

The effect of optotype presentation duration on acuity estimates revisited

Sven P. Heinrich · Katja Krüger · Michael Bach

Received: 27 July 2009 / Revised: 2 December 2009 / Accepted: 7 December 2009 / Published online: 14 January 2010
© Springer-Verlag 2010

Abstract

Background A high reproducibility of visual acuity estimates is important when monitoring disease progression or treatment success. One factor that may affect the result of an acuity measurement is the duration of optotype presentation. For times below 1 s, previous studies have convincingly shown that acuity estimates increase with presentation duration. For durations above 1 s, the situation is less clear.

Methods We have reassessed this issue using the Freiburg Visual Acuity Test with normal subjects. Presentation durations of 0.1 s, 1 s, and 10 s were assessed.

Results Confirming previous findings, in all subjects acuity estimates in the 1-s condition were higher than those in the 0.1-s condition, on average nearly by a factor of 2, equivalent to 3 lines. However, in 12 out of 14 subjects, acuity estimates increased further with a presentation duration of 10 s, on average by 23% ($P=0.002$), or roughly 1 line. Test–retest variability improved by 49% ($P=0.003$). These findings can be explained by a simple statistical model of acuity fluctuations. Cognitive processing may also be a relevant factor. Interestingly, most observers subjectively felt that they could perceive the optotypes best in the 1-s condition.

Conclusion The results highlight the importance of standardizing presentation durations when high reproducibility is required.

Keywords Visual acuity · Exposure duration · Presentation duration · Optotype · Acuity test

Introduction

Estimating visual acuity is a standard procedure that is performed as part of most optometric or ophthalmological exams, and is an important foundation for an expert's opinion on the visual status of a person. One of the factors affecting the outcome of an acuity test is the time available to the examinee to observe an optotype before he or she has to decide on a response. Over the last 80 years, several studies have investigated this issue, albeit mostly not for a range of times relevant for routine diagnostics, as the authors were interested in the so-called critical duration, i.e. in retinal integration, which most likely is not the only factor determining the effect of presentation duration on acuity.

Graham and Cook [1] assessed acuity estimates as a function of presentation duration, but not for durations longer than 1 s. Monjé and Schober [2] found a sizable increase in acuity with presentation duration in the interval of 10 ms to 50 ms, which is far below those durations typically used in practice. Zanen and Klaassen-Nenquin [3] did not find a difference between a 1-s presentation and a “pause” condition. The latter condition was not defined in detail in the article, but appears to involve a presentation duration longer than 1 s. Schwarz [4, 5] did not measure acuities, but instead varied the presentation duration for a given stimulus size to find the minimum time necessary to

S. P. Heinrich (✉) · M. Bach
Sektion Funktionelle Sehforschung, Univ.-Augenklinik,
Killianstr. 5,
79106 Freiburg, Germany
e-mail: sven.heinrich@uniklinik-freiburg.de

K. Krüger
Fakultät für Biologie, Albert-Ludwigs-Universität Freiburg,
Schänzlestr. 1,
79104 Freiburg, Germany

identify the stimulus. While related to the present question, it is difficult to interpret the results in terms of the effect of presentation duration on acuity. This is also the case for a study by Møllenbach as cited by Ehlers [6]. All of these authors report that a longer time is required for smaller optotypes. Gerbstädt [7] performed similar experiments. However, she performed additional analyses that imply higher acuity values for 2-s durations than for 1-s durations, although without statistical assessment.

The more recent literature also does not provide a fully consistent picture. Baron and Westheimer [8] found an increase in measured acuity up to 400 ms “and possibly longer”, but they did not use a forced-choice procedure. Kono and Yamade [9] found that acuity estimates in healthy subjects increase up to 1 s, and von Boehmer and Kolling [10] report an increase up to 1 s, with a non-significant trend beyond. Ng and Westheimer [11] tested the range of 33–300 ms with a white-on-black stimulus, and also found an increase. Westheimer [12] estimates that at least 0.5 s are necessary to achieve maximum acuity.

A quantitative comparison of previous studies confirms the lack of consistency. For the difference between 0.1 s and 1 s, for instance, Zanen and Klaassen-Nenquin [3] report an increase of around 10–15% (their Fig. 4), while Baron and Westheimer [8] (in their various figures) show an increase of the order of 50% to 100%, and Kono and Yamade [9] found an increase of about 100% (their Fig. 2).

Ehlers [6] presents empirical data suggesting that patients with different diseases require different presentation times for reaching maximum acuity. This does not appear to depend on whether or not the retina is the locus of the disease. The details of how these data were collected are not revealed, though. Monjé and Schober [2] suggested the use of short presentation durations to identify certain types of retinal diseases. In patients with retinal diseases, as opposed to normal subjects, Kono and Yamade [9] found acuity estimates to increase beyond 1 s.

On top of the limitations and inconsistencies of previous studies, there are theoretical considerations that suggest an increase of estimated acuity with presentation duration. It is not obvious why these would cease to have an effect above 1 s. For instance, fluctuations in accommodation and pupil diameter may modulate retinal image quality. The likelihood of getting a snapshot of the optotype under optimal conditions increases with prolonged presentation time, as the temporal dynamics of the underlying processes includes components below 1 Hz [13, 14]. Changes in tear-film properties may also contribute [15]. It seems furthermore likely to us that conscious pondering of different possible interpretations of the percept would help to decide on the correct response. This would certainly be facilitated by long presentation durations. The idea of a cognitive factor aiding

in optotype recognition is supported by experiments by Kono and Yamade [9]. They found that some patients with retinal diseases need several seconds to reach the maximum acuity estimate, much longer than the normal subjects needed in their study. While retinal pathophysiology as such seems to be an unlikely cause of a several-second difference, the effect may well be explained by a need to inspect a degraded stimulus more thoroughly than a stimulus that is not altered by disease. Interestingly, Ehlers [6] already appears to imply post-retinal mechanisms when speculating that “the collection of these more diffuse stimuli into the recognition of the right shape will then require a longer time.”

Following these considerations, we have systematically reassessed the influence of presentation duration over a time range of 0.1 s to 10 s, with the difference between 1 s and 10 s being particularly relevant for routine applications.

Methods

Subjects

Fourteen subjects (age 20–58, diverse educational and work backgrounds) with no known ophthalmological disorders participated in the experiment after providing written informed consent. The study was approved by the local review board, and adhered to the tenets of the Declaration of Helsinki. The number of 14 participants is in the typical range used in psychophysical investigations. We reasoned that this would be sufficient if the effects are large enough to be relevant in clinical practice.

Stimuli and procedure

We used the Freiburg Acuity and Contrast Test (FrACT [16], see also <<http://www.michaelbach.de/fract/>>) to obtain acuity estimates with different presentation durations. FrACT presents Landolt Cs on a computer screen. The subjects watched the stimuli monocularly with their habitual correction, and responded by pressing a corresponding key on a keypad, thereby indicating the perceived (or guessed) orientation of the Landolt C. Starting with a gap size of 30 arcmin, the size of subsequent presentations is controlled by the ‘best PEST’ algorithm [17].

FrACT ran on an iMac computer with an LCD screen that was located at a distance of 4.5 m from the subject. The optotypes were presented with a luminance of 1.2 cd/m² on a background of 118 cd/m² for either 0.1 s, 1.0 s, or 10 s. The subjects were instructed not to press a response key before the disappearance of the stimulus. Each run of the

experiment, yielding one acuity estimate, consisted of 30 trials with the same presentation duration. Three runs with different presentation durations were performed in random order. The runs were then performed again in reverse sequence, resulting in a total of two runs per presentation duration, the results of which were averaged geometrically. After the experiment, we also asked subjects to make a conjecture as to which of the three conditions allowed for the best performance.

Statistical analysis

Non-parametric statistical tests were applied to logMAR values. Because our hypothesis was that acuity increases with presentation duration, we applied one-sided tests when assessing the significance of differences. However, as can be estimated from the P values presented below, all tests that indicated significance would have done so in a two-sided version as well.

Results

In all subjects, acuity estimates obtained with a 1-s presentation were substantially higher than those obtained with a 0.1-s presentation (Fig. 1; 0.29 logMAR=94%; Wilcoxon test, $P<0.0001$). Between 1 s and 10 s, the increase was less drastic, but still present in 12 of the 14 subjects, and significant at the group level (0.091 logMAR=23%;

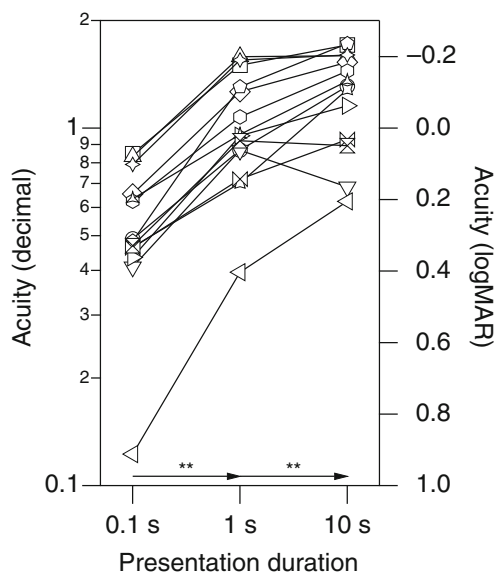


Fig. 1 Individual acuities, averaged across repetitions. In all subjects, estimated acuity increased between durations of 0.1 s and 1 s. In 12 of the 14 subjects, a further increase was found with 10 s. **, significant at $\alpha<0.01$

$P=0.0020$). Bonferroni correction suggests that both comparisons are significant at a family-wise α of 0.01.

We next asked whether the test–retest variability of the acuity estimate, computed as the difference between the runs with the same presentation duration (not taking into account the sign of the difference), improved with presentation duration (Fig. 2, left). On the group level, this was the case between durations of 1 s and 10 s (median 0.075 logMAR vs 0.041 logMAR; Wilcoxon test, $P=0.0026$, significant at a family-wise α of 0.01 with Bonferroni correction), but not to a significant degree between durations of 0.1 s and 1 s (median 0.060 logMAR vs 0.075 logMAR, $P=0.097$). However, looking not at the group median, but rather at the distribution of difference values (Fig. 2, right), and taking into account the sign of the difference, the total range of differences spanned by all subjects (1-s condition versus 10-s condition) increased, and the reduction of the range spanned by the central six subjects was not significant (bootstrap test, $P=0.36$). The 95% limits of agreement between test runs as suggested by Bland and Altman [18] cannot be reliably estimated non-parametrically due to the small sample size. A parametric approximation yields ± 0.19 logMAR for the 0.1-s condition and ± 0.13 logMAR for both the 1-s and 10-s conditions. Wondering whether there is evidence of a practice effect, we also assessed whether differences between runs were biased towards an increase in estimated acuity. This was the case in the 1-s condition (Wilcoxon test, $P=0.0046$, significant with Bonferroni correction), but not in the 0.1-s and 10-s conditions ($P=0.43$ and $P=0.88$ respectively).

Eleven of the subjects were asked after the experiment, but before they were informed of the results, to estimate in which condition they had performed best (Fig. 3). Ten subjects decided for the 1-s condition, despite eight of them performing best in the 10-s condition. Consistently, subjects attributed their difficulties in the 10-s condition to the percept of a switching or rotating optotype. One subject who performed best in the 10-s condition did not feel a difference between the 1-s condition and the 10-s condition. The discrepancy between subjective judgements regarding the condition with best performance and the objective findings was significant (Wilcoxon test, $P=0.0020$).

Discussion

We found the acuity estimates obtained with 1-s durations to be nearly twice as high as those obtained with 0.1-s durations. This compares well to previous reports by Baron and Westheimer [8] and Kono and Yamade [9]. More interestingly, we also found acuity estimates with a presentation duration of 10 s to be sizably higher than

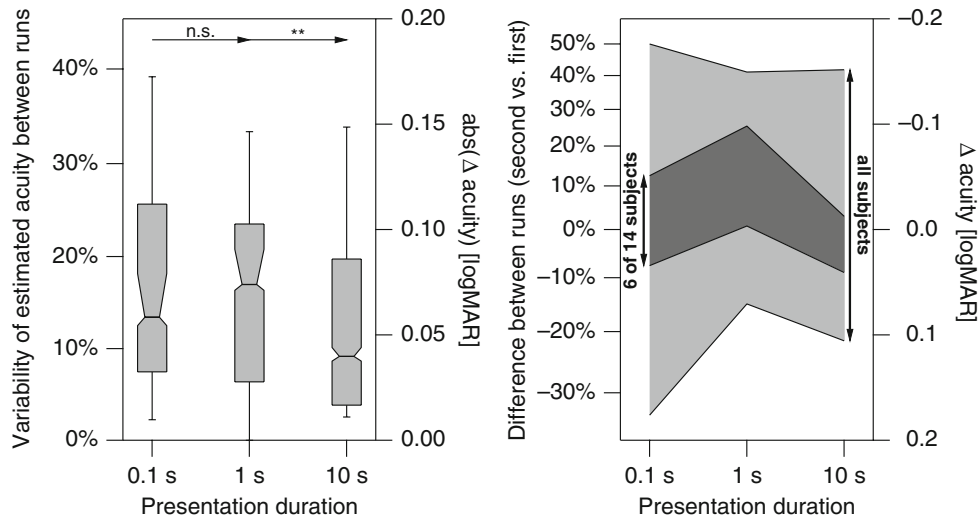


Fig. 2 *Left*: variability between the two test runs, not taking into account the sign of the differences. On the left axis, the difference is given as percentage of the mean of the two test results. The logMAR difference is displayed on the right axis. The boxplot shows the median (narrowest part of the box), the interquartile range (total extent of the box) and the total range (*whiskers*). The 67% confidence interval (non-parametric equivalent of the standard error) is indicated by the extent of the notch in the box. For the 10-s condition,

variability was significantly smaller than for the 1-s condition. **, significant at $\alpha < 0.01$; *n.s.*, not significant. *Right*: range of differences between runs (signed values). The *uppermost and lowermost lines* represent the maximum increase and maximum decrease of measured acuity, respectively, between the first and the second run in the present sample of subjects. The *dark gray area* indicates the quantile range spanned by the central six subjects

those with a presentation duration of 1 s, which is in contrast to previous reports, which only found a small non-significant trend [10] or no effect [9]. There are no obvious methodological reasons for this diversity, as von Boehmer

and Kolling also used the FrACT (in an earlier, but geometrically identical version). Our findings are consistent, however, with the results by Gerbstädt [7], which imply that acuity estimates can increase beyond presenta-

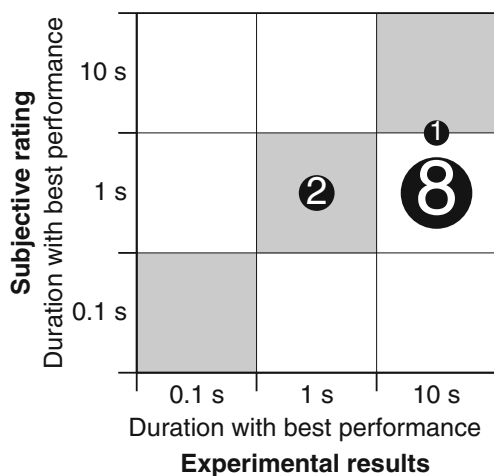


Fig. 3 Subjective ratings versus experimental results. Disk area and digits represent the number of subjects with a certain combination of subjective rating and experimental result. The one disk on the border between fields represents a subject that rated 1-s and 10-s presentations as equal. The *gray fields* on the diagonal represent correspondence between ratings and experimental results. Most subjects believed that they performed best with the 1-s duration. However, in most of them the highest acuity estimates were obtained with the 10-s duration

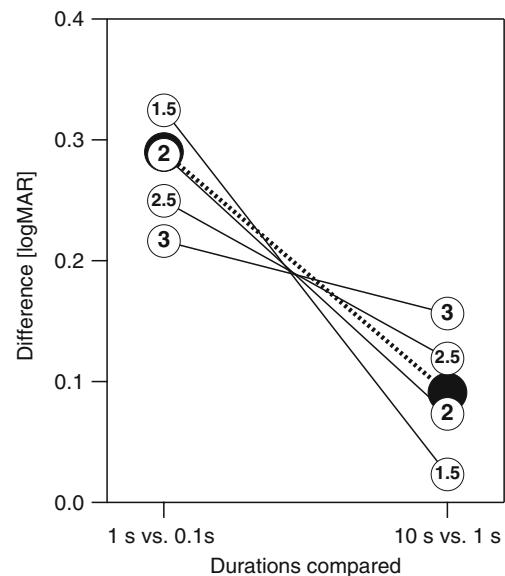


Fig. 4 *Thin lines*: simulation results, assuming a ± 0.4 -logMAR range of acuity fluctuations with different periods as indicated for each line (measured in seconds). *Thick dotted line with filled markers*: experimental results. Values indicate the differences between conditions as indicated on the abscissa. The experimental results are consistent with a fluctuation period between 2 s and 2.5 s

tion durations of 1 s. We found an increase in 12 out of 14 subjects, and based on the statistical testing, we can reject the null hypothesis of no difference between 1-s and 10-s presentation durations with rather high confidence ($P < 0.0020$). A detailed comparison with earlier studies is difficult because most of those studies did not show data of individual subjects, or included only a very small number of subjects. Statistical analysis was lacking in many of the older articles.

Retinal luminance integration operates on a time scale of around 50–100 ms (Bloch's law [19]) and is therefore unlikely to be the dominant factor. While we cannot exclude any stage of visual processing as the locus of the effect, a time scale in the range of seconds hints towards cognitive factors. As already discussed in the Introduction, it seems likely to us that prolonged inspection of the optotype would help to decide on the correct response. A second, more trivial, factor may also play a role. A longer inspection time means that there is an increased chance that a moment of optimal retinal image quality is achieved during fluctuations in, for instance, accommodation and pupil diameter, which at least partly occur at frequencies below 1 Hz [13, 14].

In order to verify whether such fluctuations could produce results similar to those reported here, we implemented a simplified model and performed a Monte Carlo simulation [20]. We approximated acuity fluctuations as a sine wave with a certain amplitude (strength of fluctuations) and a certain period. The phase of the sine wave was randomly varied. We determined the maximum acuity obtained during an optotype presentation, which could last for 0.1 s, 1 s, or 10 s as in the real experiment. For each presentation duration, 1,000 presentations were simulated, and the respective maximum acuities were averaged. Figure 4 shows the differences between presentation durations for a fluctuation strength of ± 0.4 logMAR. Larger or smaller fluctuation strengths would shift the lines up or down. Simulations with a period between 2 s and 2.5 s most closely reproduce the experimental results. These values are well within the range reported as typical frequencies for fluctuations in accommodation and pupil size. Thus, despite the model employed being very simplistic and taking into account neither retinal integration [11] nor cognitive factors, the results suggest that the experimental findings could potentially be explained by fluctuations in acuity during optotype inspection. Because the simulation used a sinusoidal model of acuity fluctuations, quantitative results may not be exact. The qualitative outcome, however, does not depend on the shape of the fluctuations.

In the present study, test–retest variability was 0.041 logMAR to 0.075 logMAR (median values), depending on

the presentation duration, with the 95% limits of agreement in the range of 0.13 logMAR (1 s and 10 s) to 0.19 logMAR (0.1 s). This is the same order of magnitude as reported by Arditi and Cagenello [21] for different acuity charts. Because the average increase in acuity by 0.91 logMAR between the 1-s condition and the 10-s condition exceeds the median variability, the presentation duration may be a relevant factor in practical applications. The main conclusion of the present study is therefore to set a fixed time available to the examinee for viewing an optotype. This time should in particular be constant across follow-up measurements, in order to avoid spurious changes of the estimated acuity. While we found 10 s to result in an improved test–retest variability, a 1-s duration appears to be the better choice considering temporal constraints in routine applications. A shorter presentation duration also offers the option of performing more measurements in the same time to improve the reliability, although part of the larger variability with the 1-s duration seems to be the result of a practice effect.

Acknowledgements This study was supported by the Deutsche Forschungsgemeinschaft (BA 877/18). We are grateful to our subjects for their participation.

References

- Graham CH, Cook C (1937) Visual acuity as a function of intensity and exposure-time. *Am J Psychol* 49:654–661
- Monjé M, Schober H (1950) Vergleichende Untersuchungen an Sehproben für die Fernvisusbestimmung. *Klin Monatsbl Augenheilkd* 117:561–570
- Zanen J, Klaassen-Nenquin E (1957) Acuité visuelle en fonction du temps d'exposition. *Bull Soc Belge Ophtalmol* 114:574–581
- Schwarz F (1947) Der Einfluß der Darbietungszeit auf die Erkennbarkeit von Sehproben. *Pflügers Arch* 249:354–360
- Schwarz F (1951) Neue Sehschärfenmessungen: Die Prüfung der Sehleistung unter Berücksichtigung der Darbietungszeit der Sehproben. *Graefes Arch Ophthalmol* 151:714–724
- Ehlers H (1948) On visual velocity. *Acta Ophthalmol* 26:115–121
- Gerbstädt U (1949) Der Einfluß der Sehprobengröße auf die minimale Darbietungszeit. *Pflügers Arch* 251:559–570
- Baron WS, Westheimer G (1973) Visual acuity as a function of exposure duration. *J Opt Soc Am* 63:212–219
- Kono M, Yamada S (1996) Temporal integration in diseased eyes. *Int Ophthalmol* 20:231–239
- von Boehmer H, Kolling GH (1998) Zusammenhang zwischen Sehschärfe und Darbietungszeit einzelner Landoltringe bei Normalpersonen und bei Nystagmus-Patienten. *Ophthalmologe* 95:717–720
- Ng J, Westheimer G (2002) Time course of masking in spatial resolution tasks. *Optom Vis Sci* 79:98–102
- Westheimer G (1987) Visual acuity. In: Moses RA, Hart WM (eds) *Adler's physiology of the eye*, 8th edn. Mosby, St. Luis, pp 415–428
- Charman WN, Heron G (1988) Fluctuations in accommodation: a review. *Ophthalmic Physiol Opt* 8:153–164

14. Stark L, Campbell FW, Atwood J (1958) Pupil unrest: an example of noise in a biological servomechanism. *Nature* 182:857–858
15. Montés-Micó R (2007) Role of the tear film in the optical quality of the human eye. *J Cataract Refract Surg* 33:1631–1635
16. Bach M (1996) The “Freiburg Visual Acuity Test” — Automatic measurement of the visual acuity. *Optom Vis Sci* 73:49–53
17. Liebeman HR, Pentlant AP (1982) Microcomputer-based estimation of psychophysiological thresholds: the best PEST. *Behav Res Methods Instrument* 14:21–25
18. Bland JM, Altman DG (1999) Measuring agreement in method comparison studies. *Stat Methods Med Res* 8:135–160
19. Bartlett NR (1965) Thresholds as dependent on some energy relations and characteristics of the subject. In: Graham CH (ed) *Visual perception*. Wiley, New York, pp 154–184
20. Metropolis N, Ulam S (1949) The Monte Carlo method. *J Am Statist Ass* 44:335–341
21. Arditi A, Cagenello R (1993) On the statistical reliability of letter-chart visual acuity measurements. *Invest Ophthalmol Vis Sci* 34:120–129