AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: Human Perception and Performance

0096-1523/19/$12.00

2019, Vol. 45, No. 9, 1271–1284
http://dx.doi.org/10.1037/xhp0000670

Complexity Can Facilitate Visual and Auditory Perception

Cameron T. Ellis and Nicholas B. Turk-Browne
Yale University

Visual and auditory inputs vary in complexity. For example, driving in a city versus the country or listening to the radio versus not are experiences that differ in complexity. How does such complexity impact perception? One possibility is that complex stimuli demand resources that exceed attentional or working memory capacities, reducing sensitivity to perceptual changes. Alternatively, complexity may allow for richer and more distinctive representations, increasing such sensitivity. We performed five experiments to test the nature of the relationship between complexity and perceptual sensitivity during movie clip viewing. Experiment 1 revealed higher sensitivity to global changes in audio or video streams for clips with greater complexity, defined both subjectively (judgments by independent coders) and objectively (information-theoretic redundancy). Experiment 2 replicated this finding but found no evidence that it resulted from complexity drawing attention. Experiment 3 provided a boundary condition by showing that change detection was unaffected by complexity when the changes were superimposed on, rather than dispersed throughout, the clips. Experiment 4 suggested that the effect of complexity, at least when defined objectively, was present without the working memory demands of the preceding experiments. Experiment 5 suggested that complexity led to richer representations of the clips, as reflected in enhanced long-term memory. Collectively, these findings show that, despite increasing informational load, complexity can serve to ground and facilitate perceptual sensitivity.

Public Significance Statement
It is commonly assumed that the perceptual systems struggle to cope with an overabundance of sensory information. To the contrary, we report certain cases in which naturalistic stimulus complexity can have an opposite, beneficial effect on perceptual sensitivity.

Keywords: change detection, compression, information theory, load, multisensory perception

In the study of perception it is typically assumed that sensory inputs are extremely complex and that this complexity impairs function. Researchers invoke the idea of "sensory overload"—the perceptual system diminishing in efficacy under the stress of complex stimulation, as a symptom of some mental disorders (Scheydt et al., 2017), a mistake made in marketing (Malhotra, 1984), and a cause of danger in some environments (Lindenmuth, Breu, & Malooley, 1980). That said, whether and how complexity diminishes perceptual sensitivity has received surprisingly little direct investigation. Our goal in this study is thus to characterize the relationship between complexity and perceptual sensitivity for naturalistic, multisensory stimuli across different task demands.

There is some suggestive evidence that complexity impairs perceptual sensitivity. In multiple object tracking, for example, every additional moving object diminishes a participant's ability to track objects (Pylyshyn & Storm, 1988). In the temporal domain, stimuli presented close together in a sequence can mask one another, revealing limits on how much processing is possible within a fixed timespan (Marcel, 1983; Raymond, Shapiro, & Arnell, 1992). Moreover, orthographic complexity affects vowel detection time, slowing reading of more complex languages like Arabic (Abdelhadi, Ibrahim, & Eviatar, 2011). Finally, working memory capacity is lower for objects with greater complexity (Alvarez & Cavanagh, 2004; Brady, Konkle, & Alvarez, 2009) and memory recall is worse for lists of numbers with greater complexity (Mathy & Feldman, 2012). Common across these tasks is the implication that complexity is detrimental to the functioning of the perceptual system.

On the other hand, increased complexity might improve perceptual sensitivity under some task demands. For instance, when complexity is measured by the number of speakers and the duration of speech, more complex auditory clips are better remembered than less complex clips (Potter & Choi, 2006). This may be because complexity provides scaffolding (e.g., schematic or hierarchical structure) that constrains and supports sensory processing. Perceptual load theory argues that complex stimuli, like those with

This article was published Online First July 18, 2019.

Cameron T. Ellis and Nicholas B. Turk-Browne, Department of Psychology, Yale University.

This work was supported by National Institutes of Health Grants R01 EY021755 and R01 MH069456. We thank Judy Fan and the Turk-Browne lab for helpful discussions on the project and Nour Malek for feedback on an early version of the article.

Correspondence concerning this article should be addressed to Cameron T. Ellis, Department of Psychology, Yale University, 2 Hillhouse Avenue, New Haven, CT 06520. E-mail: cameron.ellis@yale.edu

more distracting elements (Lavie & de Fockert, 2003) or when the targets and distractors are more similar (Beck & Lavie, 2005), drain surplus perceptual resources and thus reduce task-irrelevant interference (Lavie, 1995). But even in these examples, complexity does not improve overall performance (i.e., the participants do worse with more load), but rather perceptual load diminishes the effect distractors have on performance (Murphy, Groeger, & Greene, 2016).

Rather than always helping or hurting, it is also possible that complexity has nonmonotonic effects on perception. For example, infants are most engaged by stimuli that are neither too complex nor too simplistic (Kidd, Piantadosi, & Aslin, 2012). If this inverted-U shaped function exists in perception too, complexity may be beneficial up to a point for some tasks.

How complexity is defined may inform its relationship with perception. There are several definitions (Forsythe, 2009), including based on subjective ratings, stimulus properties, or information-theoretic measures (Yu & Winkler, 2013). Information theory is a particularly popular way to characterize complexity and makes available multiple metrics (e.g., Kolmogorov complexity, Shannon entropy, mutual information). Kolmogorov complexity is a prominent and elegant approach, quantifying complexity as the minimum description length needed to generate a stimulus. This principle has been used to assess the complexity of items retained in working memory (Mathy & Feldman, 2012) and of hieroglyphic characters (Chikhman, Bondarko, Danilova, Goluzina, & Shelepin, 2012). Kolmogorov complexity is noncomputable with audio-visual content (Cilibrasi & Vitányi, 2005) but has a tractable approximation, lossless compression—a method to reduce the file size of content without losing any information. Lossless compression has been used to quantify the complexity of images (Yu & Winkler, 2013), as well as the complexity of neural states (Casali et al., 2013). These information-theoretic measures of complexity have been found in some cases to relate to subjective estimates of visual complexity (Chikhman et al., 2012). However, the relationship between complexity and perception remains poorly understood.

To investigate how complexity and perception relate, we sought to build upon methods that have been used previously in three ways. First, we employ movie stimuli whose range and variance of complexity may better approximate naturalistic experience, compared to distilled laboratory tasks such as visual short-term memory or multiple object tracking. Second, these movie clips contain both visual and auditory information, allowing us to characterize joint effects of complexity in multisensory stimuli, rather than by considering each modality separately. Third, we compare results for two very different definitions of complexity—one psychological (coder ratings) and one computational (lossless compression).

Five experiments were conducted to understand the relationship between complexity and perception in both the visual and auditory modalities. Experiment 1 establishes the basic effect that video and audio complexity benefits visual and auditory sensitivity to global changes in a movie clip. Experiment 2 extends and replicates these findings but finds no evidence to suggest that selective attention contributes to this relationship. Experiment 3 further calls into question the role of attention in this effect by showing that detection of transient stimuli superimposed on a clip is less affected by complexity. Experiment 4 suggests the effect of complexity can be found for online processing that does not place demands on working memory. Experiment 5 indicates that complex clips are better remembered, suggesting that they afford richer, more distinctive representations. In most cases, these relationships hold for both subjective and objective measures of complexity. Overall, this research suggests that perception can be enhanced by complex stimuli.

Experiment 1

Experiment 1 aimed to establish the relationship between complexity and perception. Psychophysical metrics were used to compare sensitivity to global changes in movie clips, which differed in video and audio complexity. Because this experiment was exploratory in nature, any pattern of results that is reliable across participants would be interesting, regardless of effect size. The data and relevant stimulus details (e.g., movie clip time stamps, clip complexity scores) for this and all subsequent experiments are available online at https://osf.io/dgjyb/.

Method

Participants. Because this was a novel task and analysis approach, we first conducted a pilot study with 15 participants where we approximated the range of quality change needed to get accuracy values centered around 75%. Based on those results we adjusted the range in another pilot with 16 participants. The accuracy of the participants in this sample was in the desired range and so these levels were carried forward. The analyses of interest (psychometric threshold difference and slope) showed positive trends using a two-tailed participant-wise bootstrapping approach (described below) but were not all significant at $p < .050$. Hence we decided to increase the planned sample size to 24 prior to the start of data collection for the current experiment. To verify that this increase was appropriate, we performed power analyses by simulating what would have happened in the final pilot with a larger sample of similar participants (created by sampling participants with replacement). We found that a sample size of 24 resulted in positive trends for all analyses at least 80% of the time (range: 84–100%). Hence, 24 participants (11 females, aged 18 to 22 [$M = 19.67$, $SD = 1.01$]) reporting normal or corrected-to-normal vision were recruited. Three participants, not included in the total, were excluded because they failed to complete the experiment in time or misunderstood task instructions. All participants received course credit or $12/hr as compensation and provided informed consent to a protocol designed in accordance with the principles expressed in the Declaration of Helsinki and approved by the Institutional Review Board of Princeton University.

Stimuli. An open source film (CC 3.0), *Tears of Steel* (Blender Foundation, 2012), was partitioned into 24 five-second clips (24 frames per second) at 1080p uncompressed quality. These clips depicted a variety of scenes: indoors, outdoors, with people, without people, special effects, and so forth. They were chosen to have a range of objective complexity values for both audio and video, and so these values were uncorrelated between modalities across clips, as explained below.

After the clips were selected, their subjective complexity was rated by separate groups of participants on Mechanical Turk, who were redirected to a Qualtrics survey hosting embedded clips from YouTube. To ensure their ratings were based on the correct modality, participants were exposed to either the video ($N = 56$, 21 females, age 19–35 [$M = 29.93$, $SD = 4.61$]) or audio ($N = 51$, 13 females [1 unspecified], age 19–35 [$M = 27.20$, $SD = 4.27$]) stream. They were asked to "Evaluate the complexity of the clip" using a sliding scale, where complexity was previously defined to them as "how busy, unpredictable or elaborate the clip is." Participants could replay a clip if desired. Clips were all played with the same high quality (i.e., without lossy compression), so that load times were equated across clips. Ratings differed in reliability across modalities: intraclass correlation—a measure of interrater reliability (Shrout & Fleiss, 1979)—revealed poor reliability for audio (0.28) and fair for video (0.45) (Cicchetti, 1994). Overall subjective complexity scores for the audio and video stream of each clip were calculated by averaging ratings across participants. These average audio and video scores were highly correlated across clips $r(22) = 0.71$, $p = .004$, suggesting a relationship between the contents in both modalities, given that participants were only ever presented with one stream.

During debriefing, we asked each rater what definition of complexity they used while rating the clips. We had two blind coders categorize their responses. The blind coders were given five possible categories: (a) "Diversity: The quantity or diversity of different features, reflecting how elaborate, busy or great in detail the clip is. This also includes how easily understandable the clip is." (b) "Predictability: Whether the future of the clip could be guessed by the present moment in the clip." (c) "Feature: Some features are construed as more complex than other features (e.g., talking, special effects or cuts). Hence complexity depends on whether these features are present or not." (d) "Semantic: Are the higher-level semantics of the clips complex or not. For instance, emotional or goal driven scenes convey certain semantic content." (e) "NA: Participant did not answer the question." The coders chose which category best suited the rater's responses with moderate reliability for audio ($\kappa = 0.584$) and higher reliability for video ($\kappa = 0.672$). Averaging across the coders, *Diversity* was by far the most commonly used definition for both audio (average number of subjects: *Diversity*, 36.5; *Predictability*, 2; *Feature*, 5.5; *Semantic*, 5; *NA*, 2) and video (*Diversity*, 36.5; *Predictability*, 5.5; *Feature*, 11; *Semantic*, 0; *NA*, 2) complexity.

The objective complexity of these clips was calculated with an information theoretic measure related to Kolmogorov complexity. This refers to the length of the shortest sequence sufficient to describe an object, and in this way reflects the amount of redundancy present in a data stream (Li & Vitányi, 1997). To quantify redundancy in this way, we employed lossless compression algorithms and determined file size reduction in each modality. This specific metric can be used as a proxy for complexity (Cilibrasi & Vitányi, 2005), although many other objective metrics of complexity exist (Yu & Winkler, 2013). For video complexity, we calculated the redundancy of each frame using the PNG algorithm and then averaged file size reduction across the clip. For audio complexity, file size reduction for the clip was based on the FLAC algorithm. Note that audio is much more compressible than video and so the file size reduction in the audio modality is much greater than in the video modality. This leads to an apparent clustering of audio values when placed on the same "proportion of original file size" scale as video (Figure 2B). As alluded to above, the clips were chosen to have the lowest correlation between audio and video complexity, $r(22) = 0.03$, $p = .391$. This was done by taking a random sample of clips from the movie, correlating the audio and video complexity and then repeating 1000 times to find the set of clips with the lowest correlation. This correlation was intentionally low to quantify separate and joint effects of each modality. Only one objective measure was obtained for each stream and clip, from a deterministic algorithm, precluding analyses of reliability.

We measured perceptual sensitivity in visual and auditory modalities by presenting two clips on each given trial. One clip was the original version described above and the other clip was identical but perceptually degraded in one modality. Participants made a two-interval forced choice (2IFC) by discriminating between the clips and selecting which was degraded. To create a degraded version of each clip, the original audio and video streams were compressed using lossy algorithms. The jpeg2000 codec was used for the video stream, which reduced the file size by a given multiple: 16, 32, 64, 128, or 256. The mp3 codec was used for the audio stream, which reduced the bits per second encoding to a given rate: 128, 64, 32, 16, or 8. These degradation ranges were determined from two pilot experiments to produce comparable levels of 2IFC accuracy across modalities and ensure that we covered the full accuracy range of the psychometric function. The degraded clips were then up-sampled to a higher fidelity format (uncompressed AVI for video and FLAC for audio), preserving the degradation but equating file sizes and load times across levels.

Note that a potential source of confusion is the use of compression at two different points in the design. Specifically, lossless compression was used to estimate objective complexity, but was not applied to the stimuli used in the experiment, nor would have affected stimulus quality, by definition. On the other hand, lossy compression was applied to the stimuli for the degraded version of each clip in the 2IFC task, which was used to assess perceptual sensitivity. Lossy compression is one of many possible ways of introducing a perceptual change (e.g., we could have blurred, changed colors or pitches, inserted blanks, inserted targets, rotated the images, changed the volume, edited in or out objects/people, etc.). We chose to degrade the stimuli because this introduced subtle, parametric, and global distortions that encouraged attention to the entirety of the clips and allowed us to obtain a full range of accuracies.

One concern we had in using lossy compression to introduce perceptual changes is that it might interact with complexity. For example, low complexity clips may provide more flexibility in terms of the amount of lossy compression possible for a given discrimination accuracy (e.g., at the extreme, heavy versus minimal degradation of a solid black screen would be equally noticeable). Different approaches, such as blurring or inserting blanks, may have also interacted with complexity for similar reasons. We did consider other perceptual changes that we felt would not interact with complexity, such as a rotation of video or change of audio volume, but these are blunt and easily detectable distortions, and more importantly, they would require only superficial processing of the clips. The concern was that this would prevent participants from engaging with the content of the clips enough to be affected by stimulus complexity (this worry was supported by the findings of Experiment 3). More content-focused changes to the clips, such as moving or swapping objects or people, would be difficult to achieve seamlessly from an editing standpoint and may still impact complexity. Given these challenges, we decided to continue with lossy compression, knowing that the relationship between complexity and lossy compression is complicated and could be a potential limitation when interpreting the results.

Procedure. Participants were seated in a dimly lit room in front of a large monitor (53 cm wide) and presented with clips using Psychtoolbox (Kleiner et al., 2007) at a comfortable audio level. Figure 1A outlines the structure of this experiment. A fixation dot blinked to indicate the start of a trial. Participants were shown the same clip twice (45° visual angle), separated by a 1000-ms pause. A 1500-ms response window began at the end of the second video. Participants were asked to judge whether the first or second clip was lower in quality. They were not instructed to focus on a particular modality and did not report to us whether they noticed which modality was degraded; instead, they were asked to make a holistic judgment about quality. Whether the first or second clip was lower in quality was randomized. The accuracy of each response was the dependent variable. Upon response or at the end of this response window an intertrial interval began, ranging from 350 to 850 ms. The identity of the movie clip, the modality reduced in quality, and the level of degradation were counterbalanced to produce 240 trials, with 10 breaks interspersed throughout.

Before the start of the experiment, each participant did four practice trials in which they were given feedback about the correct answer and what modality the change was in. These trials were controlled such that participants were shown some trials in which the change was obvious and some trials where the change was almost imperceptible.

Results and Discussion

The aim of Experiment 1 was to investigate how video and audio complexity relates to visual and auditory perceptual sensitivity. We initially expected that the subjective and objective measures of complexity would differ: the subjective measure has the potential to include a variety of components ranging from pixel/bit variability (inverse of redundancy), to the number of objects/sources, to narrative complexity; whereas the objective measure only reflects pixel/bit variability in the data stream. The correlation between subjective and objective audio complexity across clips was strong and positive, $r(22) = 0.93$, $p < .001$, suggesting that subjective audio complexity is based on the temporal variability captured by objective complexity. However, the correlation between subjective and objective video complexity was not significant, $r(22) = -0.21$, $p = .250$, suggesting that subjective video complexity was largely based on different features of the stimulus.

Collapsing across the degree of degradation, participants were 74.2% correct on average when the change was visual and 74.0% when the change was auditory. To quantify the effect of complexity on accuracy, we used the Palamedes toolbox in MATLAB (Prins & Kingdom, 2018) to fit psychometric functions, with parameters for threshold, slope, and lapse rate (guessing rate was not fit because it was a forced choice task), over different levels of degradation for clips with high versus low complexity (median split). The critical parameter was the threshold: the amount of change required for accuracy halfway between chance and the lapse rate. The parameters were fit to the average accuracy across participants at each level of degradation to improve the goodness of fit. The random-effects reliability of these estimates was assessed with bootstrapping tests, by resampling participants with replacement 10,000 times prior to calculating the average accuracies and fitting the functions on each iteration. The resulting sampling distribution of the parameter estimates was used to generate 95% confidence intervals and for hypothesis testing. Note that sampling participants with replacement allows for nonparametric testing of reliability across participants in such pooled analyses (Efron & Tibshirani, 1986; Fan, Hutchinson, & Turk-Browne, 2016; Kim, Lewis-Peacock, Norman, & Turk-Browne, 2014). The intuition for why this works is that to the extent that an effect is reliable across participants, their data should be substitutable with minimal impact on the results, leading to a narrow sampling distribution. Table 1 contains the descriptive statistics for these psychometric analyses.

The threshold for detecting visual changes was lower (better performance) for clips with high compared with low video subjective complexity ($p = .001$; Figure 1B). Similarly, clips with high versus low audio subjective complexity had lower auditory thresholds ($p < .001$). The same pattern of results was obtained for objective complexity, video ($p = .002$; Figure 1C) and audio ($p = .002$). There was also a difference in lapse rate between high and low complexity for subjective and objective definitions, but our definition of threshold as halfway between chance and the lapse rate compensates for this.

To fit the psychometric functions above, we performed a median split on complexity and averaged across the clips in each bin to obtain generally smooth and monotonic data. To evaluate complexity as a continuous variable, we performed an additional analysis at the level of individual clips by calculating an average accuracy across quality levels of each clip. We used linear regression to predict these accuracies from measures of subjective and objective complexity and assessed the robustness of the slope with bootstrap resampling. As shown in Figure 2A and 2B, there was a positive relationship between clip accuracy and both subjective complexity (video: $b = 0.08$, 95% CI [0.00, 0.34], $p = .053$; audio: $b = 0.50$, 95% CI [0.55, 0.73], $p < .001$) and objective complexity (video: $b = 0.40$, 95% CI [0.13, 0.51], $p < .001$; audio: $b = 2.64$, 95% CI [0.60, 0.77], $p < .001$). That is, as the complexity of clips increased, so did the ability of participants to detect perceptual changes.

As noted earlier, there may be an interaction between the perceptual change (lossy compression) and the measure of objective complexity (lossless compression). Namely, different levels of complexity may constrain the amount or type of perceptual change possible, complicating the interpretation of our results. Although this is a potential limitation deserving of further investigation, the subjective complexity results may help. Specifically, this measure does not inherently suffer from the same problem, and so the fact that we observed a positive relationship between complexity and perceptual sensitivity for both objective and subjective definitions suggests that this confound cannot entirely explain our findings. One could make an argument that the subjective measure of complexity may also be compromised to the extent that these ratings are based on the same stimulus features as used by the lossless algorithm. However, this predicts a correlation between subjective and objective complexity, which we only observed in the auditory modality. Thus, the visual findings for subjective complexity, at least, seem to be unaffected by this concern.

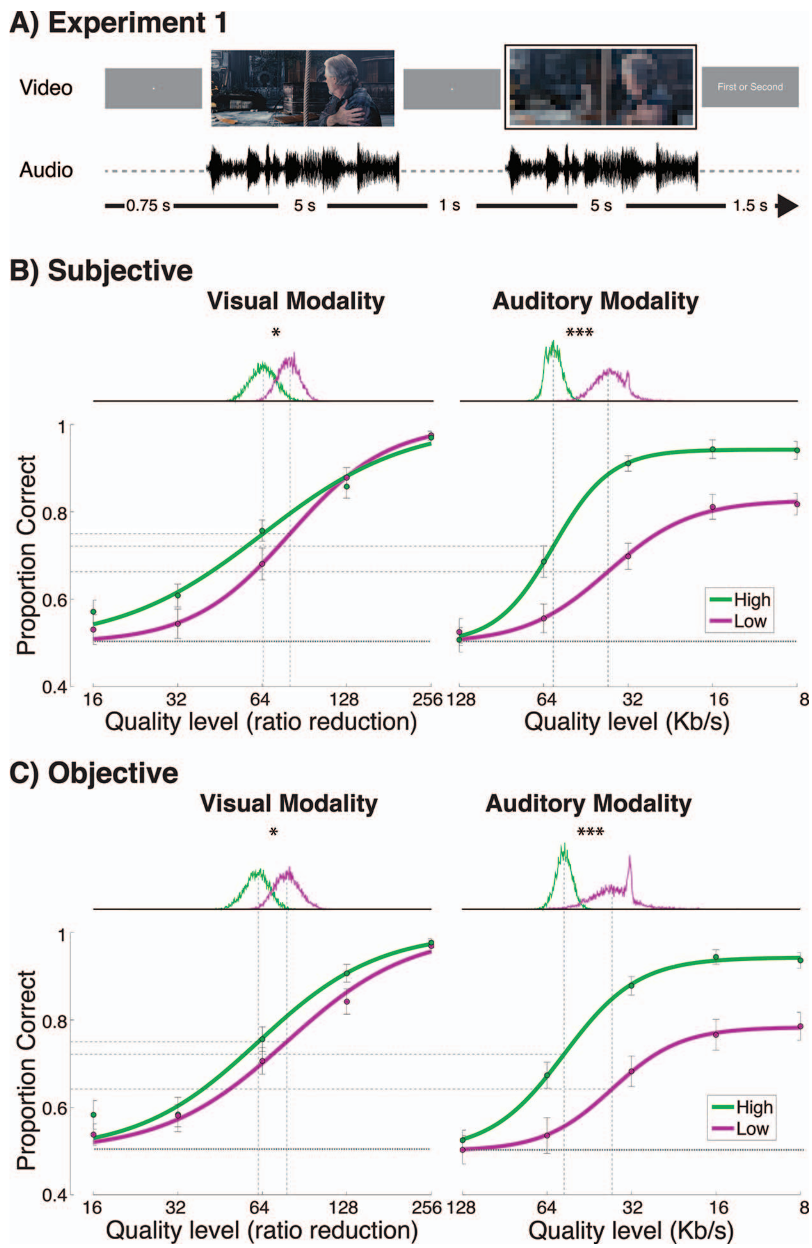

Figure 1. Methods and results for Experiment 1. (A) Example trial. Psychometric functions for high and low subjective (B) and objective (C) complexity in visual and auditory modalities. Video lossy compression steps are the proportion file size reduction (smaller values mean less compression and greater sensitivity); audio lossy compression steps are kilobits per second (larger values mean less compression and greater sensitivity). Above the plots are the resampled distributions of thresholds, significance of difference: * $p < .050$, *** $p < .001$. Still frames in this figure are permitted for reuse under CC 3.0. See the online article for the color version of this figure.

We began this experiment uncertain whether complexity would overwhelm perception and hurt performance or allow for richer representations that help performance. Our results from both psychometric and regression analyses supported the latter possibility. The results between subjective and objective measures of complexity were also consistent, especially for the psychometric analysis, despite the definitions of complexity being uncorrelated in the visual modality. This suggests the existence of two forms of complexity that impact perceptual sensitivity.

Experiment 2

Experiment 1 was an exploratory study, in the sense that we anticipated that either direction of effect was possible. Therefore, in Experiment 2 we sought to replicate these findings and at the same time test a potential explanation based on attention. Specifically, more complex video and audio streams may have attracted attention, which in turn has been shown to affect perceptual sensitivity (Pestilli & Carrasco, 2005; Störmer, Winther, Li, &

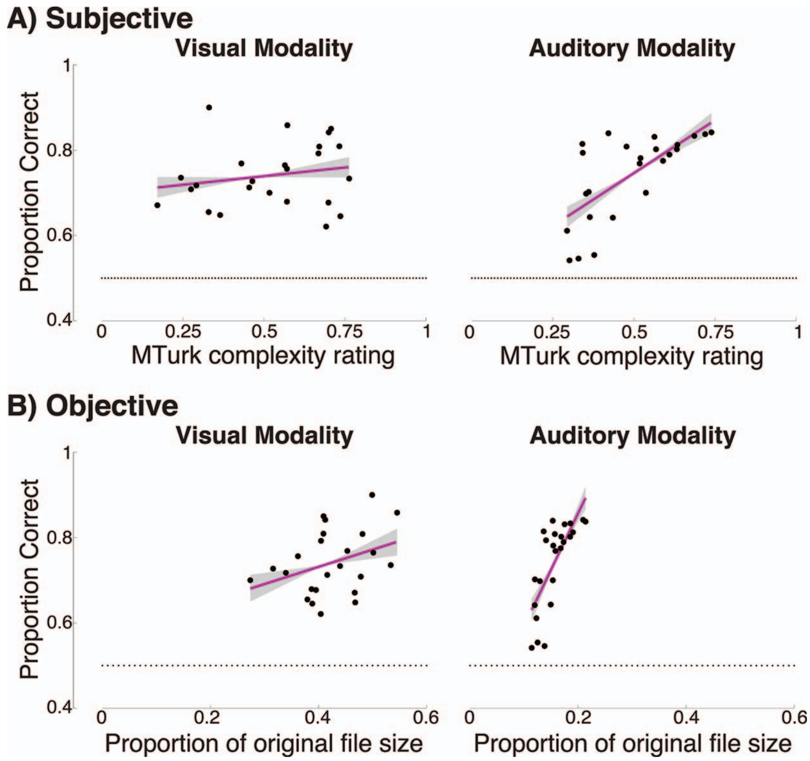

Figure 2. Regression results for Experiment 1. Relationship of (A) subjective complexity and (B) objective complexity with proportion correct in detecting visual and auditory changes. Shaded area represents the 95% confidence interval of the slope values from bootstrap resampling. The purple line is the slope without resampling. See the online article for the color version of this figure.

Andersen, 2013). Here we explicitly manipulate attention to the video or audio stream using a cuing paradigm (Posner, Nissen, & Ogden, 1978). As a sanity check, we expect improved detection of global changes that are validly versus invalidly cued. However, the key prediction of the attention account involves an interaction between complexity and cuing: If complex clips inherently attract or hold attention, we would expect a diminished impact of the cue. We tested this prediction by calculating the relationship across clips between complexity with the validity effect. Moreover, insofar as the positive relationship between complexity and sensitivity from Experiment 1 depends on attention, it should be diminished when analysis is restricted to invalid trials. In other words, does the relationship between complexity and perceptual sensitivity hold when attention is directed elsewhere and thus there is reduced attention to the stimulus.

Table 1
Psychometric Lossy Compression Thresholds

Definition	Modality	Complexity	*M*	Low CI	High CI
Subjective	Video	High	64.7	49.6	74.1
		Low	79.4	66.2	94.4
		Diff	−14.7	−27.6	−1.9
	Audio	High	59.4	51.9	66.8
		Low	37.6	28.9	48.2
		Diff	21.7	10.4	31.3
Objective	Video	High	61.4	49.6	74.1
		Low	78.6	64.0	97.2
		Diff	−17.2	−34.3	−1.9
	Audio	High	55.8	49.1	63.7
		Low	37.4	25.1	52.7
		Diff	18.4	2.1	30.2

Note. Video threshold units are the proportion file size reduction (smaller values mean less compression and greater sensitivity); audio threshold units are kilobits per second (larger values mean less compression and greater sensitivity). Thresholds are calculated by averaging participants and clips. Confidence intervals are bootstrapped distributions.

Method

Participants. The compensation and consent for the 24 participants (13 females, age 18–21 [$M = 19.12$, $SD = 1.12$]) in this experiment were the same as Experiment 1. We chose this sample size to maintain consistency across experiments. Nevertheless, simulations of the data from Experiment 1 indicated that this sample size would have yielded a positive slope at least 80% of the time (range: 97.4–100%).

Procedure. The room, computer, and movie clips were the same as Experiment 1. Before the two clips appeared on each trial, the word *Auditory* or *Visual* was presented for 750 ms to indicate with 75% probability whether the audio or video stream contained the change in quality between clips (Figure 3A). Participants were told that these cues would be accurate “most of the time.” The degree of audio or video quality change for each clip was determined using the 75% accuracy threshold

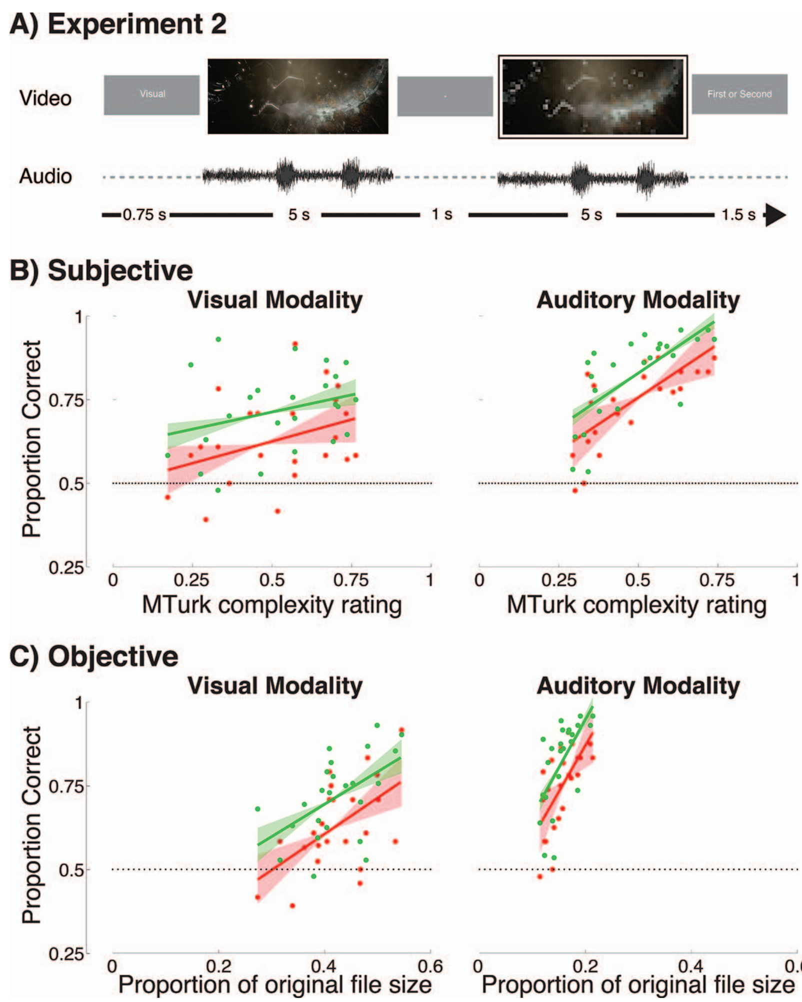

Figure 3. Methods and results for Experiment 2. (A) Example trial. Relationship of (B) subjective complexity and (C) objective complexity with proportion correct in detecting visual and auditory changes. Shaded area represents the 95% confidence interval of the slope values from bootstrap resampling. Solid lines are the slopes without resampling. Green refers to valid trials when the cued modality contained the change (75%); red refers to invalid trials (25%). Still frames in this figure are permitted for reuse under CC 3.0. See the online article for the color version of this figure.

identified in Experiment 1 (average 71:1 reduction for video, 40 Kb/s for audio). In other words, whereas we used multiple quality change levels in Experiment 1, here we picked one sensitivity level and looked at the impact of the attentional manipulation on accuracy at that level. Each of the 24 clips was shown eight times: four with video change, four with audio change, and within each modality three valid trials and one invalid trial (192 total trials). Whether the lower quality clip appeared first or second was randomized. Akin to Experiment 1 participants completed 12 practice trials before the experiment began, also with 75% cue validity.

For this and subsequent experiments, we made use of an eye-tracker (Tobii X120, sampling rate = 60Hz) available in the testing room to monitor participant alertness and compliance. We did not have any hypotheses about eye movements related to the task design, especially because the cuing task was sequential, and thus we do not report any data or analyses.

Results and Discussion

Overall participants were correct 70.2% of the time when there was a visual change and 80.3% when there was an audio change. Hence, the amount of quality change based on the 75% threshold from Experiment 1 was approximately correct.

To attempt to replicate Experiment 1, we first calculated the correlation across clips between complexity and sensitivity without consideration to cuing, and used the same resampling procedure to assess reliability. The relationship between complexity and sensitivity was positive for subjective complexity in both modalities (video: $b = 0.21$, 95% CI [0.15, 0.44], $p = .001$; audio: $b = 0.65$, 95% CI [0.60, 0.78], $p < .001$) and objective complexity in both modalities (video: $b = 1.02$, 95% CI [0.42, 0.67], $p < .001$; audio: $b = 2.84$, 95% CI [0.51, 0.71], $p < .001$). Hence, as in Experiment 1, both subjective and objective complexity seem to enable improved performance in perceptual tasks.

As verification that the attention cue was effective, accuracy was higher on valid versus invalid trials, for quality changes in both modalities (video: 71.9 versus 63.3%, $t[23] = 2.86$, $p = .009$; audio: 82.1 versus 74.7%, $t[23] = 3.17$, $p = .004$).

Insofar as complexity attracts attention, this validity effect may be attenuated for more complex clips because attention is drawn to them regardless of the cue. We calculated the validity effect for each clip as the average across participants of the difference between valid and invalid trials containing that clip. We then predicted this effect across clips with the subjective and objective complexity measures. To assess the reliability of the slopes, we resampled participants used in calculating the per-clip validity effects. There was no relationship across clips between the validity effect and subjective complexity (video: $b = -0.06$, 95% CI [−0.36, 0.27], $p = .860$; audio: $b = 0.02$, 95% CI [−0.34, 0.42], $p = .974$) or objective complexity (video: $b = -0.06$, 95% CI [−0.80, 0.57], $p = .790$; audio: $b = 0.05$, 95% CI [−1.55, 2.06], $p = .920$). These null effects fail to support an attentional explanation of the relationship between complexity and perceptual sensitivity, especially in light of the robust overall validity effect.

Indeed, the relationship between clip complexity and accuracy was immune to the attention cue. For valid trials (Figure 3B and 3C, green), there was a positive relationship for both subjective (video: $b = 0.33$, 95% CI [0.09,0.36], $p < .001$; audio: $b = 0.70$, 95% CI [0.54, 0.76], $p < .001$) and objective complexity (video: $b = 0.52$, 95% CI [0.60, 1.35], $p < .001$; audio: $b = 0.63$, 95% CI [2.34, 3.42], $p < .001$). Critically, despite having only a third as much data, the same relationships were found on invalid trials (Figure 3B–C, red), for both subjective (video: $b = 0.36$, 95% CI [0.02, 0.50], $p = .034$; audio: $b = 0.73$, 95% CI [0.26, 0.99], $p = .002$) and objective complexity (video: $b = 0.55$, 95% CI [0.53, 1.62], $p = .001$; audio: $b = 0.64$, 95% CI [0.94, 4.45], $p = .005$). Thus, even when attention is diverted to another modality, changes in complex clips are more easily noticed than changes in simple clips.

This study replicated the positive relationship between complexity and perceptual sensitivity, but also called into question the role of attention in this relationship. Although attention to the video or audio stream of a clip improved discrimination of that stream relative to the uncued stream, the complexity of the clip did not affect the benefit of this cue. In fact, even when attention was directed to the other stream, the complexity of the uncued stream still affected sensitivity to changes in that stream. Although attentional resources are not completely diverted by the cue, these results suggest that goal-directed attention is not necessary for complexity to benefit perceptual sensitivity.

Experiment 3

One interpretation of Experiment 2 is that attention does not mediate the effect of complexity on perceptual sensitivity. In particular, the obtained relationship for the uncued stream suggests no role for goal-directed attention. However, this finding is also consistent with the possibility that complex stimuli capture attention in an involuntary, stimulus-driven manner, which in turn improves sensitivity. Here we test this possibility by removing the attention cue and embedding a visual or auditory oddball within movie clips varying in complexity. This oddball is superimposed on the audio or video stream, rather than interwoven into it. Insofar as complexity drives attention, the detection of these oddballs should be enhanced for more complex video and audio streams.

Method

Participants. A new cohort of 24 participants (15 females, age 18–24 [$M = 19.43$, $SD = 1.70$]) were recruited for this experiment. One participant was excluded because for failing to finish in time. Given that this was a new task design, we were unable to predict the effect size or estimate power. However, for the sake of comparison and consistency we chose the same sample size of 24 participants as Experiments 1 and 2.

Procedure. On each trial, a single clip was presented uncompressed (Figure 4A). An oddball appeared on half of the trials, consisting of an unexpected transient superimposed in one modality on a single movie frame (42 ms). The oddball could not appear within the first or last five frames of the clip but was otherwise timed randomly. In the visual modality, the transient consisted of a circular patch of scrambled pixels 8.3° in diameter (i.e., each pixel in the patch had its position randomized) on either the left or right side of the screen. In the auditory modality, a 5,000-Hz sine wave was played in either the left or right ear of the participant. After each clip, participants were asked to respond with a key press

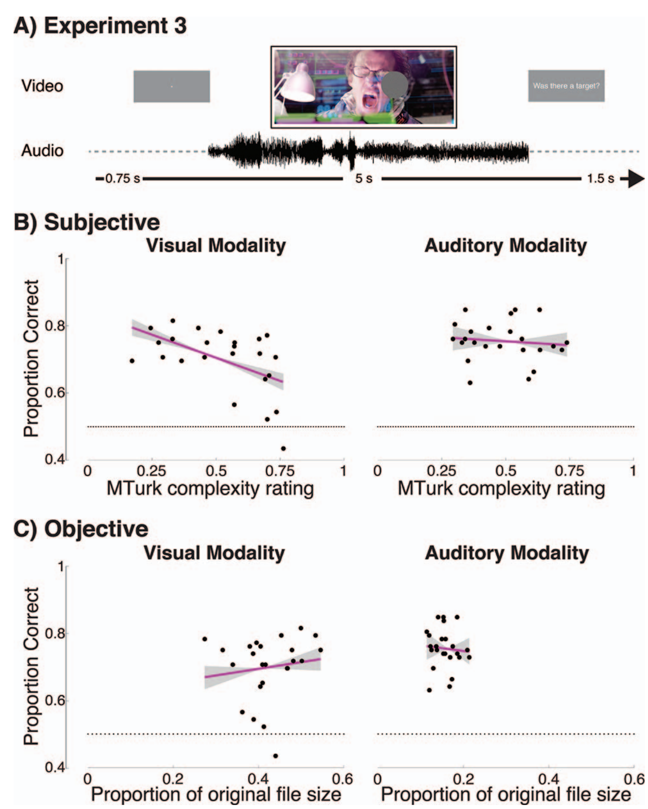

Figure 4. Methods and results for Experiment 3. (A) Example trial. The video stream has a transient on the last frame. Relationship of subjective (B) and objective (C) complexity to proportion correct in detecting the transient in visual and auditory modalities. Shaded lines represent the confidence interval of the slope from resampling. The purple line is the slope without resampling. Still frames in this figure are permitted for reuse under CC 3.0. See the online article for the color version of this figure.

whether they detected a transient in the audio or video stream, or they could omit their response to indicate that no transient was detected. The identity of the movie clip, the presence of a transient, and, if present, the modality and side of the transient were counterbalanced, resulting in 384 trials shown in random order. Akin to Experiments 1 and 2, participants completed four practice trials before the experiment began.

Results and Discussion

Participants were accurate on average in reporting the presence of a transient in the visual (69.9%) and auditory modalities (75.4%). On trials with a visual transient, they rarely reported an auditory transient (1.0%), and vice versa (1.4%). On trials with no transient, the false alarm rates were very low for both visual (1.6%) and auditory (1.9%) transients.

The critical analyses concerned whether transient detection accuracy improved with complexity. Inconsistent with this prediction, the relationship between subjective video complexity and visual sensitivity was negative ($b = -0.28$, 95% CI [−0.36, −0.19], $p < .001$; Figure 4B). Moreover, there was no reliable relationship for subjective audio complexity ($b = -0.05$, 95% CI [−0.20, 0.12], $p = .540$), objective video complexity ($b = 0.20$, 95% CI [−0.05, 0.47], $p = .120$), and objective audio complexity ($b = -0.17$, 95% CI [−0.92, 0.64], $p = .670$; Figure 4C).

Overall, we failed to extend the effect found with global change detection in Experiments 1 and 2 to the detection of local transients. In other words, the enhancement of perceptual sensitivity by complexity is selective to the complex stimulus itself rather than a general benefit for any stimuli appearing at the same location.

In Experiment 1 we raised the possibility that some perceptual changes that are alternatives to lossy compression, such as inserting a transient, would be too superficial to engage with the content of the clips and, in turn, would not affect performance. The present findings provide empirical support for this claim.

These results are inconsistent with the hypothesis that complexity captures attention in a stimulus-driven manner, and thus does not support an explanation of improved sensitivity mediated by attentional modulation. In conjunction with Experiment 2, this experiment suggests that complexity does not affect either stimulus-driven nor goal-directed attention. However, it is possible that complexity modulates attention in ways not examined here, such as in local versus global scope.

Experiment 4

We observed a benefit of complexity for perceptual sensitivity in Experiments 1 and 2 but not Experiment 3. One important difference between these studies was that the discrimination in Experiments 1 and 2 was between two clips presented sequentially, a task that engages short-term, working memory to make the comparison. The task in Experiment 3 placed no such demands. Thus, it is possible that working memory mediates the relationship between complexity and perceptual sensitivity. Here we test this possibility by returning to the design of Experiments 1 and 2 with a global distortion and two clips per trial, but present them simultaneously, allowing participants to make immediate perceptual comparisons between movie frames. Insofar as the benefit of complexity depends on working memory, there should no longer be a relationship between complexity and perceptual sensitivity.

Method

Participants. A new cohort of 24 participants (12 females, age 18–24 [$M = 9.62$, $SD = 1.44$]) were recruited for this experiment. Based on simulations of the data from Experiment 2, this sample size provided sufficient power to observe positive slopes for all analyses at least 80% of the time (range: 99.9–100%).

Procedure. Two clips were displayed simultaneously on either side of fixation and different audio was played to each ear (Figure 5A). To fit on the screen, the videos were reduced in size (13° width, 0.125° separation). To help compensate for the increased difficulty of this task from stimulus conflict and size, the amount of quality degradation for the clip that had to be detected was increased. Specifically, the upper end of the quality changes from Experiment 1 were used: factor of 64, 128, or 256 for the video and bit rate of 32, 16, or 8 Kb/s for the audio. For data analysis, we collapsed across the three levels of degradation, as it was not possible to fit psychometric curves to this small number of steps. Participants had to respond whether the left or right clip was lower in quality, and the location of this degraded clip was randomized. The fixation and intertrial interval were the same duration as Experiments 1–2. The identity of the movie clip, the modality reduced in quality, and the level of degradation were counterbalanced; each combination of parameters appeared twice, resulting in 288 trials. Participants completed eight practice trials before the experiment began.

Results and Discussion

Participants were correct 65.0% of the time when there was a visual change and 78.3% when there was an audio change. We replicated the positive relationships from Experiments 1 and 2 between subjective audio complexity and auditory sensitivity ($b = 0.34$, 95% CI [0.42, 0.65], $p < .001$; Figure 5B), objective video complexity and visual sensitivity ($b = 0.69$, 95% CI [0.36, 0.67], $p < .001$), and objective audio complexity and auditory sensitivity ($b = 1.52$, 95% CI [0.36, 0.61], $p < .001$; Figure 5C). However, there was no reliable relationship between subjective video complexity and visual sensitivity ($b = 0.03$, 95% CI [−0.12, 0.23], $p = .560$). It is possible that the relationship between subjective video complexity and perceptual sensitivity is weaker than the other relationships considered here. For instance, in Experiment 1 this relationship was only marginally significant and in Experiment 3 this relationship was significant in the opposite direction. Alternatively, it is possible that differences in the experiments drove the differences in subjective complexity. For instance, it is possible that subjective video complexity benefits perceptual sensitivity when those percepts are stored in working memory, which is required in Experiments 1 and 2 but not in Experiments 3 and 4. Future research should investigate the circumstances under which subjective complexity is and is not related to perceptual sensitivity.

Thus, the qualitative pattern of results was similar for the other relationships, but now without demands on working memory. This

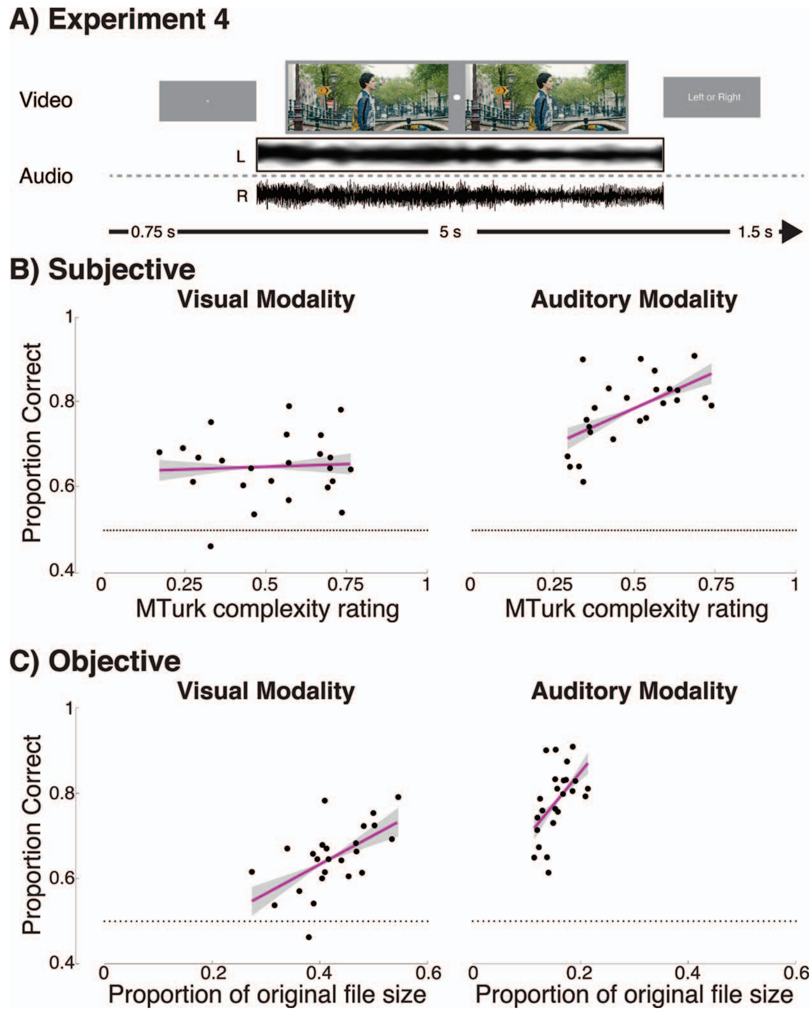

Figure 5. Methods and results for Experiment 4. (A) Example trial. Relationship between subjective (B) and objective (C) complexity and proportion correct visual sensitivity and auditory sensitivity. Shaded lines represent the confidence interval of slope values from resampling. The purple line is the slope without resampling. Still frames in this figure are permitted for reuse under CC 3.0. See the online article for the color version of this figure.

suggests that the benefit of complexity on perceptual sensitivity can manifest both when the information must be maintained over time (Experiments 1 and 2) as well as during the time of immediate perception, at least when the changes to be detected are integral to the complex stimulus (unlike Experiment 3).

Experiment 5

The fact that complexity benefits perceptual sensitivity during online processing but does not seem to depend on attention means that a different explanation is needed. One possibility is that complex clips result in richer perceptual representations, which in turn allow global distortions to be better detected. This could occur because complex clips contain more elements, which enable participants to evaluate the quality of the clip over multiple locations, time points, and feature dimensions, improving the likelihood of noticing a distortion and thereby increasing sensitivity. Here we sought corroborating evidence for this account by examining long-term memory.

Insofar as complexity enhances perceptual representations, a downstream consequence may be more detailed memories for complex clips, including as a result of the availability of additional cues at retrieval. Thus, we conducted a long-term memory task with individual clips that were encoded incidentally, followed after a delay by a surprise episodic memory test. We hypothesized that complex clips would be better remembered in the test.

Method

Participants. A new cohort of 49 participants (34 females, age 18–22 [M = 19.82, SD = 1.29]) was recruited for this experiment. Three participants, not included in this total, were excluded from the analysis because they failed to complete the experiment in time. As with Experiment 3 we could not accurately predict the effect size or estimate the power of this novel task design. Note that to estimate memory sensitivity for each clip, the clip had to be presented to some participants during encoding (so that it would be old at test and contribute to hits/misses) but not to other participants (so that it would be new at test and contribute to correct rejects/false alarms). To obtain these two sets of responses, we ran twice as many participants as Experiments 1–4.

Stimuli. Twenty-four additional 5-s clips were included in this experiment for a total of 48 clips. They were taken from the same film but at previously unused time points intermingled with the original set. The subjective complexity ratings for these new clips were collected using the same Mechanical Turk procedure as Experiment 1 ($N = 50$ for video, 16 females [two unspecified], age 19–35 [$M = 27.12$, $SD = 3.66$]; $N = 51$ for audio, 17 females [1 unspecified], age 19–35 [$M = 28.27$, $SD = 3.75$]). The ratings were as reliable as the first set: audio had poor interclass correlation (0.12) and video had fair (0.39). Like the first set of movie clips, the correlation between subjective and objective measures of audio complexity was strong and positive, $r(22) = 0.67$, $p = .007$, whereas subjective and objective video complexity were not reliably correlated, $r(22) = 0.27$, $p = .179$. The two sets of movie clips were counterbalanced across participants, with half of participants receiving the first set during encoding and as old items at test, and the second set serving as novel lures at test.

Procedure. As outlined in Figure 6A, participants completed three phases. In the Encoding phase, participants were shown 24 clips while their eyes were tracked. They were given no instruction about what to do while watching these other than to remain vigilant. In the Distractor phase, participants completed a distractor task for five minutes in which they summed two two-digit numbers at their own pace. In the Recognition phase, they were tested on their memory for the previously presented clips and 24 novel lure clips randomly intermixed. Because audio and video complexity vary with respect to each other and may impact memory differently, we tested the two modalities of each clip separately. Specifically, participants were presented with either just the video or just the audio from one of the 48 clips (96 total trials) and reported whether they had: high confidence it was previously presented (sure old), low confidence it was previously presented (unsure old), low confidence it was novel (unsure new), or high confidence it was novel (sure new). To assess episodic memory, "sure old" responses were coded as hits for previously presented clips and as false alarms for novel lures; all other responses were coded as misses and correct rejections, respectively (Kim et al., 2014; Turk-Browne, Yi, & Chun, 2006). The pattern of results was nearly identical when hits/false alarms were calculated for "old" responses irrespective of confidence. We quantified sensitivity with A', a nonparametric measure of sensitivity (Aaronson & Watts, 1987; Grier, 1971; Snodgrass & Corwin, 1988).

Results and Discussion

To verify that participants were able to remember the clips, we computed an overall A' for each participant across clips and found that it was reliably above chance (0.50) for both video ($M = 0.82$, $t(48) = 32.29$, $p < .001$) and audio ($M = 0.74$, $t(48) = 22.38$, $p < .001$) tests.

To examine the relationship between memory and complexity, we computed an A' for each clip across participants (half of participants contributed to the hit rate and the other half to the false alarm rate, based on counterbalancing). We then attempted to predict these values from subjective and objective measures of complexity, and resampled participants to obtain confidence intervals for the regression slope. Consistent with our hypothesis, complexity positively predicted subsequent memory for most conditions: subjective video ($b = 0.20$, 95% CI [0.10, 0.31], $p < .001$), subjective audio ($b = 0.66$, 95% CI [0.39, 0.93], $p < .001$; Figure 6B), and objective audio ($b = 2.52$, 95% CI [1.31, 3.70], $p < .001$; Figure 6C); however, there was an unexpected negative relationship for objective video ($b = -0.53$, 95% CI [-0.68, -0.40], $p < .001$).

This pattern of memory results is generally consistent with the perceptual benefits in Experiment 4, suggesting that they may be a downstream consequence. The exception was the objective definition of video complexity, which was positively related to perceptual sensitivity and negatively related to mnemonic sensitivity. This shows that complexity, at least defined in certain ways, can have dissociable effects on perceptual sensitivity and memory. For instance, it might be that the greater number of features in complex clips provides more opportunities for online detection of differences during perception, but that only a subset of the features can be encoded and/or more features can be encoded with lower precision, resulting in impaired memory. Indeed, research on pictorial memory suggests that delayed recognition is worse for images high versus low in complexity, at least when shown briefly during encoding (Fleming & Sheikhian, 1972). Regardless, in general we interpret these findings as evidence that complex inputs afford richer representations that can provide more opportunities for finding changes during online processing and memory retrieval.

General Discussion

The goal of this work was to investigate how complexity affects perception of naturalistic, multisensory stimuli. Using human judgment and information theory, we evaluated the impact of subjective and objective complexity on the perceptual sensitivity of movie clips. Across several experiments, perceptual sensitivity to global distortions was greater for more complex clips, regardless of whether they were presented sequentially (Experiment 1–2) or simultaneously (Experiment 4). Goal-directed attention enhanced sensitivity overall but not in a way that interacted with complexity (Experiment 2). We also did not find evidence that complexity captures attention in a stimulus-driven manner (Experiment 3). Instead, complex clips afforded better long-term memories, especially for auditory streams (Experiment 5).

Our tentative conclusion from these findings is that, rather than overwhelming perception, the additional elements and dynamics present in complex stimuli constrain and support perceptual analysis. As an analogy, consider evaluating whether the prescription of a pair of corrective eyeglasses is correct: looking at a rich and textured scene rather than a uniform surface provides more opportunities and variability over which to assess whether the scene appears correctly. It is unclear from the present work which features and dimensions in the complex clips provided this better scaffolding, though the failure to enhance detection of superimposed stimuli in Experiment 3 suggests that these features were integral to the stimulus.

We attempted to show that the findings reflect a general property of complexity by using two different definitions of complexity. However, the movie clips chosen resulted in a particular range

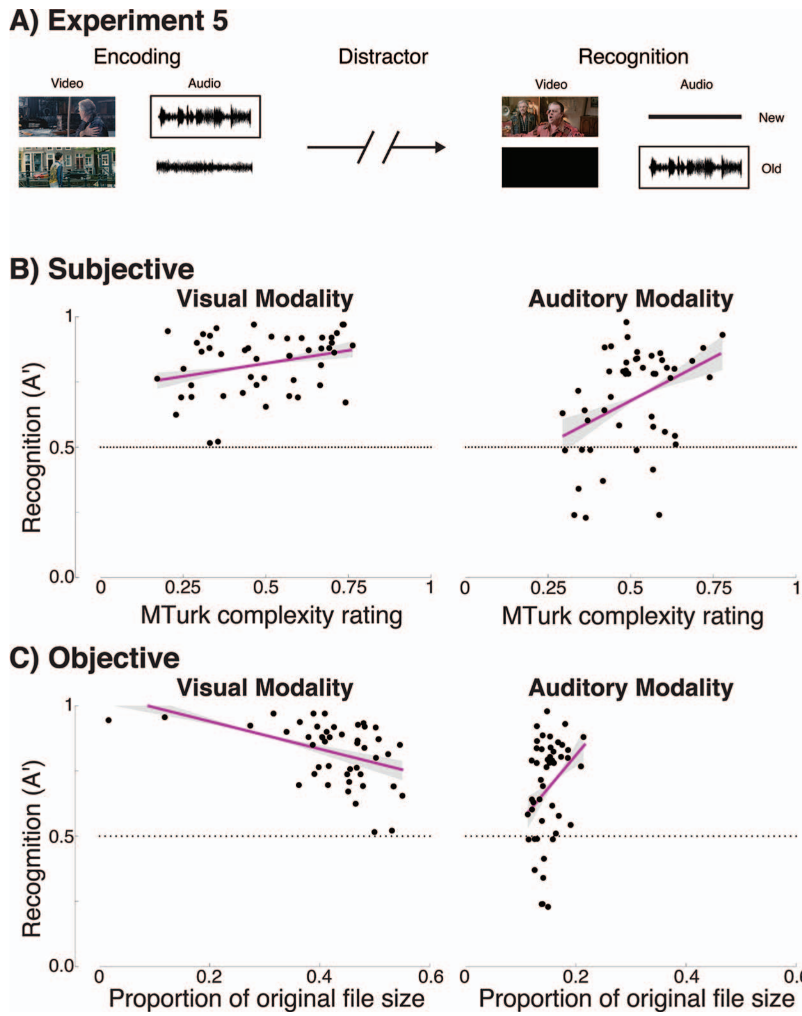

Figure 6. Methods and results for Experiment 5. (A) Schematic of the three phases. Participants incidentally encoded 24 clips, then did a distractor task, and finally had to judge whether they recognized the audio or video stream of a clip from earlier in the experiment. Relationship between subjective (B) and objective (C) complexity with A' for visual and auditory recognition memory. Shaded lines represent the confidence interval of slope values from resampling. The purple line is the slope without resampling. Still frames in this figure are permitted for reuse under CC 3.0. See the online article for the color version of this figure.

of complexities of both types, and it is possible that the findings may be different at higher or lower levels of complexity. Specifically, rather than the linear relationships observed here, the full relationship between complexity and perceptual sensitivity/memory may be nonmonotonic (Kidd et al., 2012). This range issue could also explain why certain types of complexity can be detrimental to perceptual sensitivity (Abdelhadi et al., 2011; Pylyshyn & Storm, 1988). Relatedly, artificial stimuli (e.g., white noise) can be outside the range of complexity of naturalistic images, and for such stimuli it would be difficult to notice subtle changes in quality. The impact of complexity range may also differ for auditory versus visual modalities, as well as for perceptual versus mnemonic tasks. Finally, having more elements in complex clips against which to identify changes may come at the cost of having lower precision representations of individual elements. This could be examined in future studies that involve discrimination of local rather than global changes during perception and that employ highly similar lures in the memory test, both of which require greater precision.

The extent to which the two complexity measures captured different aspects of complexity differed by modality. In the auditory modality, subjective and objective complexity were highly correlated. Although participants were unconstrained in how they made subjective judgments, we know that objective audio complexity reflected the temporal diversity of the signal, suggesting that these judgments were also based on such features. In the visual modality, subjective and objective complexity were uncorrelated or, if anything, weakly negatively correlated. We think this reflects a difference in the nature of the objective measure used for video versus audio complexity. In particular, whereas audio complexity tracked temporal diversity, video complexity was measured within each frame using lossless spatial compression and thus ignored temporal diversity across frames. Despite this difference, objective video complexity behaved like objective audio complexity, with both leading to enhanced perceptual sensitivity. They differed however in their relationship to long-term memory, which could be explored in future studies by contrasting temporal versus spatial redundancy in videos. Subjective video complexity showed the same relationship to long-term memory as objective and subjective audio complexity, suggesting that these subjective judgments may have been more sensitive to temporal redundancy.

The current study could have practical implications for compression codecs. Specifically, our findings suggest that consumers will be less likely to notice quality decrements from lossy compression when audio and/or video epochs with higher complexity are encoded with higher fidelity. Lossy compression algorithms could thus dynamically vary the amount of compression based on an information-theoretic measure of complexity/redundancy, to efficiently store multimedia content with minimal impact on user experience.

In conclusion, cognitive systems for perception and memory can operate more effectively in the presence of complex input. We provide evidence for a tentative explanation of this phenomenon, namely that it occurs because complex clips contain more elements against which to detect perceptual changes and that can serve as retrieval cues. This stands in contrast to the conventional assumption that sensory systems are overwhelmed with information and that this deluge impairs performance by exceeding perceptual and mnemonic capacities (Alvarez & Cavanagh, 2004; Lavie, 2005). At least for some types of perceptual judgments, complexity can support perceptual sensitivity. Indeed, given that naturalistic input is highly complex, the benefits reported here may be adaptive.

References

Aaronson, D., & Watts, B. (1987). Extensions of Grier's computational formulas for A′ and B′′ to below-chance performance. *Psychological Bulletin, 102,* 439–442. http://dx.doi.org/10.1037/0033-2909.102.3.439

Abdelhadi, S., Ibrahim, R., & Eviatar, Z. (2011). Perceptual load in the reading of Arabic: Effects of orthographic visual complexity on detection. *Writing Systems Research, 3,* 117–127. http://dx.doi.org/10.1093/wsr/wsr014

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15,* 106–111. http://dx.doi.org/10.1111/j.0963-7214.2004.01502006.x

Beck, D. M., & Lavie, N. (2005). Look here but ignore what you see: Effects of distractors at fixation. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 592–607. http://dx.doi.org/10.1037/0096-1523.31.3.592

Blender Foundation (Ton Roosendaal). (2012). *Tears of steel*. [Web video]. Available from https://mango.blender.org/

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General, 138,* 487–502. http://dx.doi.org/10.1037/a0016797

Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., . . . Massimini, M. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine, 5,* 198ra105. http://dx.doi.org/10.1126/scitranslmed.3006294

Chikhman, V., Bondarko, V., Danilova, M., Goluzina, A., & Shelepin, Y. (2012). Complexity of images: Experimental and computational estimates compared. *Perception, 41,* 631–647. http://dx.doi.org/10.1068/p6987

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6,* 284–290. http://dx.doi.org/10.1037/1040-3590.6.4.284

Cilibrasi, R., & Vitányi, P. M. (2005). Clustering by compression. *IEEE Transactions on Information Theory, 51,* 1523–1545. http://dx.doi.org/10.1109/TIT.2005.844059

Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science, 1,* 54–75. http://dx.doi.org/10.1214/ss/1177013815

Fan, J. E., Hutchinson, J. B., & Turk-Browne, N. B. (2016). When past is present: Substitutions of long-term memory for sensory evidence in perceptual judgments. *Journal of Vision, 16*(8), 1. http://dx.doi.org/10.1167/16.8.1

Fleming, M. L., & Sheikhian, M. (1972). Influence of pictorial attributes on recognition memory. *AV Communication Review, 20,* 423–441.

Forsythe, A. (2009). *Visual complexity: Is that all there is?* Paper presented at the International Conference on Engineering Psychology and Cognitive Ergonomics.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin, 75,* 424–429. http://dx.doi.org/10.1037/h0031246

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE, 7,* e36399. http://dx.doi.org/10.1371/journal.pone.0036399

Kim, G., Lewis-Peacock, J. A., Norman, K. A., & Turk-Browne, N. B. (2014). Pruning of memories by context-based prediction error. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 111,* 8997–9002.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception, 36,* 1–16.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 451–468. http://dx.doi.org/10.1037/0096-1523.21.3.451

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences, 9,* 75–82. http://dx.doi.org/10.1016/j.tics.2004.12.004

Lavie, N., & de Fockert, J. W. (2003). Contrasting effects of sensory limits and capacity limits in visual selective attention. *Perception & Psychophysics, 65,* 202–212. http://dx.doi.org/10.3758/BF03194795

Li, M., & Vitányi, P. M. (1997). *An introduction to Kolmogorov complexity and its applications* (Vol. 34., No. 10). Heidelberg, Germany: Springer.

Lindenmuth, J. E., Breu, C. S., & Malooley, J. A. (1980). Sensory overload. *The American Journal of Nursing, 80,* 1456–1458.

Malhotra, N. K. (1984). Information and sensory overload. Information and sensory overload in psychology and marketing. *Psychology and Marketing, 1,* 9–21. http://dx.doi.org/10.1002/mar.4220010304

Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15,* 197–237. http://dx.doi.org/10.1016/0010-0285(83)90009-9

Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition, 122,* 346–362. http://dx.doi.org/10.1016/j.cognition.2011.11.003

Murphy, G., Groeger, J. A., & Greene, C. M. (2016). Twenty years of load theory: Where are we now, and where should we go next? *Psychonomic Bulletin & Review, 23,* 1316–1340. http://dx.doi.org/10.3758/s13423-015-0982-5

Pestilli, F., & Carrasco, M. (2005). Attention enhances contrast sensitivity at cued and impairs it at uncued locations. *Vision Research, 45,* 1867–1875. http://dx.doi.org/10.1016/j.visres.2005.01.019

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. In H. L. Pick & I. J. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 137–157). Hillsdale, NJ: Erlbaum.

Potter, R. F., & Choi, J. (2006). The effects of auditory structural complexity on attitudes, attention, arousal, and memory. *Media Psychology, 8,* 395–419. http://dx.doi.org/10.1207/s1532785xmep0804_4

Prins, N., & Kingdom, F. A. A. (2018). Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the Palamedes toolbox. *Frontiers in Psychology, 9,* 1250. http://dx.doi.org/10.3389/fpsyg.2018.01250

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3,* 179–197. http://dx.doi.org/10.1163/156856888X00122

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18,* 849–860. http://dx.doi.org/10.1037/0096-1523.18.3.849

Scheydt, S., Müller Staub, M., Frauenfelder, F., Nielsen, G. H., Behrens, J., & Needham, I. (2017). Sensory overload: A concept analysis. *International Journal of Mental Health Nursing, 26,* 110–120. http://dx.doi.org/10.1111/inm.12303

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428. http://dx.doi.org/10.1037/0033-2909.86.2.420

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117,* 34–50. http://dx.doi.org/10.1037/0096-3445.117.1.34

Störmer, V. S., Winther, G. N., Li, S.-C., & Andersen, S. K. (2013). Sustained multifocal attentional enhancement of stimulus processing in early visual areas predicts tracking performance. *The Journal of Neuroscience, 33,* 5346–5351. http://dx.doi.org/10.1523/JNEUROSCI.4015-12.2013

Turk-Browne, N. B., Yi, D.-J., & Chun, M. M. (2006). Linking implicit and explicit memory: Common encoding factors and shared representations. *Neuron, 49,* 917–927. http://dx.doi.org/10.1016/j.neuron.2006.01.030

Yu, H., & Winkler, S. (2013, July). Image complexity and spatial information. In *2013 Fifth International Workshop on Quality of Multimedia Experience* (QoMEX; pp. 12–17). IEEE.

Received November 3, 2018
Revision received April 29, 2019
Accepted April 30, 2019 ■