Review of Psychology,
2010, Vol. 17, No 1, 39-42 UDC 159.9

Sequential effects are not trivial context effects in psychophysical research

ANJA PODLESEK

Sequential effects are a type of context effects in which the response to the stimulus is affected by both, stimuli presented and responses obtained on previous trials. In psychophysical studies, sequential effects are expressed through time-order errors in esthesiometry and through assimilation and contrast effects in scaling. Although these effects are robust, they are too often neglected in psychophysical research. The paper discusses whether the usual procedure with which we intend to control them, namely randomized trials, can really eliminate these effects. Possible consequences of sequential effects in psychological measurement in general are further discussed, and in the end a simple solution on how to minimize them is proposed.

Key words: psychophysics, psychological measurement, context, sequential effects, assimilation

Psychophysical judgements depend on prior events. When the response to the stimulus is affected by both, stimuli presented and responses obtained on previous trials, we speak of sequential effects. Although many psychology textbooks point out to the importance of sequential effects, most of the studies simply neglect them. In Lockhead's (1992) words, a common assumption in psychophysical scaling models is that "attribute judgments are independent of time". We can easily spot that attribute judgments are not independent of space (as in simultaneous lightness contrast) or of other attributes (as in the dependence of lightness perception on perception of depth), but sequential effects are not that obvious. They can only be observed through special data analysis following the measurements, which is probably why they remain neglected.

In the times of classical psychophysics when discrimination methods were mostly used, several studies revealed the existence of time errors on stimulus comparisons (see Guilford, 1954, for review). When stimuli are presented for comparative judgements, the second stimulus in the pair is systematically judged as greater or lesser than we should expect. For example, when the standard is compared with itself, there is an excess of judgements 'greater' for the second stimulus. This is a negative time-order error. Koehler (as cited in Guilford, 1954) found that the time-order error depends on the time interval between stimuli, changing from the positive error at short interstimulus intervals to the negative one at longer interstimulus intervals. He explained the negative time-order error as a consequence of 'reduced' memory trace of the first stimulus. When this trace is compared with the subsequent stimulus, the trace seems to have lower intensity. Lauenstein (as cited in Guilford, 1954) attributed the effect to the assimilation of the trace from the first stimulus to the level of excitation between the onsets of the first and the second stimulus. Guilford (1954) wrote that it seems that the gist of the first stimulus in the pair regresses towards the central value of all the stimuli in the experiment, and that such regression effects may also be found in scaling methods whenever there is a lack of perfect discrimination.

Later, when Stevens' new psychophysics and direct scaling was in the full swing, novel consequences of sequential effects became apparent. Holland and Lockhead (1968) asked people to make absolute judgements of the intensities of 10 tones varying in amplitude. They found that the mean response error was a function of the stimulus that occurred k trials earlier. Stimuli were overestimated when the prior stimulus was large and underestimated when the prior stimulus was small. In general, judgments tended to be similar to the value of the prior trial, which is known as assimilation, and they tended to be different from the stimuli

Anja Podlesek, Department of Psychology, Faculty of Arts, University of Ljubljana, Aškerčeva 2, 1000 Ljubljana, Slovenia,
E-mail: anja.podlesek@ff.uni-lj.si (the address for correspondence).

that occurred earlier in the sequence (two to five or more steps back), which is known as contrast. Later, most studies included both the previous signals and the previous responses in the regression equation when trying to predict the response to the current stimulus (see e. g. De Carlo & Cross, 1990; Schifferstein & Frijters, 1992; Ward, 1982).

The assimilation effect of the immediately preceding trial seems to be quite robust. It was found in magnitude estimation (DeCarlo, 2003; Ward, 1973), absolute identifications (Holland & Lockhead, 1968) and successive-ratio-judgments (Lockhead & King, 1983), and somewhat smaller for cross-modality matching (Baird, Green, & Luce, 1980; Ward, 1975) and the method of production (DeCarlo, 2003; Green, Luce, & Duncan, 1977). Lockhead (1992) even stated: "There is assimilation in every set of psychophysical scaling data that has been examined and reported, no matter what the experimental procedure." (p. 550).

Because Ward and Lockhead (1971) found assimilation in guessing studies where no stimuli were presented, Lockhead (1992) mostly attributed assimilation to response bias. In a successive-ratios-judgment task, Lockhead and King (1983) found that when two tones of equal intensity were presented in succession, their ratio was estimated higher than 1 when the stimulus just prior to the current pair was less than the intensity of the pair, and lower than 1 when the immediately preceding stimulus was greater than the intensity of the pair. Their interpretation was that the first stimulus in the pair must have assimilated in memory toward the immediately preceding tone, and the second tone in pair was compared to this biased memory. They explained contrast effect of stimuli presented more than one step prior to the estimated stimulus as being associated with response adjustments made by observers to correct for the errors caused by such assimilation.

Although the first studies on sequential effects in ratio methods (e. g., Holland & Lockhead, 1968; Wagner & Baird, 1981) reported that sequential effects extend over as many as five trials, it was shown later (see Jesteadt, Luce, and Green, 1977) that these effects are almost exclusively due to the signal and response on the previous trial and that there do not seem to be any sequential effects resulting from even earlier trials. When the effects of response on trial *N*–1 was factored out, no effect of events on trials *N*–2 and before had any direct influence on the response on trial *N* anymore. Gregson (1992) stated that if psychophysical judgments are modelled as a dynamic system rather than as a static one, the occurrence of apparent assimilation together with contrast in the same data series is to be expected whenever some feedback and instability are involved.

At the present time, the dilemma about how sequential effects should be modelled remains a challenge. Another important question is: Are these effects large enough and therefore devastating to the results of psychophysical measurements?

There is no straightforward answer to this question. One of the reasons for this is that sequential effects are not constant over different stimuli. Previous stimuli affect the judgment of the current stimulus in different ways, depending on their size. Ward and Lockhead (1971) reported that an extreme stimulus on the immediately preceding trial "attracts" the response by 0.4 of a response category, whereas assimilation effect was much smaller for the stimuli in the middle of the range of intensities. Other studies (e.g., Baird et al., 1980; Green et al., 1977; Jesteadt et al., 1977) indicated that the magnitude of stimulus presented on trial *N* may not be the most important factor for the size of sequential effects. Instead, they found that sequential effects are strong when the difference between the two successive stimuli is small, but are much weaker when this difference is large. They claimed that assimilation effect is operating only in cases where the signal falls within the attention band that tends to be located in the region of the immediately preceding signal.

Lockhead (1992) argued that each judgment (he was referring to psychophysical judgment, but we can easily extend this to judgments in other contexts) depends on so many different factors "that any prospect of removing their effects to reveal a true, underlying function is remote" (p. 549). Teghtsoonian (1992), on the other hand, questioned the importance of context factors, at least for describing psychophysical relations. He claimed that no investigator believes that judgments of univariate stimuli are independent of space, time, and other features. Instead, those factors are important, but they are not equally important as the stimulus intensity. Intensity is "so far above all other controlling factors that it is easy to identify it as the only pertinent factor" (p. 580).

Typical researchers in experimental psychology are aware of the fact that there are many validity issues regarding their methods of measurement. They are aware of the fact that psychological phenomena are too complex to be studied extensively and validly in a laboratory. Most of them would probably agree with Lockhead (1992) that different attributes of objects simply cannot be judged in isolation. They would probably agree that sequential effects should always be taken into account, because a mass of studies convincingly shows how non-trivial they are. So how can we take these effects into account?

One of the options is to include these non-sensory factors in the psychophysical functions as additional parameters (Dzhafarov, 1992). However, it is doubtful whether we are able to determine the exact quality or magnitude of sequential effects, because we cannot identify all of the participant's previous experiences that may be important for the present judgment. In his adaptation theory, Helson (1964) claimed that adaptation level which becomes the point of reference for the psychological scale according to which the stimuli are judged, is affected both by other stimuli present (the background stimuli and other contextual stimuli) and the experiences obtained prior to the experiment. Therefore, a series of pre-experimental adaptations should always be considered when forming a comprehensive frame of reference for the adaptation level, but this is extremely hard, if not impossible, to accomplish.

Other options are to counterbalance or randomize the trials to control for the sequential effects in the long run. With counterbalancing and randomization we can be comforted that in extensive experiments sequential effects are neutralized. But extensive experiments are not something that our participants could put up with easily. Who on Earth is able to do over three thousand trials, as G. T. Fechner (1860/1966) did, and maintain a constant level of motivation and avoid fatigue? In reality, experiments are limited in time, which is inevitably leading to the question whether the point at which data collection was terminated was the right one and will it allow a decent intra-personal neutralisation of sequential effects. Luckily, we include many participants in the study and these effects are then inter-personally neutralised. This may be satisfactory in nomothetic studies. But when one is doing an idiographic study, sequential effects may become extremely important. Books on qualitative research often mention that context is important, such as historical and socioeconomic context, and social setting. In this collection of context factors we should definitely add short- and long-range temporal context. Personal history and experiences, and also their temporal order, define each individual in a unique way. It seems that sequential effects are not trivial indeed, and they should be paid more attention to in studies with a smaller number of participants who are not subjected to extensive measurements within different conditions.

Even in nomothetic studies sequential effects sometimes cannot be avoided and they can have substantial impact on the results. Sequential effects can be avoided in a carefully planned study, such as in studies of sequential context, where numerous judgements of a certain stimulus are obtained in two conditions and then compared. In one condition preceding stimuli are smaller and in the other they are larger than the judged stimulus. Probabilities of the two conditions are balanced. In such studies, the target stimulus is not the most extreme stimulus presented, but instead lies somewhere in the middle of the series of all presented intensities. In most other studies where sequential effects are not monitored extreme stimuli (stimuli at both sides of the range of intensities) are treated as any other stimuli in the pool and contribute to the outcome of the study. Let us elaborate on what can happen in such cases.

For example, let us imagine a magnitude estimation study on brightness perception. The exponent of psychophysical power function is around 1/3 for brightness perception. Let us imagine that we present 11 different stimuli to observers in the randomized trials and that the sensation magnitude increases uniformly. It is probably safe to say that when the sixth intensity is presented, the probability that the immediately preceding stimulus was smaller equals the probability that the preceding stimulus was larger. Therefore, if assimilation to the prior stimulus occurs, it is reasonable to assume that it will be neutralised over many trials. But when the smallest stimulus is presented, the probability that prior stimulus was smaller equals zero. Therefore, if the effect of assimilation towards the stimulus presented on trial *N*–1 is much stronger than the contrast effect of stimuli presented on trials *N*–2 or before (and it usually is; see e.g. Lockhead, 1992), the estimated magnitude will be larger than it should be (i.e., than it would have been had there been no sequential effects). The eleventh intensity will, equivalently, be judged smaller than it would have been with no sequential effects taking place. Therefore, stimuli smaller than the medium one would all be overestimated, increasingly so for the more extreme ones, and stimuli larger than the medium one would all be underestimated. This outcome would resemble the regression to the mean. In the end, sequential effects could result in a changed (i.e., too decelerated) psychophysical function.

If the finding that sequential effects are stronger for more extreme stimuli could be generalized to other psychological domains, this would have some important consequences. For example, in case of psychological assessment, regression of judgements of extreme items (e. g., items expressing extremely positive or negative attitude) towards the mean would result in too narrow dispersions of responses, which may in the next step lead to too low correlations among items and lower internal consistency of scales. It could also compromise their validity. Specifically, it could lead to underestimation of criterion validity of the scale. High correlations among different psychological constructs are rare, and perhaps sequential effects, although indirectly, add something to this misfortune. One must also recall that in a typical psychometric measurement items are presented in the same order to every person. Even if there is no regression-to-the-mean effect, assimilation or contrast effects are very difficult to neutralise across participants.

In psychophysics, the sequential effects are usually studied over extensive periods of time, and their average effect is extracted from data. However, it was already Guilford (1954) who wrote that learning, as well as other background stimuli, can affect the time-order errors. It is reasonable to assume that sequential effects themselves vary in time, so although we may be able to determine how they operate throughout a longer period, we may never be able to determine how they operate at a specific moment in time. Consequently, we may not be able to include them in the psychophysical equation properly. It seems that a *post hoc* analysis of those effects and a subsequent adjustment of results may not be the optimal procedure for controlling these effects.

What would be a better way to control for the sequential effects? Perhaps a better way would be to ensure that in our studies they do not emerge at all. Jesteadt et al. (1977) presented several "erase" stimuli, randomly selected from the entire signal range, between different trials, and they found that sequential effects were eliminated. In cross-modality matching, the changed initial value of the variable stimulus in each trial reduced sequential effects (Teghtsoonian, Teghtsoonian, & DeCarlo, 2008). Thus, it seems that loading attentional resources, inserting some new, irrelevant stimuli in the series, and erasing the memory trace of prior stimuli and responses may be helpful in controlling the sequential effects. Future studies will have to show what kinds of stimuli work best at achieving this goal.

REFERENCES

Baird, J. C., Green, D. M., & Luce, R. D. (1980). Variability and sequential effects in cross-modality matching of area and loudness. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 277–289.

DeCarlo, L. T. (2003). An application of a dynamic model of judgement to magnitude production. *Perception & Psychophysics, 65*, 152–162.

DeCarlo, L. T., & Cross, D. V. (1990). Sequential effects in magnitude scaling: Models and theory. *Journal of Experimental Psychology: General, 119*, 375–396.

Dzhafarov, E. N. (1992). Can brightness be related to luminance by a meaningful function? *Behavioral and Brain Sciences, 15*, 565–566.

Fechner, G. T. (1966). *Elements of psychophysics* (Vol. I) (E. G. Boring & D. H. Howes, Eds.; H. E. Adler, Trans.). New York: Holt, Rinehart & Winston. (Original work published 1860).

Green, D. M., Luce, R. D., & Duncan, J. E. (1977). Variability and sequential effects in magnitude production and estimation of auditory intensity. *Perception & Psychophysics, 22*, 450–456.

Gregson, R. A. M. (1992). Walking in a psychophysical dustbowl creates a dustcloud. *Behavioral and Brain Sciences, 15*, 568–569.

Guilford, J. P. (1954). *Psychometric methods*. New York: McGraw-Hill.

Helson, H. (1964). *Adaptation level theory.* New York: Harper & Row.

Holland, M., & Lockhead, G. (1968). Sequential effects in absolute judgments of loudness. *Perception & Psychophysics, 3,* 409–414.

Jesteadt, W., Luce, D. R., & Green, D. M. (1977). Sequential effects in judgements of loudness. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 92–104.

Lockhead, G. R. (1992). Psychophysical scaling: Judgments of attributes or objects? *Behavioral and Brain Sciences, 15*, 543-601.

Lockhead, G. R., & King, M. C. (1983). A memory model of sequential scaling tasks. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 461–473.

Schifferstein, H. N. J., & Frijters, J. E. R. (1992). Contextual and sequential effects on judgments of sweetness intensity. *Perception & Psychophysics, 52*, 243–255.

Teghtsoonian, R. (1992). Selecting one attribute for judgement is not an act of stupidity? *Behavioral and Brain Sciences, 15*, 580–581.

Teghtsoonian, M., Teghtsoonian, R., & DeCarlo, L. T. (2008). The influence of trial-to-trial recalibration on sequential effects in cross-modality matching. *Psychological Research, 72*, 115–122.

Wagner, M., & Baird, J. C. (1981). A quantitative analysis of sequential effects with numeric stimuli. *Perception & Psychophysics, 29*, 359–364.

Ward, L. M. (1973). Repeated magnitude estimations with a variable standard: Sequential effects and other properties. *Perception & Psychophysics, 13,* 193–200.

Ward, L. M. (1975). Sequential dependencies and response range in cross-modality matches of duration to loudness. *Perception & Psychophysics, 18,* 217–223.

Ward, L. M. (1982). Mixed-modality psychophysical scaling: Sequential dependencies and other properties. *Perception & Psychophysics, 31*, 53-62.

Ward, L. M., & Lockhead, G. R. (1971). Response system processes in absolute judgment. *Perception & Psychophysics, 9,* 73–78.