Integration of serial sensory information in haptic perception of softness

Anna Metzger, Alexandra Lezkan and Knut Drewing

Justus-Liebig University, Giessen Germany

Submitted to Journal of Experimental Psychology: Human Perception and Performance

Address correspondence to: Anna Metzger, Institute for Psychology, Giessen University, Otto-Behaghel-Str. 10F, 35394 Giessen, Germany, Phone: +49 641 9926113, e-mail: anna.metzger@psychol.uni-giessen.de

AUTHORS' NOTE

This work was supported by the Deutsche Forschungsgemeinschaft SFB TRR 135, A05. Response and movement data of individual participants from all experiments presented here is available at [link will be pasted in final version]. We are grateful to Matteo Toscani for helpful discussions and comments, Lorilei Alley for helping to improve grammar and style and Steffen Bruckbauer, Maria Ebsen, Claire Weyel and Laura Wilhelm for help with running the experiments.

ABSTRACT

Redundant estimates of an environmental property derived simultaneously from different senses or cues are typically integrated according to the Maximum Likelihood Estimation model (MLE): Sensory estimates are weighted according to their reliabilities, maximizing the percept's reliability. Mechanisms underlying the integration of sequentially derived estimates from one sense are less clear. Here we investigate the integration of serially-sampled redundant information in softness perception. We developed a method to manipulate haptically perceived softness of silicone rubber stimuli during bare finger exploration. We then manipulated softness estimates derived from single movement segments (indentations) in a multi-segmented exploration to assess their contributions to the overall percept. Participants explored two stimuli in sequence, using 2-5 indentations and reported which stimulus felt softer. Estimates of the first stimulus' softness contributed to the judgments similarly, whereas for the second stimulus estimates from later as compared to earlier indentations contributed less. In line with unequal weighting, the percept's reliability increased with increasing exploration length less than predicted by the MLE model. This pattern of results is well explained by assuming that the representation of the first stimulus fades when the second stimulus is explored, which fits with a neurophysiological model of perceptual decisions (Deco et al., 2010).

STATEMENT OF PUBLIC SIGNIFICANCE:

When exploring the softness of an object people usually indent it repeatedly, generating redundant estimates of its softness. We investigated how such serially-sampled estimates are combined. Participants repeatedly indented two rubber stimuli one after the other to estimate and compare their softness. While estimates from the exploration of the first stimulus contributed equally to this decision, later estimates from the exploration of the second stimulus contributed less than earlier ones. This is in contrast to typical observations from simultaneous information integration suggesting that the contribution of sensory estimates is solely determined by the estimate's reliability. Instead, our results indicate that serial integration is affected by memory limitations. We propose that the memory representation of the first stimulus decays during the exploration of the second stimulus and hence, later estimates contribute less to the decision regarding which of the two stimuli feel softer.

INTRODUCTION

Perception is the process of estimating the properties of our environment. If there are redundant signals from an environmental property, meaning a signal is repeated (sequential redundancy), or there are simultaneous signals in different dimensions (simultaneous redundancy), the property can be better detected (Mulligan & Shaw, 1980; Shaw, 1982; Swets, Shipley, McKey, & Green, 1959; Swets & Birdsall, 1978). Consequently, perception of an environmental property is more reliable with repeated estimates or estimates derived from different senses or cues, as compared to when perception is based on a single estimate. For example, sequential viewing (e. g. Oostwoud Wijdenes, Marshall & Bays, 2015) or touching (e. g. Louw, Kappers & Koenderink, 2005) usually increases perceptual reliability. An object's position and its properties can be estimated more reliably using different senses simultaneously, e. g. when estimating the position of the hand from vision and proprioception (van Beers, Sittig & van der Gon, 1998) or the size of an object from vision and touch (Ernst & Banks, 2002). Also combining different cues from a single sense increases perceptual reliability e. g. estimating the slant of a plane from the texture gradient and the linear perspective (Oruc, Maloney & Landy, 2003).

The integration of redundant information is often modeled as maximum likelihood estimation (MLE, Ernst & Buelthoff, 2004): single estimates are integrated by linear weighted averaging with the weights being proportional to the single estimates' relative reliabilities (Cochran, 1937). Reliability is defined as the inverse of variance $r = \sigma^{-2}$.

$$\hat{S} = \sum_{i=1}^{n} w_i s_i, \text{with } w_i = \frac{\sigma_i^{-2}}{\sum_{i=1}^{n} \sigma_i^{-2}},\ w_i \geq 0 \text{ and} \sum_{i=1}^{n} w_i = 1. \quad (1)$$

$\hat{S}$ denotes the combined estimate (i.e., the percept), s_i the different single estimates, w_i their individual weights and n the number of available estimates. If the single estimates are independent Gaussian variables, the combined estimate would have the maximal possible reliability of

$$R_{\hat{S}} = \sum_{i=1}^{n} r_i . \qquad (2)$$

Several studies on the integration of simultaneously available estimates (overview: Landy, Banks & Knill, 2011) support the MLE model for multisensory and cue integration in human perception (e. g. Alais & Burr, 2004; Ernst & Banks 2002; Fetsch, DeAngelis & Angelaki, 2010; Hartcher-O'Brien, Di Luca & Ernst, 2014; Helbig & Ernst, 2007; Hillis, Watt, Landy & Banks, 2004; Moscatelli et al. 2016). Nevertheless, there are also some reports of integration of simultaneous estimates, where perceptual reliability was not maximized or where weights could be changed by feedback without changing reliability (e.g. Cellini, Kaim & Drewing, 2013; Ernst, Banks, & Buelthoff, 2000; Jacobs & Fine, 1999; Rosas, Wagemans, Ernst & Wichmann, 2005; van Beers, van Mierlo, Smeets & Brenner, 2011). There is evidence that multisensory integration also depends on the exploration mode (visual and haptic both parallel or serial vs. visual parallel and haptic serial or vice versa, Plaisier, van Dam, Glowania & Ernst, 2014).

In contrast, studies investigating the mechanisms underlying the integration of sequential estimates from one sense do not yet reveal concordant models. For visual perception of color Oostwoud Wijdenes et al. (2015) found that pre- and postsaccadic estimates were integrated consistent with the MLE model with a higher weight given to the more reliable estimate (derived when the disc was closer to the fovea). Wolf and Schuetz (2015) showed that the orientation of a visually presented grating and the estimate's reliability could be predicted by the MLE model from the reliabilities of estimates sequentially derived in the periphery and the fovea. Also in haptic perception sequential redundancy is exploited to increase reliability. For example, extended exploration or longer stimulus presentation decreased thresholds in haptic discrimination of surface orientation (Giachritsis, Wing and Lovell, 2009), or in the detection of sine-wave gratings and vibro-tactile stimuli (Gescheider, Bolanowski, Pope and Verrillo, 2002; Louw et al., 2005). However, Drewing, Lezkan and Ludwig (2011) showed for virtual sine-wave gratings that the gain of sequential redundancy was significantly lower than the gain predicted by the MLE model.

A Kalman filter (Bryson & Ho, 1975) could potentially be more appropriate than the MLE model to account for sequential integration of information. The Kalman filter is an optimal model for estimating the state of dynamic linear systems over time. It is a recursive algorithm estimating the current state of a system by combining the current measurement with prior information, maximizing the estimate's reliability. Within the Kalman filter framework sequential integration of sensory estimates can be considered as recursive combination of the current sensory estimate (measurement) with the information obtained from prior sensory estimates. The dynamics of a (one-dimensional) system are predicted in the Kalman filter by a transition function from the previously estimated state $\hat{s}_{i-1}$ with the estimated noise $\hat{\sigma}_{i-1}$ to a hypothetical current state: $s_i = A\hat{s}_{i-1} + \varepsilon_p$ (ε_p being the process noise drawn from $N[0,\sigma_p]$). This prediction is combined with the current measurement $y_i = S + \varepsilon_m$ (ε_m being the measurement noise drawn from $N[0,\sigma_m]$ and S the true state of the system) by weighted averaging to obtain a new state estimate:

$$\hat{s}_i = \frac{\sigma_m^{-2}}{A^{-2}\hat{\sigma}_{i-1}^{-2} + \sigma_p^{-2} + \sigma_m^{-2}} y_i + \frac{A^{-2}\hat{\sigma}_{i-1}^{-2} + \sigma_p^{-2}}{A^{-2}\hat{\sigma}_{i-1}^{-2} + \sigma_p^{-2} + \sigma_m^{-2}} A\hat{s}_{i-1}\,. \tag{3}$$

The updated reliability is then given by

$$r_i = r_{i-1} + r_m \text{ with } r_{i-1} = A^{-2}\hat{\sigma}_{i-1}^{-2} + \sigma_p^{-2} \text{ and } r_m = \sigma_m^{-2}\,. \tag{4}$$

Similar to the MLE model weights that maximize reliability are proportional to the relative reliabilities of the prior information and the measurement. In contrast to the MLE model, the Kalman filter also includes the process noise that influences the prior sensory information and it can be extended to model non-linear systems (Anderson, & Moore, 1979). The Kalman filter was shown to approximate well aspects of state estimation of dynamic systems in human visual perception and learning (e.g. Kwon, Tadin & Knill, 2015; Rao, 1999) and visuomotor behavior (e. g. Burge, Ernst & Banks, 2008; Koerding & Wolpert, 2004). However, sequential integration in the perception of a constant stimulus corresponds to the estimation of the state of a static system (constant state). For the estimation of the state S of an one-dimensional static system (A = 1) with negligible process noise ($\sigma_p \sim 0$), n iterations of the Kalman filter (Eqs. 3 & 4) result in an final estimate $\hat{S}$ as given by Eq. 1 and a final reliability $R_{\hat{S}}$ as given by Eq. 2 in the MLE model. Hence in this case, the estimate and its reliability obtained after n Kalman filter iterations are identical to the ones obtained from MLE integration of n estimates.

The question thus remains why and under which conditions in haptic perception sequential redundancy did not result in maximal possible reliability (Eq. 2). For discrimination of virtual gratings in extended exploration Drewing et al. (2011) had predicted maximal reliability under the assumption that the repeated estimates from one sense are equally weighted, because they are all gathered in the same manner and thus have similar reliability. The authors suggested that the observed lower reliability is associated with unequal weighting. Lezkan and Drewing (2014) hypothesized that memory decay might cause unequal weighting of sequential estimates. In haptic perception, the accumulation of sensory information spans longer times than in visual perception. Due to memory decay, early estimates might become less reliable over time, so that they are weighted less. Lezkan and Drewing (2014) measured the contribution (weight) of every estimate to the discrimination of spatial frequency of two virtual sine-wave gratings (explored successively using 2-5 strokes, each stroke regarded to provide one estimate) by slightly changing spatial frequency during one stroke on the standard. Strokes' weights decreased monotonically as a function of their temporal distance to the comparison stimulus being highest around the time point when participants switched between the stimuli. The findings provide important first hints to the sequential integration during haptic perception: In contrast to simultaneous integration, estimates that were gathered with equal reliability were not equally weighted, and perceptual performance was correspondingly lower than predicted by the MLE model (Eq. 2).

However, the perceptual situation investigated in the study by Lezkan and Drewing (2014) might represent a highly specific case: Only virtual stimuli were used that were explored with a thimble connected to a force feedback device, thereby omitting several cutaneous cues, typically dominating haptic perception of softness (e.g. Bergmann Tiest & Kappers, 2009 for softness). Furthermore, participants could not explore the stimuli in a natural manner, because force and velocity were prescribed. These conditions differ considerably from everyday haptic perception, in that cutaneous information is highly relevant and exploration is hardly constrained. Hence the question arises, whether the results are representative for natural haptic perception. In the present study we investigated the integration of sequential sensory information in haptic perception for naturalistic conditions and stimuli, using softness perception as an example. We used real stimuli (silicon rubber) and constrained the stimulus' exploration only according to its length.

Softness refers to the perception of an object's compliance (ratio between displacement of the object's surface and the force applied to the object). Active exploration of softness usually involves successive manipulation of the object using a stereotypical movement classified as the *Exploratory Procedure* of *Pressure*, in which the object is repeatedly squeezed between the fingers or palpated with a finger or a tool (Lederman &

Klatzky, 1987; Kaim & Drewing, 2011). Active haptic perception of an object's softness could thus be thought of as a multi-segmented exploration, where the movement segments refer to single indentations of the object. A single indentation comprises an increase of finger force up to a peak force, followed by a force decrease down to a local minimum, yielding a pattern of increasing and decreasing deformation of the object by the finger. We considered a single indentation as the basis of a single "indentation-specific" estimate of softness. The combined percept of softness would then be based on the integration of these multiple estimates.

Since in our experiment participant explored real stimuli, whose physical softness was given, we could not present a different stimulus for a single indentation as in Lezkan and Drewing (2014) to assess indentation-specific weights. We developed a paradigm to manipulate perceived softness of real deformable objects during the exploration with bare fingers (Experiment 1). We transmitted subtle external forces to the exploring finger of the participant (Figures 1 and 2) which pressed the finger more into the stimulus (pushing force) or pulled it away (pulling forces). External forces (proportional by factor α to the forces participants applied themselves) changed perceived softness proportional to α. Results from Experiment 1 have been presented in a conference article (Metzger & Drewing, 2015), where they were discussed in the context of the cutaneous and kinesthetic integration. We reconsider the results here, because they are the basis for the following experiments: We used external forces to manipulate softness estimates from single indentations.

In four connected Experiments 2a-d, differing in the length of the exploration (2-5 indentations) we studied the integration of sequential softness estimates during multi-segmented exploration of deformable silicon rubber stimuli. We first investigated how perceptual reliability depends on the number of exploratory movement segments. We hypothesized that perceptual reliability would increase with an increasing number of indentations, but less than predicted by the MLE model (Eq. 2; cf. Drewing et al., 2011).

Second, we studied how estimates from single exploratory segments are weighted in perceived softness. Based on the findings of Lezkan and Drewing (2014), we expected to find the highest weights of indentation-specific estimates around the time point when participants switched between the stimuli in order to compare them, and that weights would decrease with increasing temporal distance to this time point. The MLE model predicts equal weights w_i for the integration of estimates gathered with equal reliabilities.

General Methods

Participants

Participants were right-handed (with the exception of one), naïve to the purpose of the experiments, volunteered to participate, and were reimbursed for their time. No participant reported sensory or motor impairments of the index finger of the dominant hand, which we confirmed by measuring a two-point discrimination threshold lower than 3 mm on this finger (Johnson & Phillips, 1981). In total 70 participants took part in the experiments; 10 participants were excluded from the analysis (six due to misdetection of indentations in more than 5% of trials (see *Data analysis*, Experiments 2a-d); two due to repeated loss of connection between the finger and the force-feedback device during external force transmission); two due to reporting in the post-experimental survey that the restrictions imposed by our setup strongly impeded the exploratory behavior). The main characteristics of the final sample are listed in Table 1. More participants were recruited for Experiments 2c-d, because subtler effects in weights were expected with longer explorations. The study was approved by the local ethics committee LEK FB06 at Giessen University and was in line with the declaration of Helsinki from 1964. Written informed consent was obtained from each participant.

[Insert Table 1 about here]

Apparatus and setup

The experiments were conducted at a visuo-haptic workbench (Figure 1), which displayed haptic stimuli by a PHANToM 1.5A haptic force feedback device, and 3D visual stimuli using a 22"-computer screen (120 Hz, 1280x1024 pixel) viewed via stereo glasses indirectly through a mirror from 40 cm viewing distance. The mirror enabled spatial alignment of visual and haptic displays and prevented participants from seeing their hand when touching the stimuli. Force feedback was limited to the transmission of subtle external forces during the exploration of real stimuli. Two rubber stimuli were placed side-by-side in front of the participants on a force sensor (produced by ME-Messsysteme GmbH) consisting of a bending beam load cell (LCB 130) and a measuring amplifier (GSV-2AS, resolution 0.05 N, temporal resolution 682 Hz) to record exerted forces. Participants touched the stimuli with the index finger of their dominant hand using a downward directed movement. The finger was connected to the PHANToM arm. The visual 3D scene comprised a schematic representation of the finger (sphere of 8 mm diameter) and the two stimuli. Visual information was used only to guide participants through the experiment. Importantly, no visual information about the finger movement and stimulus compliance was presented when the stimuli were touched (force > 0.1 N). The head was fixated by a chin rest.

[Insert Figure 1 about here]

A custom-made gimbal-like adapter was used to connect the participant's index finger to the PHANToM (Figure 2). It was designed to have no vertical degrees of freedom to ensure the transmission of forces and to allow (despite this restriction) natural and comfortable exploration of the stimuli. The adapter left the finger pad uncovered and was adjusted to the preferred inclination of the finger between 0° and 45° before the experiment, by rotating the main gimbal of the adapter, which was then fixed by two screws during the experiment. We used a circular design for the adapter to ensure that the direction of external forces as well as the calibrated zero-position of the PHANToM would not change with inclination. To exclude finger movement independent of the adapter, the dorsal side of the finger was affixed to the adapter by adhesive deformable glue pads (Pritt Multi-Fix). The weight of the adapter was counterbalanced with a constant upward force (0.2 N) produced by the PHANToM. While attached to the PHANToM participants were able to move freely in a 38x27x20 cm workspace.

[Insert Figure 2 about here]

Via the adapter external forces were transmitted by the PHANToM to the index finger vertically, either pushing the finger into the stimulus orthogonal to the stimulus' surface or pulling it out of the stimulus. The amounts of external force were fixed fractions α of the force applied by the participant. The total vertical force was measured every 3 ms with the force sensor. We calculated the force applied by the participant by subtracting the external force transmitted at the previous time point. To avoid amplification of noise produced by the force sensor (causing vibrations), external forces were calculated based on the force average from the last 15 ms. External forces were transmitted only during one indentation of the standard stimulus, i.e., in Experiment 1 during the single performed indentation and in Experiments 2a-d during one of the indentations.

The algorithm to detect and count the indentations was developed and trained using samples of trajectory data (force and vertical position) from free softness explorations (described in Lezkan & Drewing, 2015). The algorithm distinguished between three states in the temporal course of exploratory movements:

1. The state without indentation.
2. Force increase and downward directed vertical displacement.
3. Force decrease and upward directed vertical displacement.

These states alternated circularly as long as the participant explored the stimulus and were detected using different threshold combinations of time, force, vertical position, and derivatives of force and vertical position (Figure 3). When a full cycle of these states was completed, one indentation was counted. The beginning of "force increase" marked the beginning of the indentation and the end of "force decrease" marked its end.

[Insert Figure 3 about here]

In the beginning of a trial the state was set to "no indentation". If a minimal force (1N, measured by the force sensor below the stimulus, Figure 1) was reached and minimal increase in force (0.01) and decrease in vertical position (-0.01) were detected, the state changed to "force increase". If after a minimal time from the beginning of the indentation (200ms) the force minimally decreased again (-0.01), the state changed to "force decrease". Finally, if after the minimal time following the onset of "force decrease" (30ms) the force fell below 17N and stopped to decrease (force derivative > -0.01), the indentation was considered to be terminated ("no indentation") and the next cycle could begin. The thresholds used by the algorithm were optimized to produce less than 5% misses (absent or wrong detection of an indentation) as compared to careful visual inspection.

If the indentation was the one to be manipulated, transmission of external force started with the beginning of the indentation. To prevent sudden on- and offsets at the indentation' beginning and end the fraction of participants' forces used to calculate external forces linearly increased to α or decreased to 0 (respectively) within 30 ms.

A custom-made software (C++) controlled the experiment, collected responses, and recorded finger positions and reaction forces. White noise presented via headphones masked possible sounds from the PHANToM's engines when transmitting external forces.

Softness Stimuli

The stimuli were made from two-component silicon rubber solution (AlpaSil EH 10:1), which was mixed with varying amounts of a diluent (polydimethylsiloxane, viscosity 50 mPa·s) to obtain different compliances. The solution was poured into cylindrical plastic dishes (75 mm diameter x 38 mm high) avoiding the formation of air pockets, to obtain flat surfaces without any discriminable differences in texture. After the stimuli were completely cured, compliance was measured using the experimental apparatus but replacing the finger-adapter at the PHANToM arm by a flat–ended cylindrical probe of 1 cm^2 area (‘standard finger’). The probe was repeatedly pressed into the stimulus using sufficiently high forces (15-25N), to warrant that enough data was sampled in the analyzed force range (0-9 N). The compliance was calculated as the slope of the regression line, fitted to the measured displacement–force traces (example plot in Figure 4). For analysis we used only the trajectories caused by the increase of force, to exclude hysteresis effects during the decrease of force. Possible biases from non-uniform data sampling due to manual indentation of the stimuli were reduced by calculating the mean displacement in 2N steps for bins of +/-0.4 N. For further details and discussion on the measurement method see Kaim and Drewing (2011).

We produced two sets of rubber stimuli, each consisting of one standard and ten comparison stimuli. As standards we used one rather hard (0.32 mm/N) and the other rather soft (0.67 mm/N) stimulus. In each set half of the comparisons had increasingly lower and the other half increasingly higher compliance as compared to the standard. The compliance difference between two neighbored comparison stimuli was about 1/2 Weber fraction (0.03 mm/N step for the hard and 0.05 mm/N step for the soft set) and the range covered by the comparisons was about 2.5 Weber fractions in each direction. Different Weber fractions of about 20% and 15% were used for the hard and the soft stimulus sets respectively (values taken from Kaim & Drewing, 2011). To reduce the traces of usage we produced each stimulus in two similar versions and alternated them. The compliance values of the stimuli are listed in Table 2.

[Insert Table 2 about here]

Design

Experiment 1. The experimental design comprised two within-participant variables: *Compliance of Standard* $c \in [0.32, 0.67]\frac{mm}{N}$ and the fraction of *External Force* $\alpha \in [-.16, -.11, 0, .11, .16]$, resulting in 10 conditions. We measured individual Points of Subjective Equality (PSE) of the manipulated standard stimulus as compared to non-manipulated comparison stimuli for each condition. For that purpose, we used a Two-Interval Forced Choice (2IFC) task combined with a 1-Up-1-Down staircase paradigm. We also estimated individual Just Noticeable Differences (JND) from fitted psychometric functions. In Experiment 1 each stimulus was indented once.

Experiments 2a-d. The Experiments 2a-d differed by the *Exploration length*, as instructed by the number of indentations of each stimulus, $N \in [2,3,4,5]$. The experimental design of each Experiment 2a-d comprised the two within-participant variables *Number of Intervening Indentations* (0,…, *N-1*) and the fraction of *External Force* $\alpha \in [-.16, 0, .16]$. The compliance of the standard was always c = 0.67 mm/N. An external pulling or pushing force was applied during a single indentation of the standard stimulus or was not applied at all. We distinguished between different indentations by their distance to the exploration of the comparison stimulus, as expressed by the *Number of Intervening Indentations*. For example, when the standard was explored first the standard's last indentation had the minimum distance (0 intervening indentations) to the exploration of the comparison and when it was explored second, its first indentation had the minimum distance. The order in which the standard and the comparison were presented was balanced and randomized. Each Experiment 2a-d comprised 2**N* conditions with external forces and one baseline condition without. Again, we measured individual PSEs for each standard stimulus using a 2IFC softness discrimination task and a 1-Up-1-Down staircase paradigm. From the PSEs we calculated the weights of single indentation-specific estimates. To assess the gain of sequential redundancy we estimated for each Experiment 2a-d JNDs from fitted psychometric functions.

Procedure

There were up- and downwards directed staircases. In the first trial of the downwards directed staircase the standard was paired with the comparison stimulus of highest compliance in the set. The upwards directed staircase started with the comparison stimulus of lowest compliance. The comparison for trial *j* in a staircase depended on the participant's response in trial *j-1* of this staircase. If the comparison had felt softer (harder) than the standard, the next comparison in the staircase was less soft (hard). If the calculated comparison was out of the range of the staircase, the same comparison was presented again in the next trial. The estimation of the PSE and JND by one staircase was considered terminated after 10 reversals (changes of direction in the staircase because participants changed their judgment from softer to harder and vice versa), which were reached on average after 18.43 trials (SD = 1.67).

Each of the five experiments was split into two sessions, which were completed on two days within one week, except for Experiment 2a which was completed on the same day.

In each session the estimation of the PSE and JND of each condition was completed by one upward and one downward directed staircase. There were thus in total four staircases, two per condition and session. Each session consisted of blocks in which the current step of each staircase was presented once in random order (number of trials = number of conditions*2). Toward the end of the session, the number of trials in one block decreased, because staircases were increasingly terminating. The sessions were interspersed with pauses (1 min every 45 trials, about every 15 min). In the first session participants completed a practice session prior to the experiment consisting of 8 trials to familiarize them with the setup and the task. After the last session participants completed a survey in which they reported whether they noticed differences between the trials, which technique they used to compare the softness of the two stimuli, and how they experienced the experiment overall.

In the beginning of each trial the stimulus to be touched first (standard or comparison) was displayed on the computer screen (left or right). A tone presented via the headphones signaled the participant to start the exploration. Participants were instructed to touch each stimulus in its center, which was visually rendered as a cross on the stimulus representation. Depending on the design of each experiment (1, 2a-d) participants were instructed how many times to indent each stimulus. The standard and the comparison stimuli were always indented the same number of times. As soon as the first stimulus was touched, the second stimulus appeared in the visual scene and after its exploration, participants reported which stimulus felt softer by moving their finger to one of the two virtual decision buttons displayed above the stimuli. They did not receive any feedback about the correctness of their response. If participants used more or less indentations than instructed, the trial was repeated later in the block. Between trials, participants moved their finger to an indicated position in the corner of the 3D-scene to wait until the experimenter had manually changed the stimuli. The position (left, right) of the standard was randomized by the computer program.

Data analysis

Psychometric functions. In both Experiments 1 and 2a-d in order to assess PSEs and JNDs, we calculated for each participant, each condition and each comparison stimulus the percentage of trials in which it was perceived to be softer than the standard. These values, combined for all comparisons composed the individual psychometric data, to which we fitted cumulative Gaussian functions using the psignifit4 toolbox (Schuett, Harmeling, Macke, & Wichmann, 2016). From the fitted psychometric functions, we estimated the PSE as the 50% discrimination threshold and the JNDs as the 84% discrimination threshold, corresponding to the standard deviation of the cumulative Gaussian function (Helbig & Ernst, 2007). The goodness of fit of the psychometric functions was assessed by comparing the measure of *deviance* (D) to the critical χ^2 value for 10 comparisons, $\chi^2_{10;95\%} = 18.31$ (Wichmann & Hill, 2001). *Deviance* is the log-likelihood ratio between the full model (one parameter for every observation) and the fitted model (achieved by maximum likelihood) - the smaller the *deviance*, the better the fit.

We estimated PSEs using psychometric functions instead of averaging over compliances at reversals, in order to eliminate the influence of trials in which the manipulation of force was not warranted and to be able to analyse the data separately for the cases that the standard was explored first or second. Estimates from both methods are highly correlated (R = 0.95).

Experiment 1. We assumed that external forces (calculated as the fraction α of force applied by participants) would proportionally increase (pushing forces, $+\alpha$) or decrease (pulling forces, $-\alpha$) perceived softness ($\hat{c}_m$) as compared to perceived softness without external forces ($\hat{c}_0$; Metzger & Drewing, 2015):

$$\hat{c}_m = (w_f \alpha + 1)\hat{c}_0 . \qquad (5)$$

with w_f being the extent by which external forces are translated into a perceptual change. To calculate w_f we performed linear regressions of individual PSE shifts with external forces relative to the PSE without external force ($\frac{\hat{c}_m - \hat{c}_0}{\hat{c}_0}$) on α. According to Eq. (5) the relative PSE shift caused by external force is a linear function of α (cf. Landy, Maloney, Johnston & Young, 1995):

$$\frac{\hat{c}_m - \hat{c}_0}{\hat{c}_0} = w_f \alpha \ . \tag{6}$$

with w_f being the slope. We expected the slope to be positive and used a one-tailed *t*-test to test this hypothesis. Two-sided *t*-tests were used to analyze the regression intercepts and to compare the regression parameters between the *Compliance of standard* conditions. To verify that external forces do not affect perceptual reliability, we additionally assessed individual JNDs for each *External Force* condition and performed a one-way repeated measures ANOVA.

Experiments 2a-d. In order to verify that the manipulation of perceived softness was successful in Experiments 2a-d, i.e., that external force was transmitted in the target segment and only in the target segment, force applied by the participant and external force were visually inspected for every trial. Trials in which the manipulation was not successful were excluded for the fit of psychometric functions (total of 0.9% trials). Participants for whom more than 5% of trials had to be excluded were excluded from the analysis.

To analyze perceptual reliability we assessed individual JNDs for each *Exploration length* (1-5 indentations, Experiments 1,2a-d), by averaging over the *External Force* conditions in both Experiments 1 and 2 and additionally over the *Number of Intervening Indentations* conditions in Experiments 2a-d. Individual JNDs were entered in a one-way

ANOVA with the between participant factor *Exploration length.* To test whether perceptual reliability decreased with increasing length of the exploration, we performed a trend analysis using linear contrasts. We wanted to analyze whether the JNDs decreased as predicted by the MLE model. According to Eq. 2 integration of N statistically independent estimates with the same variance σ^2 results in a variance of $\frac{\sigma^2}{N}$ for the combined estimate, which predicts that the JNDs, which are proportional to $\sqrt{\sigma^2}$ (Ernst & Banks, 2002), decrease with the number of indentations N by the factor of $N^{-1/2}$ (Quick, 1974; Drewing et al., 2011). We log-transformed the JNDs and the total number of indentations in order to linearise the relationship, fitted a linear regression model to this data and compared the decrease in JNDs to the predicted decrease which corresponds to a slope of -0.5 in of the regression function.

In our experiments JNDs were estimated from psychometric data sampled with the 1-Up-1-Down staircase. Simulations show that JNDs are proportionally underestimated in similar cases (Leek, Hanna & Marshall, 1992). We replicated the simulation of Leek et al. (1992) for the estimation of JNDs with our 1-Up-1-Down staircase and found a small proportional bias of 12%. However, a proportional bias in the JND estimates might change the intercept but not the slope of the analyzed decrease of JNDs with the extension of the exploration, because we perform it on log-transformed data. Nevertheless, due to this proportional distortion caution is advised when generally comparing JNDs estimated from a 1-Up-1-Down staircase to JNDs estimated with more appropriate methods (e.g. constant stimuli).

For each condition in Experiments 2a-d we estimated individual PSEs. To confirm that the manipulation of perceived softness was successful, we entered the PSEs in ANOVAs with the within-participant factor fraction of *External Force*, separately for each Experiment 2a-d.

From the PSEs we calculated individual indentation-specific weights w_i for each indentation i and each *Exploration length* (2-5 indentations, Experiments 2a-d). We assumed that the overall percept (as assessed by the PSE) was the result of a weighted linear combination of indentation-specific estimates (Eq. 1) and that indentation-specific estimates of non-manipulated indentations equal the perceived softness without external force $\hat{c}_0$. According to the results of Experiment 1, external forces transmitted during the exploration, shifted perceived softness following Eq. (6). In a multi-segmented exploration, the manipulation of a single indentation i would then shift the overall percept $\hat{c}_{mi}$ relatively to $\hat{c}_0$ by:

$$\frac{\hat{c}_{mi} - \hat{c}_0}{\hat{c}_0} = w_i w_f \alpha \ . \qquad (7)$$

To assess $w_i w_f$ we performed a linear regression of the relative PSE shift $\frac{\hat{c}_{mi} - \hat{c}_0}{\hat{c}_0}$ on α. w_f (0.3) was estimated from the results of Experiment 1 as the average of the factors w_f for rather soft and rather hard stimuli.

To confirm that the weights sum up to 1 as predicted by Eq. (1) we performed a *t*-test on individual weight sums against one. To analyze whether the weights decreased with the *Number of Intervening Indentations*, we performed linear regressions of the weights on the number of intervening indentations, for each participant and exploration length and computed a *t*-test of individual slopes against 0.

We additionally investigated whether the weight of each estimate depended on whether the stimulus was explored first or second. For this purpose, we estimated the PSEs, using only trials in which the standard was the first or the second stimulus (respectively) and repeated the analysis reported above. Additionally, we analyzed for each exploration length whether the pattern of weights differed between the first and the second stimulus (interaction *Number of Intervening Indentations* X *Presentation of the Standard* as first vs. as second stimulus in two-way ANOVA) and whether the decrease of weights with the number of intervening indentations depended on the length of the exploration (one-way ANOVA on the slopes of the weights with between-participant factor *Exploration Length*). We did this analysis only for the second stimulus, because only there we found a decrease of weights.

EXPERIMENT 1

Results

Participants did not report in the surveys that they noticed any differences between the trials. Thus, it can be assumed that they were not aware of external forces. For all external force conditions, we observed a shift in the PSEs in the predicted direction: to a softer percept with pushing external forces and to a harder percept with pulling external forces. In Figure 5, the PSEs with and without force manipulation are plotted as a function of α (fraction of force participants applied themselves used to calculate external forces). The average regression slopes (w_f, Eq. 6) were significantly positive for both standards: hard: $t(9) = 4.80$, $p < .001$; soft: $t(9) = 5.93$, $p < .001$ (one-tailed tests). The intercepts were not significantly different from zero for both, the hard, $t(9) = 1.18$, $p = 0.268$ and the soft, $t(9) = 0.95$, $p = 0.365$ standards. The regression functions were $\frac{\hat{c}_m - \hat{c}_0}{\hat{c}_0} = 0.28 * \alpha + 0.03$ for the hard standard and $\frac{\hat{c}_m - \hat{c}_0}{\hat{c}_0} = 0.33 * \alpha + 0.01$ for the soft standard (dashed black lines in Figure 5). Paired *t*-tests on individual slopes and intercepts showed no significant difference in the average slope, $t(9) =$ -

0.72, p = .489[1], nor in the average intercept, $t(9) = 0.74$, p = .476, between the hard and soft conditions.

The one-way repeated measures ANOVA on the JNDs did not reveal a significant effect of the within-participant factor *External Force*, $F(4,36) = 0.22$, $p = 0.928$, indicating that external forces did not change perceptual reliability.

[Insert Figure 5 about here]

Discussion

Force feedback has been previously used for virtual and augmented softness (Biggs & Srinivasan, 2002; Jeon & Choi 2009, 2011). However, the effect of external forces had not yet been studied for the exploration of compliance stimuli with bare fingers. In Experiment 1, we showed that haptically perceived softness of deformable silicon rubber stimuli can be manipulated during the exploration with bare fingers. Perceived softness was changed by transmission of subtle external forces (calculated as a fraction of the forces α participants exerted themselves). PSEs shifted as a linear function of α. The same stimuli were judged to be softer when pushing forces were transmitted, and judged to be harder with pulling forces. We argued that the linear relationship between α and the resulting PSE shift reflected the exclusive perturbation of the force estimate provided by the kinesthetic system and the efference copy (Metzger & Drewing, 2015). As kinesthetic afferent subsystem we subsume the receptors in the muscles and tendons and the efference copy of the motor command (Wolpert & Flanagan, 2001) whereas mechanoreceptors innervating the skin are referred to as the cutaneous afferent subsystem. It is likely that the kinesthetic force receptors and the efference copy only inform about self-produced forces (Jones, 1986). Hence, if assuming

[1] Please note, average slopes were erroneously reported to be significantly different for the soft and the hard standard in Metzger & Drewing, 2015.

further that the kinesthetic and the cutaneous subsystem provide separate softness estimates which are integrated with constant weights (Eq. 1), the slope (w_f) in the regression function (Eq. 6) corresponds to the weight of the kinesthetic estimate.

However, the impact of external forces might have been not as exclusive. The two estimates (cutaneous and kinesthetic) might not have been perfectly disentangled. For instance the change in contact area as sensed by the cutaneous afferent subsystem could have been also interpreted as displacement of the finger (Moscatelli et al, 2016). It is also possible that external forces triggered resistance forces (similar to resistance forces reported for the ocular system, Stark and Bridgeman, 1983), resulting in an additional efference copy, which would have reduced the impact of external forces on the kinesthetic estimate. Furthermore, averaging and ramping in the calculation of external forces (cf. *Apparatus and setup*) introduced small asynchronies and nonlinearities of external forces with respect to participants' forces possibly modulating the perceptual effect (cf. impact of delays in softness perception; Di Luca, Knoerlein, Ernst & Harders; Knoerlein, Di Luca & Harders, 2009; Leib, Karniel & Nisky, 2015; Pressman, Welty, Karniel & Mussa-Ivaldi, 2007). Here, we used the effect of external forces on perceived softness, and the average extent w_f (0.3, Eq. 6) to which α is translated into a perceptual difference in order to study the integration of sequential information in Experiments 2a-d.

EXPERIMENTS 2A-D

Results

Psychometric functions: Overall, the participants' responses were well fit by cumulative Gaussian functions: The *deviance* values were below the critical $\chi^2_{10;95\%} = 18.31$ in more than 99% of all fitted psychometric functions. When splitting the data into the sets where the standard was explored as the first or as the second stimulus, the *deviance* values were still below the critical χ^2 values in more than 95% of the psychometric fits in these datasets.

JNDs. The one-way ANOVA on the JNDs revealed a significant main effect of the between participant factor *Exploration length*, $F(4,55) = 3.27$, $p = 0.018$ and the trend analysis confirmed that the JNDs linearly decreased with an increasing number of indentations, $F(1,55) = 7.89$, $p = .007$ (Figure 6). Averages and standard errors of logarithmized JNDs are plotted as a function of logarithmized exploration length (number of indentations) in Figure 6. The slope of the linear regression function (solid line) was -0.141. The 95% confidence interval of the slope, ranging from -0.884 to -0.048, did not include the value of -0.5 predicted by the MLE (dotted line), indicating a shallower decrease of JNDs with increasing length of the exploration as compared to the prediction of the MLE.

[Insert Figure 6 about here]

PSEs. For each Experiment 2a-d average PSEs for each fraction of external force transmitted in each indentation are plotted in Figure 7. As expected from the results in Experiment 1 overall the pushing external force shifted the PSEs to a softer percept, and a pulling external force shifted it to a harder percept. The separate one-way ANOVAs revealed a significant main effect of *External Force* in all Experiments 2a-d (all p's $< .008$), indicating successful manipulation of perceived softness.

[Insert Figure 7 about here]

Weights. The average weights are plotted as a function of the number of intervening indentations to the comparison, separately for each length of the exploration (Experiments 2a-d) in Figure 8. A *t*-test showed that (consistent with Eq. 1) for none of the exploratory lengths (*N*) the average sum of the weights was significantly different from 1 (all p's > 0.5). The average slope of the regression of weights on the number of intervening indentations was significantly negative for the exploration lengths of 2 indentations, $t(9) = -4.96$, $p < .001$ and 3 indentations, $t(7) = -3.23$, $p = .015$ but not for 4 indentations, $t(12) = .06$, $p = .95$ and 5 indentations, $t(18) = -2.09$, $p = .051$.

[Insert Figure 8 about here]

In Figure 9 the weights of indentation-specific estimates are plotted separately for the first and the second stimulus. In both cases the sum of the weights was not significantly different from 1 for all exploration lengths (all p's > 0.2). The ANOVA on individual weights with the within-participant factors *Number of Intervening Indentations* and *Presentation of the Standard* (as first vs. as second stimulus) revealed a significant interaction for the exploration lengths of 2 indentations, $F(1,9) = 36.07$, $p < .001$, 3 indentations, $F(2,14) = 10.07$, $p = .002$, was at the edge of significance for 4 indentations, $F(3,36) = 2.81$, $p = .053$ and not significant for 5 indentations, *F*(4~~3~~,72) = 1.06, *p*= .~~372~~385, indicating differences between the first and the second stimulus in the pattern of weights for the shorter exploration lengths. When the standard was explored first the slopes of linear regressions of the weights on the *Number of Intervening Indentations* were not significantly different from 0 for all exploration lengths: 2 indentations, $t(9) = 0.23$, $p = .820$; 3 indentations, $t(7) = 1.19$, $p = .274$; 4 indentations, $t(12) = 0.84$, $p = .417$; 5 indentations: $t(18) = -0.35$, $p = .733$. When the standard was explored as the second stimulus, the slopes were significantly negative for 2 indentations, $t(9) = -9.35$, $p < .001$, 3 indentations, $t(7) = -4.53$, p = . 0.003, and 4 indentations $t(12) = -3.03$, $p = .010$ but not for 5 indentations $t(18) = -1.22$, $p = .240$. A one-way ANOVA on the slopes of the weights in the exploration of the second stimulus with the *Exploration length* as the between-participant factor revealed that the slopes were affected by the *Exploration length*, $F(3,46) = 49.65$, $p < .001$. A trend analysis showed that the slopes increased with the decreasing exploration length, $F(1,46) = 134.48$, $p < .001$.

[Insert Figure 9 about here]

Discussion

As predicted, JNDs in softness discrimination decreased with an increasing length of the exploration. This finding demonstrates that participants were able to increase perceptual reliability by accumulating information about softness over time, and is in concordance with earlier reported redundancy gains as a result of spatiotemporal extension of the exploration (Drewing et al., 2011; Giachritsis et al., 2009; Lezkan & Drewing, 2014; Louw et al., 2005). Also as expected, the decrease in the JNDs with increasing exploration length was significantly smaller than would have been predicted by the MLE model ($N^{-0.14}$ vs. $N^{-0.5}$ for N indentations). Drewing et al. (2011) and Lezkan and Drewing (2014) found similarly flat decreases of JNDs ($N^{-0.12}$, $N^{-0.15}$) with increasing length of the exploration of virtual gratings. Hence, we can extend the conclusion that the integration of sequential information in haptic perception does not follow the MLE model, to a naturalistic task.

To approach the question, why sequential information was not integrated according to the MLE model, we assessed the weights of softness estimates derived from single exploratory segments (i.e., indentations). In agreement with previous observations on the perception of virtual gratings (Lezkan & Drewing, 2014), we found that the weights of indentation-specific estimates decreased significantly as a function of the indentation's distance to the comparison - at least for the short explorations comprising 2-3 indentations. Distinguishing between the indentation-specific weights on the first and the second stimulus in the 2IFC task, we found that the estimates from indentations of the first stimulus were weighted approximately equal, and that only the estimates gathered during the exploration of the second stimulus were weighted unequally, decreasing with increasing distance to the comparison stimulus (significant for explorations comprising 2-4 indentations). We did not find a significant decrease in weights over time for the exploration of the second stimulus comprising 5 indentations. Smaller effects in weights were expected with an increasing number of indentations (weights sum to 1), thus the differences between the weights might have been too small to be measured. We conclude that in the natural integration of sequential haptic information, softness estimates from different movement segments of the exploration of the second stimulus are unequally weighted, at least for explorations comprising less than 5 indentations.

The question arises why observers did not use all of the information provided by the repeated exploration of the stimulus. Estimates with correlated noises do not provide as much information as statistically independent estimates (Oruc et al., 2003). However, correlations between signal estimates do not predict that estimates with similar reliabilities are integrated with unequal weights, as we find for the integration of information on the second stimulus' softness.

Unequal weights of estimates derived in the same sense on the same environmental quality by repeated exploration (comparable in length to a typical haptic exploration [~2s]) were also observed in several studies on summary statistics (review: Hubert-Wallander & Boynton, 2015), which addressed the ability of humans to estimate the mean value of serially presented visual cues (e.g. mean size of a dot that changed size over time). Some of these studies observed so-called *recency effects* (late information is weighted higher; Cheadle et al., 2014; Toscani, Zdravković & Gegenfurtner, 2016) others observed *primacy effects* (early information is weighted higher; Drugowitsch, Moreno-Bote, Churchland, Shadlen & Pouget, 2012). Whereas recency effects cannot explain the pattern of our results, a primacy effect could be responsible for higher weights of early estimates of the second stimulus' softness.

For primacy effects it is hypothesized that early information is weighted higher, due to increasing costs of sampling and processing of information. In our experiments participants were instructed to indent each stimulus a fixed number of times. It is possible that in order to reduce the costs and effort of sampling and processing, participants had decided which stimulus felt softer before the end of the exploration of the second stimulus. At the same time, it is a feasible assumption that resources were not saved during the exploration of the first stimulus, because the difficulty of the decision was unknown at this time point. From such a strategy, equal weights during the exploration of the first stimulus and a decrease of weights during the exploration of the second stimulus over time might be explained. However, this explanation does not fit with observations on the muscular effort that participants spend over the time course of the exploration: participants significantly increased peak force in their last indentation as compared to all previous indentations of the first $t(49) = -9.14$, $p < .001$, as well as the second stimulus, $t(49) = -4.76$, $p < .001$ (paired t-tests comparisons of peak forces averaged over indentations $1,\ldots,n-1$ and peak forces in the n-th indentation).

As already outlined, in the introduction for the integration of information over time a Kalman-filter approach might be more appropriate than the MLE. Remember, that for the estimation of a static variable the Kalman-filter approach yields the same estimate and the same reliability as the MLE model, if the process noise is negligible. However, unequal weights of estimates could be expected in our experiment if the process noise were large. According to Eq. (3) high process noise would bias the estimate towards the current sensory measurement in every iteration. In our experiment where we manipulated always one of the estimates this would result in higher weights if the manipulated estimate was acquired later in the exploration. However, we instead observed a decrease of weights towards the end of the exploration and only in the exploration of the second stimulus. Thus, high process noise cannot explain the pattern of weights we found in our experiment.

Given different patterns of weights of indentation-specific estimates of the first and the second stimulus, it might be assumed that there are different underlying mechanisms in information processing for these two stimuli. There is indeed an important difference - participants could compare the softness of both stimuli only when they explored the second stimulus. Deciding which stimulus is softer requires them to keep the perceived softness of the first stimulus in memory while they explore the second stimulus. Romo, Hernández and Zainos (2004) studied the neural correlates of such perceptual decisions. Monkeys were trained to compare two vibrotactile stimuli with frequencies *f1* and *f2*, which were sequentially presented to the fingertips, and to report which stimulus had higher frequency. The authors recorded single neurons in the ventral premotor cortex (VPC).

The VPC receives input from the sensory areas and sends projections to motor areas (Godschalk, Lemon, Kuypers & Ronday, 1984; Matelli, Camarda, Glickstein & Rizzolatti, 1986; Luppino, Murata, Govoni & Matelli, 1999; Lu, Preston & Strick, 1994). Its neurons exhibit both sensory (Rizzolatti & Luppino, 2001; Graziano, Hu & Gross, 1997; Graziano, Reiss & Gross, 1999) and motor (Gentilucci, Fogassi, Luppino, Matelli, Camarda & Rizzolatti, 1988) related response properties, and inactivation of VPC impairs sensorimotor tasks (Fogassi, Gallese, Buccino, Craighero, Fadiga & Rizzolatti, 2001). Therefore, the VPC is thought to likely be involved in the mechanisms that link sensory events with actions (Romo et al., 2004).

Romo et al. (2004) found neurons in the VPC which encoded *f2* and two other populations representing the decisions $f1 < f2$ and $f1 > f2$. Yet other neurons which the authors labeled "partially differential neurons", encoded *f1* during the presentation of the first stimulus and were modulated by *f2* afterwards. However, there were no neurons coding *f1* during the presentation of the second stimulus. The question arose how the difference between the frequencies of the two stimuli was computed.

On the basis of this work, Deco et al. (2010) suggested a synaptic mechanism to model the activity of the "partially differential neurons" and their role in decision making. They proposed that the memory of $f1$ is realized by short-term synaptic facilitation of the "partially differential neurons". The phenomenological model of short-term synaptic facilitation is based on residual calcium accumulated during the firing of the cells increasing the probability of transmitter release (Mongillo, Barak & Tsodyks, 2008). This means that as long as $f1$ is presented, recurrent connections between the selective neurons are strengthened. When $f2$ is presented, the activity of these selective neurons depends on the synaptic history and on $f2$. Deco et al. (2010) proposed that the activity of these neurons might represent the sum of $f1$ and $f2$, from which the difference between $f1$ and $f2$ can be read out by comparing it with the activity of neurons representing $f2$. This difference might then be the input to a standard attractor based-decision-making network comprising competing neuronal populations representing the decisions $f1 < f2$ and $f1 > f2$.

Our observations on information processing in a comparison of haptic softness fit well with the model of perceptual decision making proposed by Deco et al. (2010). If we assume that for softness discrimination there are "partially differential neurons" in the VPC or an analogous area, they would encode the softness of the first stimulus c_1 during its presentation. Our results suggest that in this case, the gathered information equally contributes to the overall percept. This might be implemented by e. g. computing the mean over the activity of neurons in lower areas, which show a tuning to certain softness values (possible model in Knudsen, Lac & Esterly, 1987; Metzger & Drewing, 2016). When the second stimulus with softness c_2 is presented, the "partially differential neurons" would reflect the sum of $c1$ and $c2$ which would be compared to the activity of the neurons encoding $c2$. Since the short-term memory of $c1$ is implemented by synaptic properties of "partially differential neurons", it decays over time. Accordingly, the estimates from the indentations in the beginning of the exploration of the second stimulus would be weighted higher, because they can reliably be compared to the estimated softness of the first stimulus. Furthermore, also the observation that weights decreased less rapidly in a longer exploration is consistent with the synaptic model of perceptual decisions. If the memory of *c1* is implemented as calcium based synaptic facilitation of the "partially differential neurons", it could be expected that in a certain range, longer explorations facilitate the neurons more than shorter ones (Deco et al., 2010; Mongillo et al., 2008). Thus the decay of the memory of *c1* after the exploration of the first stimulus is more rapid the shorter the exploration.

Taken together, the model of perceptual decision making by Deco et al. (2010) can well explain the observed pattern of unequal weights during softness discrimination. We correspondingly suggest that our observation of sequential integration with unequal weighting leads back to the the decay of the representation of the softness of the first stimulus in memory, which is stronger the longer the stimulus is explored.

General Discussion and conclusions

In the present study we showed that haptically perceived softness can be manipulated using subtle external forces. The value of perceived softness can be shifted to softer or to harder percepts by pushing or pulling forces. For the multi-segmented exploration of softness, we found that the redundant information is not being exploited according to the MLE model. Unequal weights of indentation-specific estimates of the second stimulus are most likely the reason for the lower reliability gain than predicted by the MLE. The results are well explained by a model of synaptic dynamics in a perceptual decision task (Deco et al., 2010), that attributes our findings to decay of the memory about the first stimulus during the discrimination task. Noise introduced by memory effects is considered neither in the MLE nor in the Kalman filter model. These models in their simplest form assume constant measurement noise (Eq. 1 and 3), which might change over time. However, in an extended Kalman filter model further assumptions on the sources of noise e.g. memory decay could be included to model the integration of estimates of the second stimulus' softness. For this purpose it is necessary to determine the relevant factors influencing memory decay (e.g. does it progress with time or with acquisition of information about the second stimulus?). Our results further indicate that also decision making processes might be involved in this kind of information integration, so a combination of an information integration model (e.g. MLE/Kalman filter) with decision making demands might be required. Such model of the integration of serially sampled information might then represent a theoretically optimal integration in the sense of exploiting *actually* available information in a decision making task.

REFERENCES

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current biology*, *14*, 257-262.

Anderson, B. D., & Moore, J. B. (1979). Optimal filtering. *Englewood Cliffs*, *21*, 22-95.

Bergmann Tiest, W. M., & Kappers, A. M. (2009). Cues for haptic perception of compliance. IEEE Transactions on Haptics, 2, 189-199.

Bryson, A. E., & Ho, Y.C. (1975). *Applied optimal control*. New York: Wiley.

Burge, J., Ernst, M. O., & Banks, M. S. (2008). The statistical determinants of adaptation rate in human reaching. *Journal of Vision, 8*, 20-20.

Cellini, C., Kaim, L., & Drewing, K. (2013). Visual and haptic integration in the estimation of softness of deformable objects. *i-Perception*, *4*, 516-531.

Cheadle, S., Wyart, V., Tsetsos, K., Myers, N., De Gardelle, V., Castanon, S., H., & Summerfield, C. (2014), Adaptive gain control during human perceptual choice. *Neuron, 81*, 1429–1441.

Cochran, W. G. (1937). Problems arising in the analysis of a series of similar experiments. *Supplement to the Journal of the Royal Statistical Society, 4*, 102-118.

Di Luca, M., Knoerlein, B., Ernst, M. O., & Harders, M. (2011). Effects of visual–haptic asynchronies and loading–unloading movements on compliance perception. *Brain research bulletin, 85*, 245-259.

Deco, G., Rolls, E. T. & Romo, R. (2010). Synaptic dynamics and decision making. *Proceedings of the National Academy of Sciences, 107*, 7545-7549.

Drewing, K., Lezkan, A. & Ludwig, S. (2011). Texture discrimination in active touch: Effects of the extension of the exploration and their exploitation. In *World Haptics Conference (WHC), 2011 IEEE*, 215-220.

Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., & Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *The Journal of Neuroscience, 32*, 3612-3628.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*, 429-433.

Ernst, M. O., Banks, M. S., & Buelthoff, H. H. (2000). Touch can change visual slant perception. *Nature Neuroscience, 3*, 69–73.

Ernst, M. O., & Buelthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in cognitive sciences*, *8*, 162-169.

Fetsch, C. R., DeAngelis, G. C., & Angelaki, D. E. (2010). Visual–vestibular cue integration for heading perception: applications of optimal cue integration theory. *European Journal of Neuroscience, 31*, 1721-1729.

Fogassi, L., Gallese, V., Buccino, G., Craighero, L., Fadiga, L., & Rizzolatti, G. (2001). Cortical mechanism for the visual guidance of hand grasping movements in the monkey. *Brain, 124*, 571-586.

Gentilucci, M., Fogassi, L., Luppino, G., Matelli, M., Camarda, R., Rizzolatti, G. (1988). Functional organization of inferior area 6 in the macaque monkey. I. Somatotopy and the control of proximal movements. *Experimental Brain Research, 71*, 475-490.

Gescheider, G. A., Bolanowski, S. J., Pope, J. V., & Verrillo, R. T. (2002). A four-channel analysis of the tactile sensitivity of the fingertip: frequency selectivity, spatial summation, and temporal summation. *Somatosensory and motor research, 19*, 114-124.

Giachritsis, C. D., Wing, A. M., & Lovell, P. G. (2009). The role of spatial integration in the perception of surface orientation with active touch. *Attention, Perception, and Psychophysics, 71*, 1628-1640.

Godschalk, M., Lemon, R.N., Kuypers, H.G., & Ronday H.K. (1984). Cortical afferents and efferents of monkey postarcuate area: an anatomical and electrophysiological study. *Experimental Brain Research, 56*, 410–424.

Graziano, M.S., Hu, X.T., & Gross, C.G. (1997). Visuospatial properties of ventral premotor cortex. *Journal of Neurophysiology, 77*, 2268–2292.

Graziano, M.S., Reiss, L.A., & Gross, C.G. (1999). A neuronal representation of the location of nearby sounds. *Nature, 397*, 428–430.

Hartcher-O'Brien, J., Di Luca, M., & Ernst, M. O. (2014). The duration of uncertain times: audiovisual information about intervals is integrated in a statistically optimal fashion. *PloS one, 9*, e89339.

Helbig, H. B., & Ernst, M. O. (2007). Optimal integration of shape information from vision and touch. *Experimental Brain Research, 179*, 595–606.

Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of vision, 4*, 967–992.

Hubert-Wallander, B., & Boynton, G. M. (2015). Not all summary statistics are made equal: Evidence from extracting summaries across time, *Journal of vision, 15*, 1-12.

Jacobs, R. A., & Fine, I. (1999). Experience-dependent integration of texture and motion cues to depth. *Vision Research, 39*, 4062–4075.

Johnson, K.O., & Phillips J.R. (1981). Tactile spatial resolution: I. Two-point discrimination, gap detection, grating resolution, and letter recognition. *Journal of Neurophysiology, 46*, 1177-1191.

Jones, L. A. (1986). Perception of force and weight: theory and research. *Psychological bulletin, 100*, 29-42.

Kaim, L., & Drewing, K. (2011). Exploratory strategies in haptic softness discrimination are tuned to achieve high levels of task performance. *IEEE Transactions on Haptics, 4*, 242-252.

Klein, S. A. (2001). Measuring, estimating, and understanding the psychometric function: a commentary. *Perception & Psychophysics, 63*, 1421–55.

Knoerlein, B., Di Luca, M., & Harders, M. (2009). Influence of visual and haptic delays on stiffness perception in augmented reality. In *Proceedings of the 8th IEEE International Symposium on Mixed and Augmented Reality*, Orlando, FL, 2009 (pp. 49-52).

Knudsen, E. I., Lac, S. d., & Esterly, S. D. (1987). Computational maps in the brain. *Annual review of neuroscience, 10*, 41-65.

Koerding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, *427*, 244-247.

Kwon, O. S., Tadin, D., & Knill, D. C. (2015). Unifying account of visual motion and position perception. *Proceedings of the National Academy of Sciences, 112*, 8142-8147.

Landy, M. S., Banks, M. S., & Knill, D. C. (2011). Ideal-observer models of cue integration. In J. Trommershaeuser, K. P. Koerding & M. S. Landy, (Eds.), *Sensory cue integration*. (pp. 5‑29). New York, NY: Oxford University Press.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision research, 35*, 389-412.

Lederman, S. J., & Klatzky, R. L. (1987). Hand movements: A window into haptic object recognition. *Cognitive psychology*, *19*(3), 342-368.

Leib, R., Karniel, A., & Nisky, I. (2015). The effect of force feedback delay on stiffness perception and grip force modulation during tool-mediated interaction with elastic force fields. *Journal of neurophysiology, 113*, 3076-3089.

Lezkan, A., & Drewing, K. (2014). Unequal but Fair? Weights in the Serial Integration of Haptic Texture Information. In M. Auvray and C. Duriez, (Eds.), *Haptics: Neuroscience, Devices, Modeling, and Applications: 9th International Conference, EuroHaptics 2014, Versailles, France, June 24-26, 2014, Proceedings, Part I* (pp. 386-392). Vol. 8618. Berlin Heidelberg: Springer.

Lezkan, A., & Drewing, K. (2015). Predictive and sensory signals systematically lower peak forces in the exploration of softer objects. *World Haptics Conference (WHC), 2015 IEEE* (pp. 69-74). IEEE.

Louw, S., Kappers, A.M.L., & Koenderink, J.J. (2005). Haptic detection of sine-wave gratings. *Perception, 34*, 869-885.

Lu, M. T., Preston, J. B., & Strick, P. L. (1994). Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. *Journal of Comparative Neurology, 341*, 375-392.

Luppino, G., Murata, A., Govoni, P., & Matelli, M. (1999). Largely segregated parietofrontal connections linking rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4). *Experimental Brain Research, 128*, 181–187.

Matelli, M., Camarda, R., Glickstein, M., & Rizzolatti, G. (1986). Afferent and efferent projections of the inferior area 6 in the macaque monkey. *Journal of Comparative Neurology, 251*, 281-298.

Metzger, A., & Drewing, K. (2015). Haptically perceived softness of deformable stimuli can be manipulated by applying external forces during the exploration. In *World Haptics Conference (WHC), 2015 IEEE* (pp. 75-81). IEEE.

Metzger, A., & Drewing, K. (2016). Haptic Aftereffect of Softness. In F. Bello, H. Kajimoto & Y. Visell (Eds.), *Haptics: Perception, Devices, Control, and Applications: 10th International Conference, EuroHaptics 2016, London, UK, July 4-7, 2016, Proceedings, Part I* (pp. 23-32). Berlin Heidelberg: Springer.

Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic theory of working memory. *Science*, *319*, 1543-1546.

Moscatelli, A., Bianchi, M., Serio, A., Terekhov, A., Hayward, V., Ernst, M. O., & Bicchi, A. (2016). The Change in Fingertip Contact Area as a Novel Proprioceptive Cue. *Current Biology*, *26*, 1159–1163.

Mulligan, R. M., & Shaw, M. L. (1980). Multimodal signal detection: Independent decisions vs. integration. *Perception & Psychophysics*, *28*, 471-478.

Oostwoud Wijdenes, L.O., Marshall, L., & Bays P.M. (2015). Evidence for Optimal Integration of Visual Feature Representations across Saccades. *The Journal of Neuroscience, 35*, 10146 –10153.

Oruç, I., Maloney L.T., & Landy M.S. (2003). Weighted linear cue combination with possibly correlated error. *Vision research, 43*, 2451-2468.

Plaisier, M. A., van Dam, L. C., Glowania, C., & Ernst, M. O. (2014). Exploration mode affects visuohaptic integration of surface orientation. *Journal of Vision*, *14*, 22.

Pressman, A., Welty, L. J., Karniel, A., & Mussa-Ivaldi, F. A. (2007). Perception of delayed stiffness. *The International Journal of Robotics Research*, *26*(11-12), 1191-1203.

Quick Jr, R. (1974). A vector-magnitude model of contrast detection. *Kybernetik, 16*, 65-67.

Rao, R. P. (1999). An optimal estimation approach to visual perception and learning. *Vision research*, *39*, 1963-1989.

Rizzolatti, G., & Luppino G. (2001). The cortical motor system. *Neuron, 31, 889–901.*

Romo, R., Hernández, A., & Zainos, A. (2004). Neuronal correlates of a perceptual decision in ventral premotor cortex. *Neuron, 41, 165–173.*

Rosas, P., Wagemans, J., Ernst, M. O., & Wichmann, F. A. (2005). Texture and haptic cues in slant discrimination: reliability-based cue weighting without statistically optimal cue combination. *Journal of the Optical Society of America, 22*, 801-809.

Schuett, H. H., Harmeling, S., Macke, J. H., & Wichmann, F. A. (2016). Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision research*, *122*, 105-123.

Shaw, M. L. (1982). Attending to multiple sources of information. I. The integration of information in decision making. *Cognitive Psychology, 14*, 353-409.

Stark, L., & Bridgeman, B. (1983). Role of corollary discharge in space constancy. *Perception & Psychophysics*, *34*, 371-380.

Swets, J. A., Shipley, E. F., McKey, M. J., & Green, D. M. (1959). Multiple observations of signals in noise. *The Journal of the Acoustical Society of America*, *31*, 514-521.

Swets, J. A., & Birdsall, T. G. (1978). Repeated observation of an uncertain signal. *Perception and Psychophysics, 23*, 269-274.

Toscani, M., Zdravković, S., & Gegenfurtner, K. R. (2016). Lightness perception for surfaces moving through different illumination levels. *Journal of Vision*, *16*, 1-18.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception and psychophysics, 63*, 1293-1313.

Van Beers, R. J., Sittig, A. C., & van der Gon, J. J. D. (1998). The precision of proprioceptive position sense. *Experimental Brain Research*, *122*, 367-377.

Van Beers, R. J., van Mierlo, C. M., Smeets, J. B., & Brenner, E. (2011). Reweighting visual cues by touch. *Journal of vision*, *11*, 1-16.

Wolf, C., & Schuetz, A. C. (2015). Trans-saccadic integration of peripheral and foveal feature information is close to optimal. *Journal of Vision, 15,* 1-18.

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current biology, 11*, R729-R732.

TABLES

Table 1

Main characteristics of the final sample.

Experiment	Participants	Mean age	Age range	Female/Male	Duration (h)
1, 1 indentation	10	23.8	19-29	4/6	5
2a, 2 indentations	10	24.3	19-27	6/4	2.5
2b, 3 indentations	8	26.9	21-32	4/4	4
2c, 4 indentations	13	24.5	21-29	9/4	5
2d, 5 indentations	19	25.3	21-35	12/7	6

Table 2

Compliances of the silicone rubber stimuli.

Set	Compliance (mm/N)										
Hard	0.16	0.19	0.23	0.26	0.29	0.32	0.36	0.39	0.43	0.46	0.49
Soft	0.41	0.47	0.52	0.56	0.62	0.67	0.72	0.77	0.82	0.87	0.92

FIGURES

Figure 1

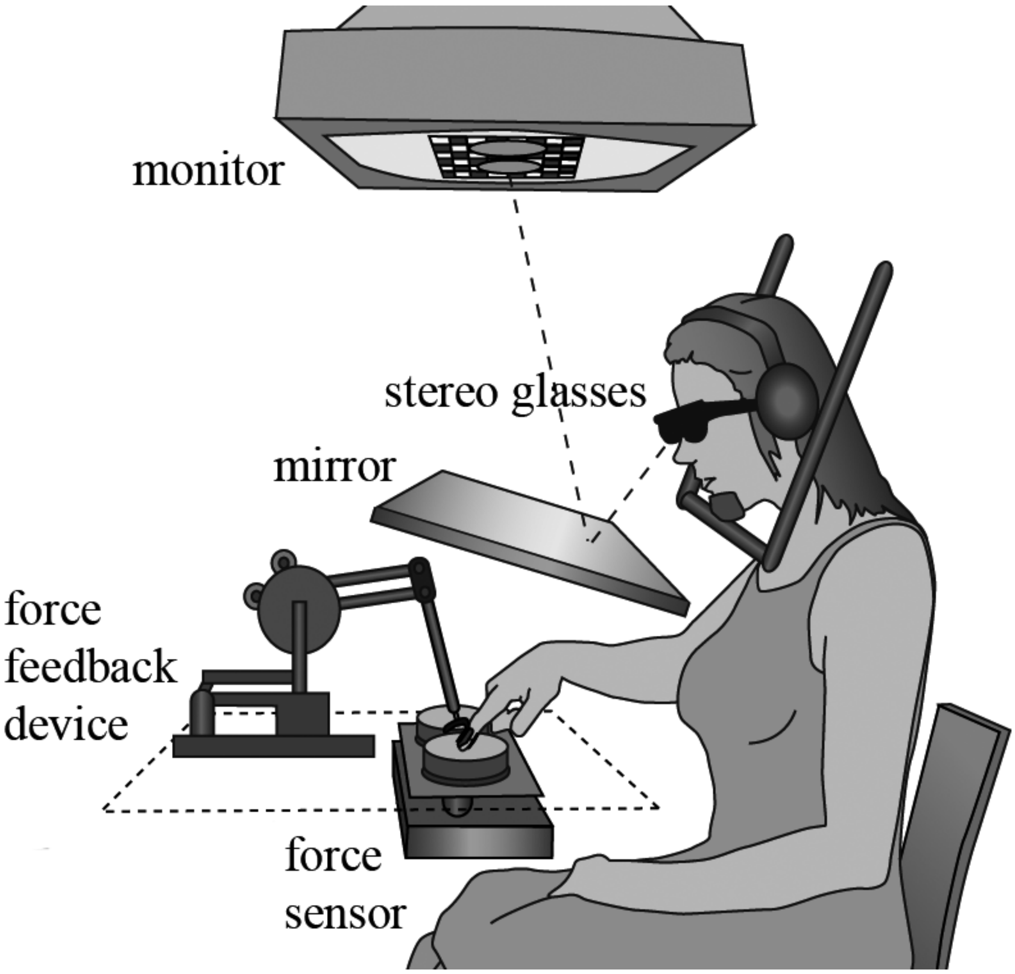

Figure 1. Visuo-haptic workbench. Real stimuli were placed in front of the participant on the force sensor next to each other (distance 2 cm). A visual representation of the stimuli was displayed on the screen and viewed via stereo glasses. The head of the participant was stabilized by a head and a chin rest. The index finger of the dominant hand was connected via a custom-made adapter to the PHANToM, which was used to measure the position of the finger and to apply external forces during the exploration. White noise was presented via headphones to mask sounds from the motors of the PHANToM.

Figure 2

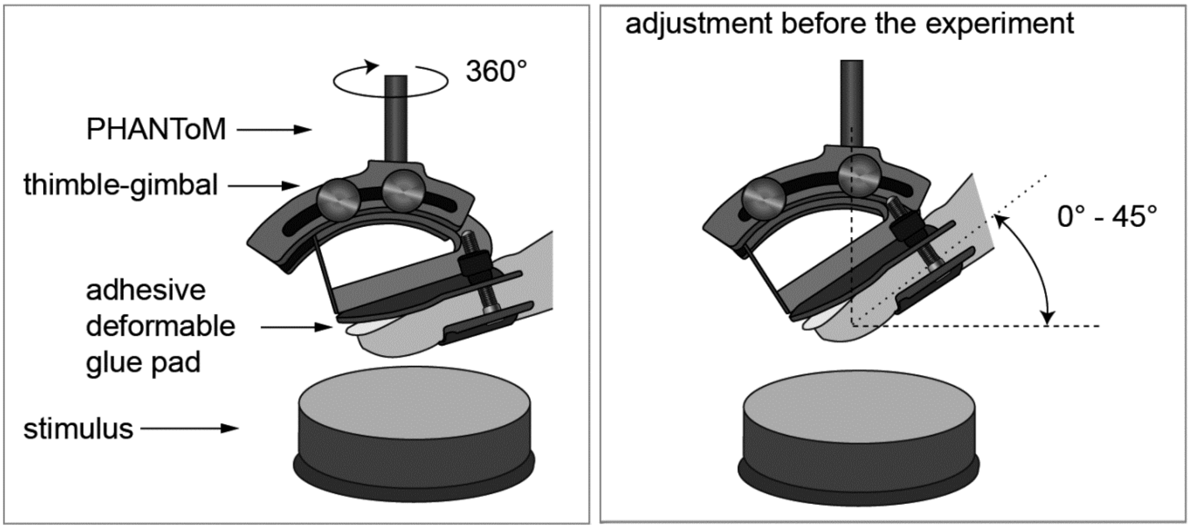

Figure 2. Adapter to connect the index finger of the participant to the PHANToM. The adapter was designed to leave the finger pad uncovered and to ensure the transmission of forces, by restricting the vertical degrees of freedom. Preferred vertical inclination between 0° to 45° was adjusted before the experiment. Due to a circular design the direction of the external force and the calibrated zero-position was unchanged by the adjustment. The adapter was additionally affixed to the dorsal side of the finger by adhesive deformable glue pads.

Figure 3

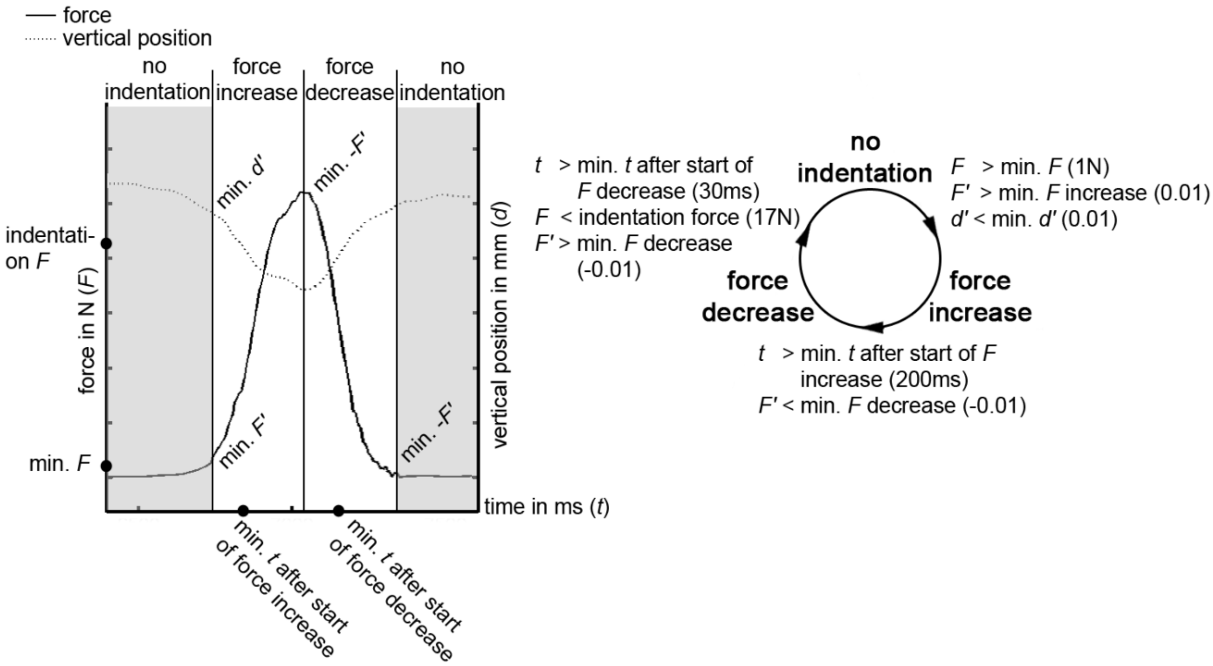

Figure 3. Detection of indentations. Depicted are the force (solid line) and vertical finger position (dotted line) during a single indentation of a stimulus in Exp. 1 as a function of time. The thresholds (absolute values and first derivatives) used to detect the indentation are symbolically indicated by black dots in the axes and by text inside the plot. The right part outlines schematically how these thresholds were used to distinguish the different states (min. = minimum/minimal). The states are depicted in bold font. The circle indicates in what order the states changed. Conditions for the change from one state to another are listed between the corresponding two states.

Figure 4

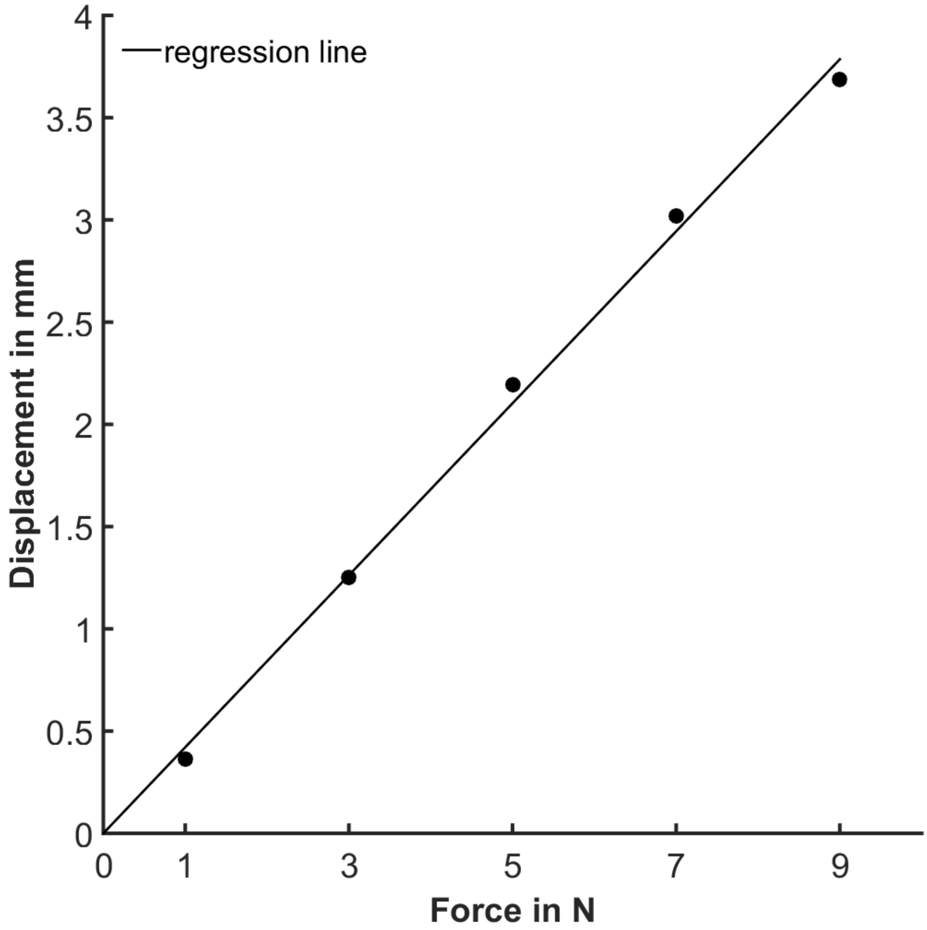

Figure 4. Compliance measurement. A force-displacement function of an exemplar stimulus as used in our Experiments. The average displacements calculated for every 2N steps for bins of +/-0.4 N are plotted as black solid dots. The regression line with the slope corresponding to the compliance of the stimulus is plotted as a black line.

Figure 5

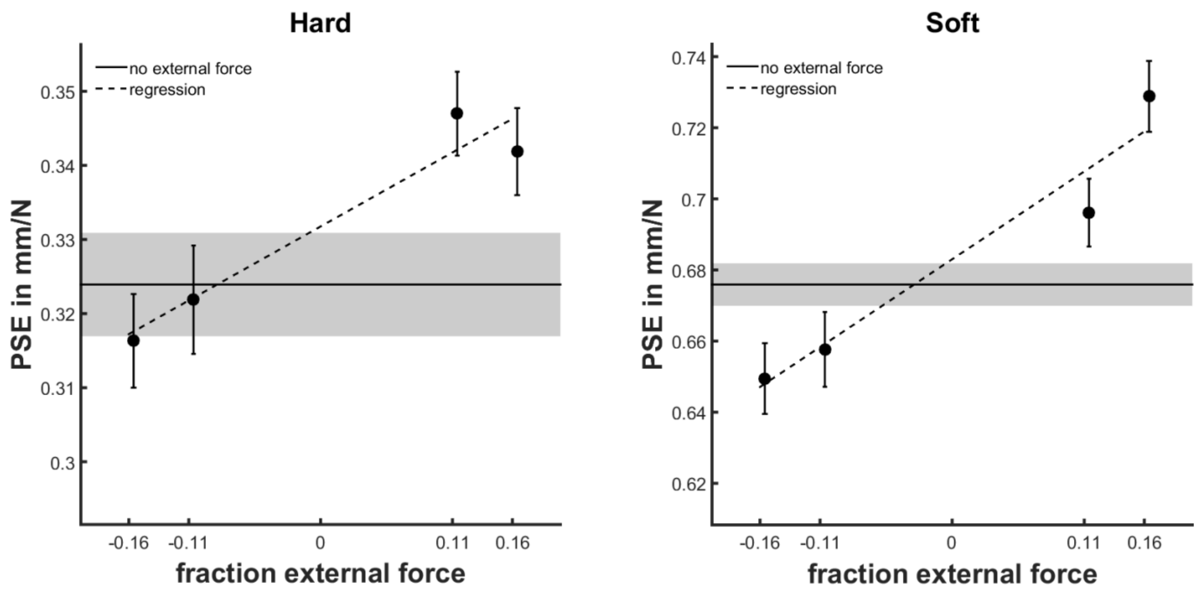

Figure 5. Average PSEs (black dots) with standard errors as a function of the fraction of external force separately for the soft and the hard standard. The average PSE without external forces is plotted as a solid black line. The respective standard error is indicated by the grey area. Additionally regression lines are plotted as black dotted lines.

Figure 6

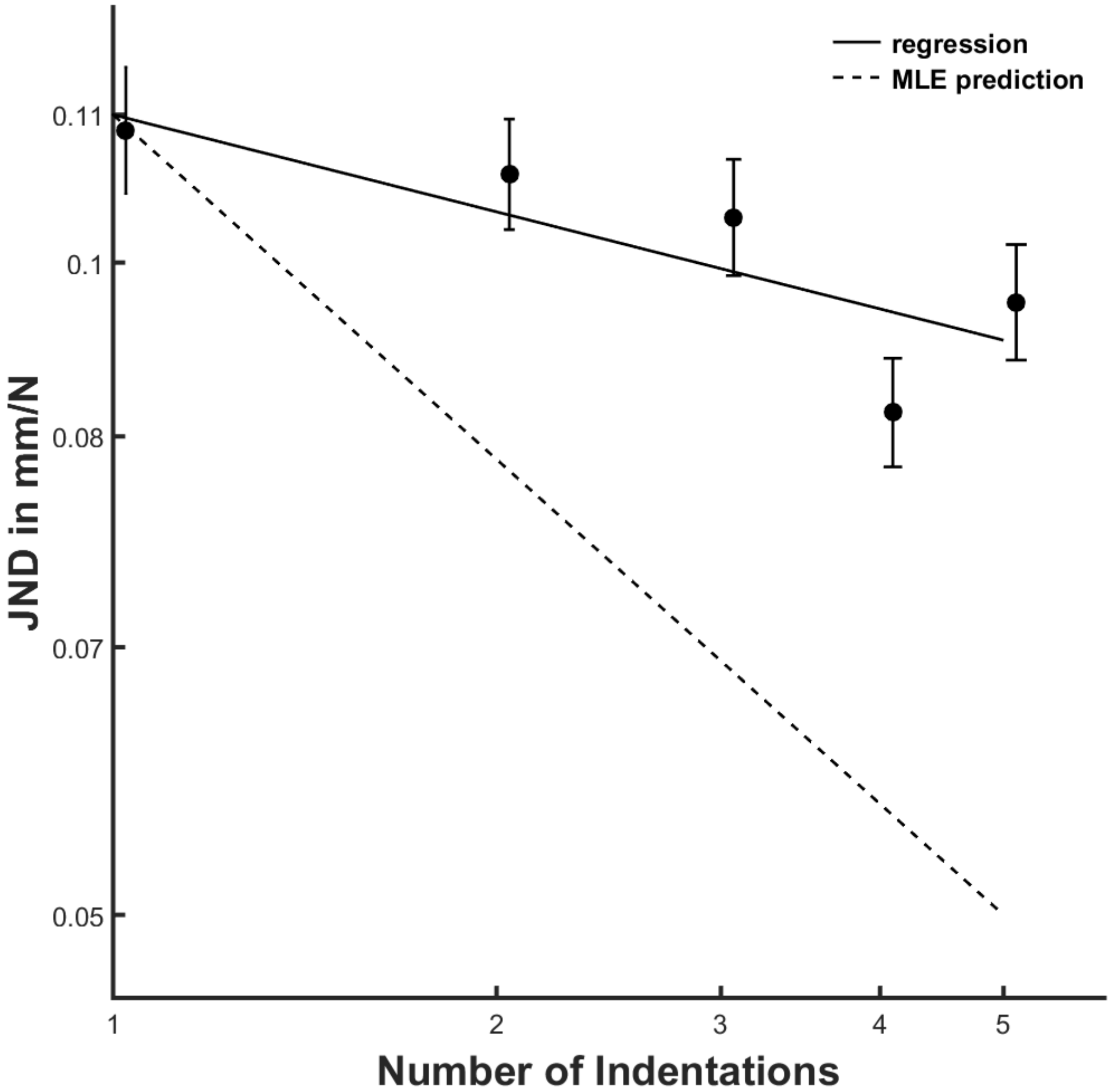

Figure 6. Average JNDs and standard errors as a function of the Number of Indentations in Log-Log space. The regression line is plotted as a solid line and the dashed line indicates integration yielding maximal possible reliability.

Figure 7

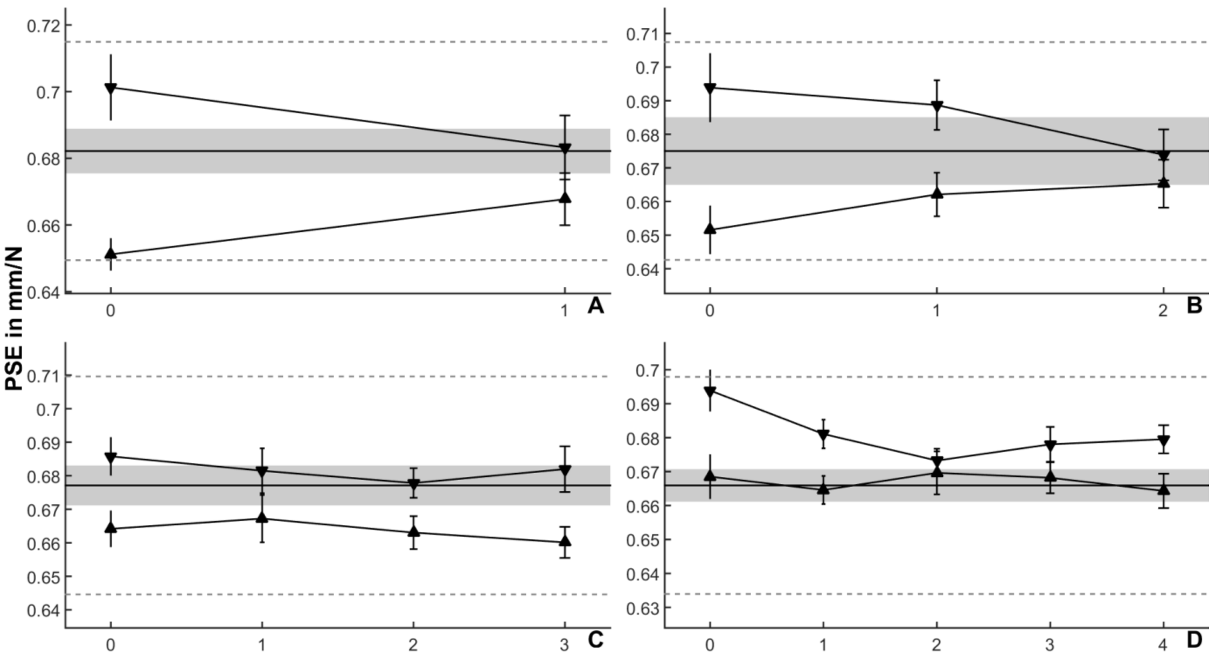

Figure 7. PSE-shift as a function of external force and the number of intervening indentations. The perceived softness is plotted separately for each Experiment 2a-d (exploration lengths of 2-5 indentations) as a function of the number of intervening indentations to the comparison. The mean values are plotted as black and individual measurements as grey triangles. Average and individual PSEs as well as the standard error are plotted for both external force conditions (pushing force and pulling force, downward and upward pointing triangles respectively). For each length of the exploration average perceived softness in the condition without external forces is plotted as a solid line. The respective standard error is indicated by the grey area. Additionally, the two grey dotted lines indicate the two PSEs, which would result if the manipulation of one indentation would entirely determine perceived softness (maximal possible shift, according to the effect of manipulating a single-indentation exploration in Experiment 1). This value could have been reached if the softness estimate of one indentation would be weighted by 1.

Figure 8

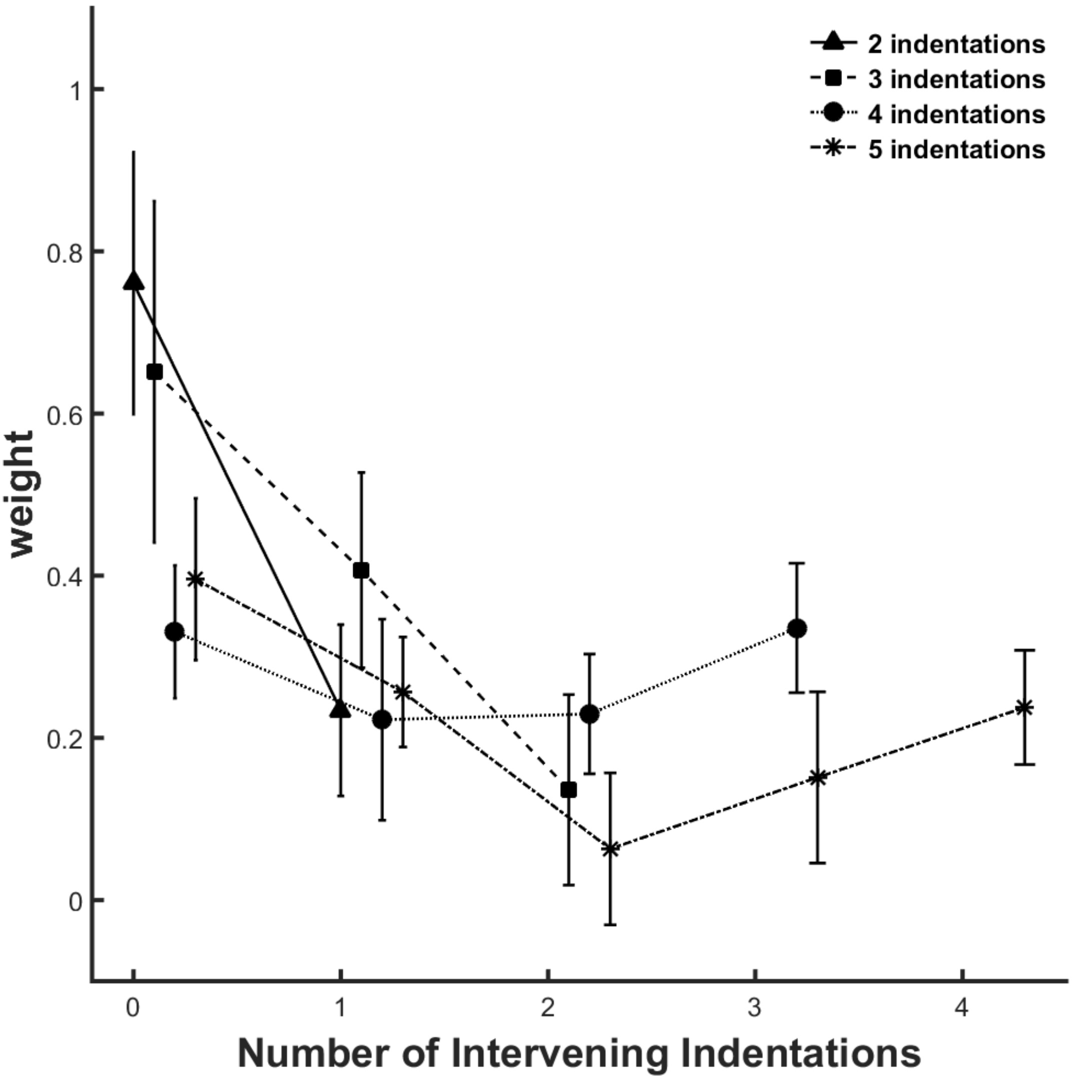

Figure 8. Weights of the estimates derived from single indentations as a function of their temporal distance to the comparison stimulus. The average weights of the estimates and their standard errors are plotted separately for each length of exploration.

Figure 9

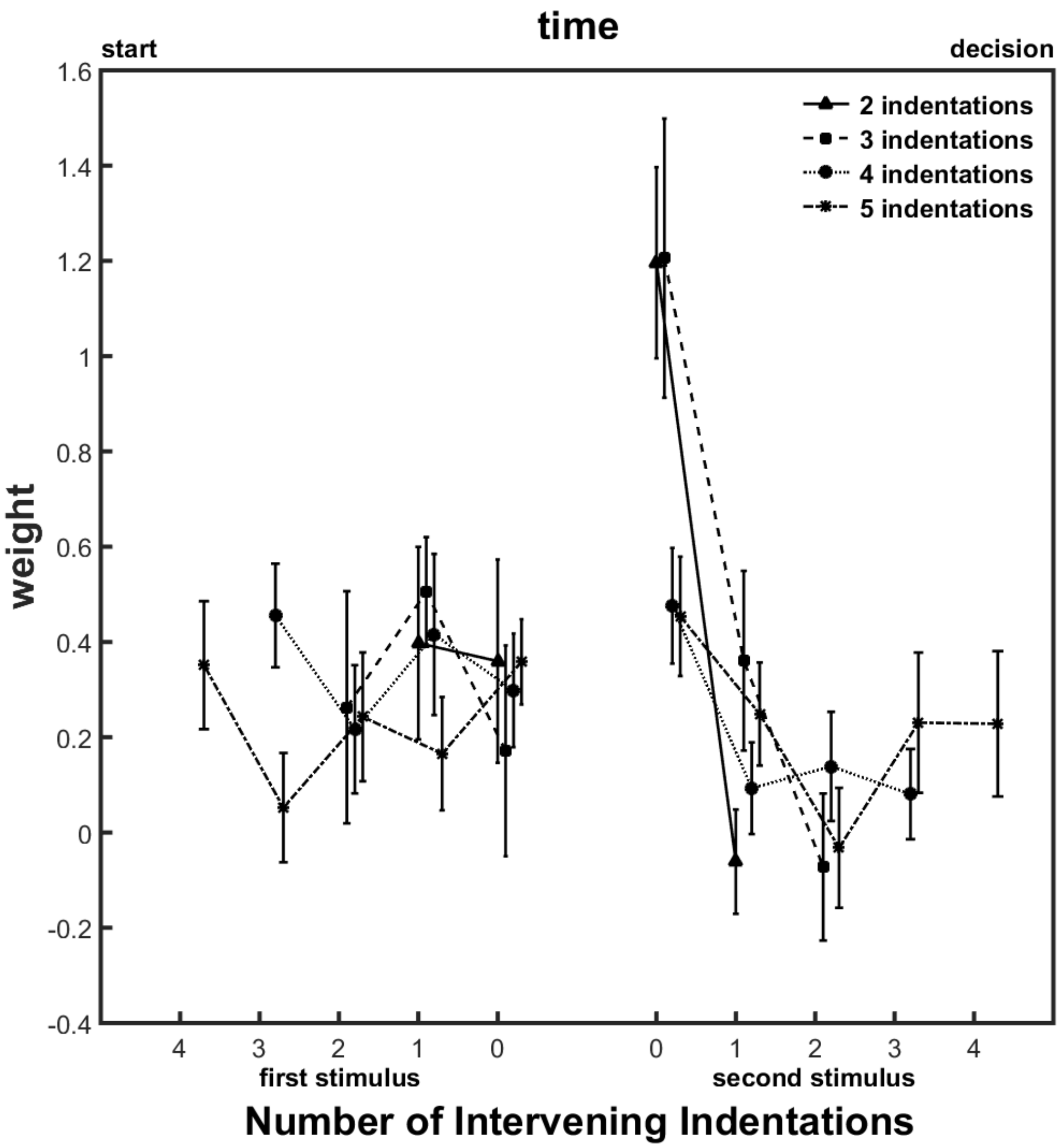

Figure 9. The weights of the estimates from single indentations on the first and the second stimulus separately as a function of their distance to the comparison stimulus and time. The average weights of estimates from single indentations of the first and the second stimulus and their standard errors are plotted separately for each length of the exploration. Time is depicted in arbitrary units.