RECALIBRATION OF LOUDNESS: SENSORY VS. DECISIONAL PROCESSES

Yoav Arieh and Lawrence E. Marks
John B. Pierce Laboratory and Yale University, New Haven, CT, USA
yarieh@jbpierce.org, marks@jbpierce.org

Abstract

Recalibration of loudness comprises a difference in relative responsiveness over sound frequency brought about by contextual differences in prior stimulation. Evidence from studies of both loudness perception and auditory response times, as well as from studies of intensity perception in taste, smell, haptic touch, and vision, supports three general hypotheses: (1) Recalibration in loudness judgments and response times consists of modifications in the underlying sensory representations, and not (just) shifts in response criteria. (2) The modifications in sensory representations themselves comprise adaptation-like decrements in suprathreshold responsiveness. (3) Recalibration arises when stimulus magnitudes (e.g., auditory intensity) are processed across a second domain (e.g., auditory frequency) in distinct channels (e.g., "critical bands"); it arises from transient stimulation in one or more channels at intensity levels that are sufficiently great to "adapt" that channel or those channels.

When listeners respond as quickly as possible to the onset of a tone, the resulting simple reaction time (SRT) provides a measure of intensity processing in the auditory system (e.g., Luce & Green, 1972). Not only does SRT decline with increasing signal intensity, much as loudness increases with intensity – the higher the SPL, the greater the loudness and the smaller the SRT – but many parameters that influence loudness, such as the presence of masking noise (Chocholle & Greenbaum, 1966), have comparable effects on SRT. The inverse relation between SRT and loudness is not perfect (Kohfeld, Santee, & Wallace, 1981), as these measures depend on overlapping but presumably not identical mechanisms, but the relation is close nevertheless. Most pertinently, SRT and loudness depend similarly on the intensity levels of other stimuli to which the listener has recently been exposed.

Consider the following experiment: Listeners respond as quickly as possible, after a variable foreperiod, to the onset of a 500-Hz or 2500-Hz tone, each of which may take on one of three possible SPLs. In condition A, the SPLs at 500 Hz are relatively low (35, 50, and 65 dB) and the SPLs at 2500 Hz relatively high (45, 60, and 75 dB), and in condition B the relation between frequency and intensity is reversed, with SPLs at 500 Hz relatively high (50, 65, and 80 dB) and those at 2500 Hz relatively low (30, 45, and 60 dB). A sample of results (6 listeners) appears in the left panel of Figure 1. Two features of the results are important. First, SRT declines as intensity increases (note that the scale of SRT is inverted so the curves rise as SPL increases). And second, the relation between SRT at 500 Hz and SRT at 2500 Hz varies with condition, SRT being relatively longer at 2500 Hz in condition A and at 500 Hz in condition B. This pattern of SRTs strongly resembles the results obtained time and again when subjects rate or directly compare the loudness of brief stimuli presented under comparable stimulus conditions (e.g., Marks, 1988, 1992a, 1992b, 1993, 1994, 1996; Marks & Warner, 1991). The right panel of Figure 1 replots loudness judgments of 16 listeners obtained more than a decade ago (Marks, 1988) using the method of magnitude estimation.

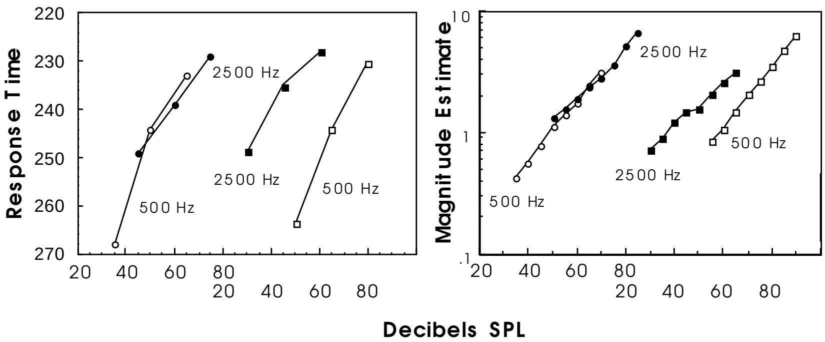

Figure 1. Left panel: Simple response times to detect 500-Hz and 2500-Hz tones in different context sets of SPLs (Arieh & Marks). Right panel: Magnitude estimates of loudness of 500-Hz and 2500-Hz tones in comparable contextual sets (Marks, 1988).

Consider a pair of acoustic signals that, under "normal" conditions, are judged to be about equally loud – say, a 500-Hz tone at 75 dB and a 2500-Hz tone at 70 dB. Loudness at 500 Hz, relative to 2500 Hz, becomes notably greater, and SRT shorter, when the stimulus ensemble contains soft 500-Hz tones and loud 2500-Hz tones. But loudness at 2500 Hz becomes relatively greater, and SRT shorter, when the assignment of low and high SPLs to the two frequencies reverses. Why?

At first glance, these changes in loudness judgment and SRT resemble the kinds of context-induced changes subsumed under Helson's (1964) adaptation-level (AL) theory, and this resemblance raises two fundamental, and connected, issues. First, taking a Helsonian perspective, it might be tempting to assume that loudness at a given frequency is reduced when the mean SPL is high, and hence the AL is high, but enhanced when the mean SPL and AL are low. By this token, changes in the stimulus levels may effect both increases and decreases in loudness. Alternatively, it is possible that loudness undergoes only reductions but not enhancements, with the magnitude of the reduction depending on the mean SPLs presented. These alternatives, and they are by no means the only possibilities, are closely connected to the second issue. What is it that is modified? Do the changes in loudness judgment and SRT reflect modifications in the underlying sensory representations of intensity at one frequency or both? Or do the changes reflect modifications in decisional processes that listeners use in judging and comparing stimuli?

Contextual effects are often attributed to adjustments in post-perceptual processes, in decision criteria or response production, and not to changes in the sensory representations of stimulus intensity (e.g., Anderson, 1975), and, in fact, Marks (1988) initially sought to explain (away) the loudness judgments in terms of the ways that people use numbers in magnitude estimation – for instance, to a tendency to apply a constant range of responses to signals at each frequency even when the mean levels shift across conditions. Unfortunately, this account cannot explain the comparable effects on SRT, nor can it explain why context-induced changes can be measured by various psychophysical methods, including direct comparison of loudness differences (Schneider & Parker, 1990), direct comparison of loudness (Mapes-Riordan & Yost, 1999; Marks, 1992a) and selective adaptation (Marks, 1993). Processes of overt numerical responding are insufficient to explain these findings. Nevertheless, it is conceivable that the contextual changes do result from shifts in decisional criteria. For instance, presenting strong signals at frequency 1 (f1) might lead listeners to shift their criterion at f1, relative to criteria at other frequencies, so that relatively greater "loudness" or "information" is needed at f1 to produce a loudness match, or to initiate a simple response.

Three hypotheses regarding recalibration

The evidence at hand, accumulated over the past decade, supports three main hypotheses:

(1) The context-dependent changes in loudness judgments and response times – dubbed "recalibration" (Marks, 1994) – reflect modifications in the underlying sensory representations, and not (just) shifts in response criteria.

(2) These modifications in sensory representations themselves consist of adaptation-like decrements in suprathreshold responsiveness.

(3) Recalibration arises when stimulus magnitudes (e.g., auditory intensity) are processed across a second domain (e.g., auditory frequency) in distinct channels (e.g., "critical bands"); it results from transient stimulation in one or more channels at intensity levels that are sufficiently great to "adapt" that channel or those channels.

These three hypotheses are couched in broad form. Support for them comes not only from studies of recalibration in hearing, but also from numerous analogous studies in other sense modalities. Recalibration seems to represent the results of adaptive processes found throughout the sensory realm.

Recalibration as a sensory process: Evidence from choice response times

Findings described thus far, both for loudness judgment and SRT, are consistent with both sensory and a decisional interpretations of recalibration. It is possible to account for the context-dependent changes in SRT, for example, in terms of differential changes in response criteria at the two signal frequencies. To eliminate a decisional explanation, it would be necessary to control response criteria, or at least to quantify their effects and demonstrate that changes in their location cannot explain the results. One approach is to embed the varying-context design within a task in which the listener must identify the signal on each trial as low or high in frequency. To do this, we capitalize on the finding of Keuss and van der Molen (1982) that choice response time (CRT) is sensitive to signal intensity (CRT declines as intensity decreases) if one uses a foreperiod that is long and variable. For this purpose, we used a foreperiod with a constant hazard function; that is, the probability of presenting a signal within a given small time interval is constant, no matter how long the listener has already waited. Importantly, the choice procedure provides measures of accuracy as well as response speed. In different test conditions, we manipulated not only the contextual set of SPLs at 500 Hz and 2500 Hz but also the instructions, emphasizing speed and accuracy to different degrees. Six listeners participated in two sessions in which they received alternating blocks of trials containing tones from stimulus set A (low SPLs at 500 Hz and high SPLs at 2500 Hz) or set B (high SPLs at 500 Hz and low SPLs at 2500 Hz). Listeners pressed one key if the tone was low in frequency and another if the tone was high. One session emphasized the speed of responding ("fast" instructions) and the other speed and accuracy ("slow" instructions), the aim being to get a handle on the speed-accuracy trade-off at each frequency.

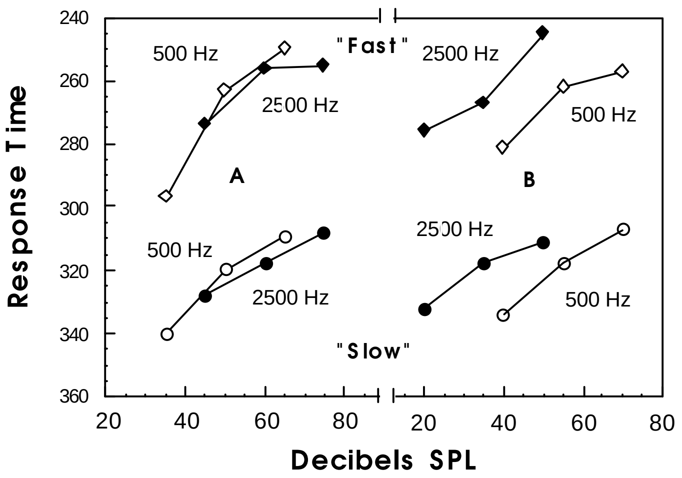

Figure 2. Response times to classify 500-Hz and 2500-Hz tones in two contextual conditions (A and B), each under two different response instructions (“fast” and “slow).

Table 1

Error rates for classifying 500-Hz and 2500-Hz tones in the two contextual conditions (A and B), each under the two sets of instructions (“fast” and “slow”).

	500 Hz		2500 Hz	
Session	Context A	Context B	Context A	Context B
“Fast”	19%	23%	17%	17%
“Slow”	4%	7%	5%	5%

Figure 2 shows the results. In both sessions, CRTs were monotonically declining functions of SPL, indicating a role of intensity processing in generating the responses. Further, the CRTs, like the SRTs of Figure 1, reveal recalibration: Whichever frequency had the lower SPLs was classified faster. Next consider the measures of accuracy, shown in Table 1. If recalibration reflects shifts in criteria, then the relatively fast CRTs at 500 Hz with stimulus set A and at 2500 Hz with set B should be accompanied by relatively high error rates. This did not happen. At 2500 Hz, the error rates with sets A and B were essentially the same in each session, and at 500 Hz the error rates are actually smaller with set A compared to set B. Taken together, the measures of CRT and accuracy imply that the recalibration has a sensory rather than decisional basis.

Recalibration as channel-selective "adaptation"

Evidence supporting the view that recalibration comprises channel-specific "adaptation" comes from Experiment 15 of Marks (1993), which used a "selective adaptation" method. The experiment had two parts. In one, listeners were passively exposed to one of two different adapting regimens: alternations of briefly presented 500-Hz tones at 53 dB and 2500-Hz tones at 68 dB (set A) or 500-Hz tones at 73 dB and 2500-Hz tones at 48 dB (set B). Exposure to set A substantially increased the probability of judging a 500-Hz tone louder than a 2500-Hz tone previously equated to it, and exposure to B substantially decreased the probability. In other words, relative loudness is affected simply by listening to soft tones of one frequency and loud tones of another. Even more telling was the outcome of the second part of the experiment, where listeners were exposed to just one frequency at one SPL. In this case, the louder tone produced recalibration comparable in magnitude to the recalibration measured in the first part of the experiment. Exposure to the softer tone, however, had essentially no effect at all.

The results just outlined are especially compatible with the notion that recalibration involves a depression in responsiveness ("adaptation") resulting from channel-specific stimulation at relatively high signal levels. Note that most of the experiments in the second author's laboratory have used methods in which subjects are exposed to two different kinds of signal, such as tones of different frequency. Using two different signals makes it possible to measure recalibration as a change in relative response to the signals under different contextual conditions (eliminating the "response assimilation effect" that dominates judgments of loudness when only one frequency is presented: Marks, 1993), but can make it difficult to determine whether two different signals are required to produce recalibration. The results of Marks (1993, Experiment 15) show that recalibration requires only one signal (hence activation of just one channel), a result consistent with other findings (e.g., Mapes-Riordan & Yost, 1999). Nevertheless, whenever a second frequency is presented, even if just for direct comparison, repeated presentation of pairs of stimuli may produce "adaptation" in both frequency channels.

The prominent role of high versus low signal levels in recalibration is not restricted to loudness perception. Armstrong and Marks (1997) reported similar findings in vision, when subjects judged the lengths of lines presented in horizontal and vertical orientations. Subjects perceived a given horizontal line to be relatively longer than a given vertical line when the stimulus set comprised short horizontals and long horizontals, but perceived the same horizontal to be relatively shorter when the set comprised long horizontals and short verticals. Most critically, recalibration in the perception of length seems to depend on exposure to the long lines but not the short ones, and to consist solely in reductions in perceived length. Although details have not yet been worked out in all modalities, recalibration appears to characterize perception in most if not all sensory modalities. Recalibration has been observed in the perception of taste intensity of stimuli varying in quality (Rankin & Marks, 1991, 1992, 2000), olfactory intensity of stimuli varying in quality (Rankin & Marks, 2000), and haptic perception of length of movements in different directions (Marks & Armstrong, 1996), as well as visual extent and loudness. The study of Rankin and Marks (2000) is particularly noteworthy because those investigators included conditions in which tastants and odorants were sipped, and thus perceived as "flavors." The results showed that recalibration depended on activating different sensory channels (gustation and olfaction), rather than perceived dissimilarity in the qualities of the stimuli.

Acknowledgement

Preparation of this article was supported by grant R01 DC03852 from the National Institute of Deafness and Other Communicative Disorders, NIH, to LEM.

References

Anderson, N.H. (1975). On the role of context effects in psychophysical judgment. *Psychological Review*, **82**, 462-482

Armstrong, L., & Marks, L.E. (1997). Stimulus context, perceived length, and the horizontal-vertical illusion. *Perception & Psychophysics*, **59**, 1200-1213.

Chocholle, R., & Greenbaum, H.B. (1966). La sonie de sons purs partiallement masqués. Étude comparative par une méthode d'égalisation et par la méthode des temps de réaction. *Journal de Psychologie Normale et Pathologique*, **63**, 387-414.

Helson, H. (1964). *Adaptation level theory: An experimental and systematic approach to behavior.* New York: Harper and Row.

Keuss, P.J.G., & van der Molen, M.W. (1982). Positive and negative effects of stimulus intensity in auditory reaction tasks: Further studies on immediate arousal. *Acta Psychologica*, **52**, 61-72

Kohfeld, D.L., Santee, J.L., & Wallace, N.D. (1981). Loudness and reaction time: I. *Perception & Psychophysics*, **29**, 535-549.

Luce, R.D., & Green, D.M. (1972). A neural timing theory for response times and the psychophysics of intensity. *Psychological Review*, **79**, 14-57.

Mapes-Riordan & Yost (1999). Loudness recalibration as a function of level. *Journal of Acoustical Society of America*, **106**, 3506-3511.

Marks, L.E. (1988). Magnitude estimation and sensory matching. *Perception & Psychophysics*, **43**, 511-525.

Marks, L.E. (1992a). The slippery context effect in psychophysics: Intensive, extensive, and qualitative continua. *Perception & Psychophysics*, **51**, 187-198.

Marks, L.E. (1992b). The contingency of perceptual processing: Context modifies equal-loudness relations. *Psychological Science*, **3**, 285-291.

Marks, L.E. (1993). Contextual processing of multidimensional and unidimensional auditory stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, **19**, 227-249.

Marks, L.E. (1994). "Recalibrating" the auditory system: The perception of loudness. *Journal of Experimental Psychology: Human Perception and Performance*, **20**, 382-396.

Marks, L.E. (1996) Recalibrating the perception of loudness: Interaural transfer. *Journal of the Acoustical Society of America*, **100**, 473-480.

Marks, L.E., & Armstrong, L. (1996). Haptic and visual representations of space. In T. Inui and J. L. McClelland (Eds.), *Attention and Performance XVI* (pp. 263-287). Cambridge, MA: MIT Press.

Marks, L.E., & Warner, E. (1991). Slippery context effect and critical bands. *Journal of Experimental Psychology: Human Perception and Performance*, **17**, 986-996.

Rankin, K.M., & Marks, L.E. (1991). Differential context effects in taste perception. *Chemical Senses*, **16**, 617-629.

Rankin, K.M., & Marks, L.E. (1992). Effects of context on sweet and bitter tastes: Unrelated to sensitivity to PROP (6-n-propylthiouracil). *Perception & Psychophysics*, **52**, 479-486.

Rankin, K.M., & Marks, L.E. (2000). Chemosensory context effects: Role of perceived similarity and neural commonality. *Chemical Senses*, **25**, 748-659.

Schneider, B., & Parker, S. (1990). Does stimulus context affect loudness or only loudness judgment? *Perception & Psychophysics,* **48**, 409-418.