer" on a subsequent trial). In an experiment using the method of single stimuli to investigate spatial frequency discrimination, Lages and Treisman (1998) found strong sequential dependencies (see their Fig. 4) in the direction predicted by Criterion-setting theory.

Almost all of the psychophysical studies conducted over the past century using the method of single stimuli have asked participants to judge test stimuli relative to a single implicit standard within individual blocks of trials (as was mentioned earlier, participants can perform such judgments with the same or better precision as when explicitly presented standards are available). Interestingly, Morgan and colleagues (Morgan, 1992; Morgan, Watamaniuk, & McKee, 2000) have also demonstrated that human observers can successfully compare individual test stimuli to a variety of different standard magnitudes within single blocks of trials (4 and 8 different standards within a block in the experiments of Morgan, 1992; 9 different standards within a block in Morgan, Watamaniuk, & McKee, 2000). In Experiment 1 of Morgan (1992), for example, participants judged the magnitude of spatial separations between a single pair of parallel lines on any given trial (i.e., judged whether the separation was larger or smaller than a standard value). The orientation of the lines on each trial indicated which standard was to be used for comparison. Morgan (1992) found that two experienced psychophysical observers could effectively compare the test separations with multiple implicit standards within a block with no loss of precision (compared to judgments made with respect to a single implicit standard). When the number of multiple implicit standards within a block was increased to nine, Morgan, Watamaniuk, and McKee (2000, see their Fig. 4) found reductions in precision (difference thresholds increased by up to a factor of two) for two additional highly experienced psychophysical observers.

In the current study, we examined the effects discussed here for a group of twenty younger and older adults, none of whom were psychophysically experienced observers. First of all, we wanted to determine the extent to which inexperienced observers can compare test stimuli against multiple implicit standards within single blocks of trials. It is known from the results of Morgan (1992) and Morgan, Watamaniuk, and McKee (2000) that experienced observers can effectively perform such judgments with only modest (or sometimes no) reductions in precision. Even if younger inexperienced observers can effectively perform such judgments, it is not clear whether older adults (in our study, their ages ranged from 61 to 79 years) could do the same. In our previous research concerning aging and visual 3-D shape discrimination, we have often found that while older adults can perform similarly to younger adults in some circumstances, their performance suffers disproportionately when tasks become challenging and difficult (e.g., see Norman, Dawson, & Butler, 2000; Norman et al., 2012, 2013). In addition, as described earlier, researchers (Lages & Treisman, 1998; Vogels & Orban, 1986; also see Ward, 1982) have found significant sequential dependencies, strong effects of prior stimuli and responses upon responses in subsequent trials. The experiments of Morgan (1992) and Morgan, Watamaniuk, and McKee (2000) did not investigate sequential dependencies – while observers are judging test stimuli relative to multiple standards within a block of trials, do sequential dependencies exist? If so, are they greater in magnitude than those that occur when observers judge test stimuli relative to single standards? The purpose of the current study is to answer these questions. The abilities of younger adults were investigated in the current Experiment 1, while those of older adults were evaluated in the current Experiment 2.

2. Experiment 1

2.1. Method

2.1.1. Apparatus

The stimulus displays were generated by an Apple PowerMacintosh G4 computer and were presented on a 22-in. Mitsubishi Diamond Plus 200 monitor. The resolution of the monitor was 1280×1024 pixels. The viewing distance between the participants and the monitor was 100 cm.

2.1.2. Experimental stimuli

The experimental stimuli were yellow and blue antialiased line segments (Foley et al., 1996; pp. 132–137) presented in the frontoparallel plane, against a black background. The longer implicit standard was 9.0 cm, while the shorter implicit standard was 6.0 cm (the same standard lengths as those used by Norman et al., 1996). For each standard, there were a total of six test lengths (whose lengths were invariant across orientation). Three of the test lengths were physically shorter than the standard (by 1.6%, 4.8%, and 8.0%), while three were longer than the standard (also by 1.6%, 4.8%, and 8.0%). Thus, the absolute test lengths were 8.28, 8.57, 8.86, 9.14, 9.43, and 9.72 cm for the longer standard and 5.52, 5.71, 5.90, 6.10, 6.29, and 6.48 cm for the shorter standard.

2.1.3. Procedure

In this experiment, there were two conditions. In the initial condition (300 trials/participant, 2 blocks of 150 trials; 50 total trials for each of the 6 test lengths, all presented in a random order), the participants compared test lengths with the longer (i.e., 9.0 cm) implicit standard. A single yellow test line segment was presented for 2.0 s on each trial. Each test line possessed a random orientation and was randomly offset from the center of the display (both horizontally and vertically) by up to 5 cm (2.9 deg visual angle). The participants' task was to judge whether each test line was longer or shorter than the implicit standard (which they never saw explicitly); if their judgments were correct, the participants received feedback (during both experimental and practice blocks) in the form of a short auditory beep. It has been repeatedly demonstrated that performance is identical regardless of whether feedback is or is not provided (e.g., see Fig. 2 of Morgan, Watamaniuk, & McKee, 2000 and Fig. 3 of Norman, Holmin, & Bartholomew, 2011). As was the case in the experiments of Norman and Todd (1998) and Norman et al. (2008, 2011), the participants were given 20 practice trials (with feedback) at the beginning of each block of 150 trials; Morgan, Watamaniuk, and McKee (2000) have demonstrated that participants' knowledge of the standard magnitude is derived from the running average of the most recent 10–20 trials. The 20 practice trials in the current study, therefore, gave our participants the opportunity to effectively learn the standard length (or appropriate response criterion in CST) before beginning the experimental trials.

In the second (i.e., subsequent) experimental condition, the procedures were the same as those used in the initial condition. The only difference was that the participants compared the test lengths against two implicit standards (6.0 and 9.0 cm) within the same blocks of trials (600 trials/participant, 2 blocks of 300 trials; 50 total trials for each of the 6 test lengths for each of the 2 implicit standards, all presented in a random order). On any given trial