

Universitat Pompeu Fabra

Perceptual oddities: assessing the impact of continuity editing on
perceptual conflict mechanisms

Alice Drew, Salvador Soto-Faraco



Background

Our ability to watch and understand films effortlessly is a remarkable and understated feat of human cognition. In films, events differing in time, space, goals, actions, characters or situations are spliced, forming sequences that are radically different from the continuous and gradual stream of events in everyday life (Cutting, 2005). It is hard to imagine what would happen if, in real life, the background of the visual scene we are looking at changed entirely every few seconds, the angle from which we see people and objects suddenly shifted by 180° without a motive, or the events that unfold before our eyes would instantly jump to a later time. Movies contain such transitions every few seconds (of course, depending on the movie and its style). For instance, according to the *cinemetrics* database (<http://www.cinemetrics.lv>) *North by Northwest* by Alfred Hitchcock has an average shot length of 6.1 s whereas *Sunset Boulevard* by Billy Wilder's average shot length is of 15.5 s (data submitted by Barry Salt).

One would think that such contrived changes at such a high rate would pose a tremendous challenge to our perception. Instead, during film viewing, we not only parse these transitions mostly seamlessly, but weave individual sequences into a coherent narrative structure and segment different parts of a story into its component episodes. This narrative coherence, in spite of unnatural sensory changes, is partly induced by continuity editing techniques (Bordwell, 2006) which consist in the process of ordering and transitioning between filmed sequences to engender a sense of flow in line with an overarching narrative. It is through the use of continuity editing techniques, that filmmakers create continuity or discontinuity around editing boundaries to be perceived as such by their viewers. However, for these techniques to work, perceptual processing must work toward filling in gaps and smoothing over gross breaches in sensory continuity. We believe that addressing these processes will not only help understand film viewing, but also reveal perceptual function in general. In fact, cinematic fragments are more 'naturalistic' than many laboratory stimuli and therefore, according to many authors, help making a step towards more ecologically valid situations (Hasson et al. 2008). Here, we advance the idea that the perceived continuity and segmentation across film editing boundaries is achieved through the interplay of two brain processes: perceptual inference and cognitive conflict. We will introduce these concepts as usually described in cognitive neuroscience, together with relevant findings in neurocinematics. We will then lay out a proposal of how perceptual inference and cognitive conflict may intervene in film viewing.

Perceptual inference

In the flow of everyday sensory experience, we are constantly inferring (often unconsciously) what will come next (Clark, 2013). For instance, if we watch a bird fly in the sky, we expect it to carry on flying, and adjust sensory predictions according to this expectation. In perception research, this process is commonly labelled predictive coding. Theories of predictive coding hold that the brain actively anticipates upcoming sensory input rather than passively register it (Summerfield & Egner, 2009). Drawing on prior experience as well as the current sensory context, our brain infers which inputs to expect based on internal models that are continually updated. Top-down predictions generated by the internal models are checked against bottom-up information acquired from the senses, and the disparity between the two generates prediction errors (mismatch responses) (Friston, 2005, Press et al. 2020). Prediction errors are large when internally generated predictions do not match incoming sensory evidence, in which case

processing resources may be directed to the part of the stimulus that has not been properly predicted (believed to be done via attention) (Press et al. 2020). Unexpected events have also been shown to produce increased neural responses. In the example of the bird above, we may devote some attention to infer the causes that could lead the bird to stop midway and update our internal model to account for the current state of the world. The loop between prediction, comparison, and adjustment of the model to minimise error leads to constant fine-tuning of the internal model (Press et al., 2020; Noppeney & Lee, 2018). Over time, the brain extracts statistical regularities in the sensory environment and can infer how likely observations are to reoccur. This loop encompasses our various senses and, according to several current theories, the subjective experience of perception emerges precisely from this process (Friston, 2012 or Clark, 2013).

Inferences about upcoming events are also constantly drawn as we watch films. Information within a given shot in a scene flows, for the most part, continually just as everyday life perception. However, transitions across adjacent shots involve sensory changes that can be very dramatic, as explained above. As mentioned before, a typical movie will contain many transitions between shots, with the average shot duration declining over the years, from 12 seconds in 1930 to about 2.5 seconds today according to Barry Salt (http://www.cinematics.lv/salt_on_cutting.php). Fortunately, within one scene, information in one shot can often be used to make predictions about the next shot via inference. If a shot of a bird flying off a tree is immediately followed by a shot of the same bird arriving at a nest, we, as viewers, infer that the bird flew to the nest, despite never seeing such abrupt transitions in real life. This ability to interpret cinematic editing improves with experience (Schwan & Ildirar, 2010). Unlike everyday life inferences, those we generate when watching films are in keeping with an overarching narrative that can supersede the expectations based on the commonly perceived behaviour of people and objects, such as for instance the laws of Newtonian physics (Smith & Casati, 2008). As such, viewers often end up bridging abrupt changes in visual input at the service of narrative, although such spans would be disconcerting, if not physically impossible, in real life. Given the prevalence of narrative therefore, one might expect changes in visual input to be more striking to viewers when these have a bearing on the narrative (i.e. when they have been carefully devised to draw the viewer's attention to a particular aspect of the scene relevant to the unfolding narrative) than when they have no bearing on it.

Event segmentation and continuity editing

Neurocinematics emerged as an interdisciplinary field of research drawing on filmmaking techniques and seeking to uncover the cognitive processes associated with viewers' experiences. One of these studies, by Magliano and Zacks (2011), addressed how continuity editing permits viewers to perceive event structures and boundaries, and which associated cognitive and neural mechanisms are at play. Magliano and Zacks asked participants to segment a film into meaningful events. The identified event boundaries were compared to the actual edits in the film which had previously been marked based on the three levels of continuity editing theory (Bordwell, 1985): continuity edits, spatial-temporal discontinuities or action discontinuities. Although all edits involve a large visual transient, they span different depths of the narrative structure, with action discontinuities involving the most profound narrative transitions (these edits are described below, under *Types of edits used for analysis*).

Behaviourally, action discontinuities had the biggest impact on event segmentation. Subjects identified boundaries marking the beginning of new meaningful events more often across action discontinuities than across spatial-temporal discontinuities or continuity edits. In terms of cognitive mechanisms, which were explored with fMRI, action changes reflected very different activations compared to spatial-temporal changes as well as continuity edits. Continuity edits and spatial-temporal edits showed responses to changes in stimulus input (increases in the activation of a network linked to early visual processing) and to changes of location (transient increases in activation in the parahippocampal cortex) respectively, as well as the latter showing patterns of activation linked to attention-driven downregulation across these moments of edits (decreases in early visual areas and in the lateral parietal cortex). Although they found little increase or decrease around action discontinuity edits and no involvement of high-level visual processing areas, the transient decreases seen in the other edit types (spatial-temporal and continuity) point to selective neural processing activated in order to bridge low-level feature discontinuities. Overall, Magliano & Zacks' results suggest that viewers may actively stifle sensory and spatial-temporal discontinuities at continuity edits in order to follow meaningful events (i.e. the overarching narrative of the film) that bridge breaks in low-level visual continuity as well as breaks in the continuity of space and time. We believe these findings are compatible with the role of perceptual inference during film watching and give us reason for probing the mechanisms involved at editing boundaries further. In particular, we believe that brain responses at editing boundaries may reflect the degree of error, or sensory input unaccounted for by perceptual predictions. In addition, we propose that there is a further mechanism at play once these discontinuities are detected: conflict monitoring and detection. These mechanisms are described hereafter.

Perceptual conflict

Cognitive control generally refers to a set of functions by which the brain monitors and flexibly adjusts to changes in the environment (Botvinick et al 2001). Incoming input is monitored in order to ensure coherent processing and control of action. If incompatible representations are simultaneously activated, then adjustments are triggered in order to aptly adapt and respond to them, minimising the consequences of the disruption in behaviour. The monitoring and detection of conflict is believed to be carried out by the anterior cingulate cortex (ACC) invoking areas in the dorso-lateral prefrontal cortex (dlPFC) for adjustments in control. Conflict detection has also been mirrored by increases in oscillatory power in the theta band (4-7Hz) in mid-frontal areas over the ACC (Botvinick et al 2001; Kerns et al 2004; Pastötter et al. 2013, Cohen and Ridderinkhof, 2013, Cavanagh and Frank, 2014). Most studies to date surrounding the phenomenon have focused on laboratory-specific experiments, inducing the cycle of conflict monitoring, detection and resolution by classic sensorimotor protocols such as the Stroop and Eriksen flanker tasks (cf. above references). However, a lesser explored phenomenon concerns conflict occurring purely between perceptual representations, although it has been stipulated that these conflicts may operate in a similar vein as the more studied, sensorimotor kind (cf. Botvinick et al. 2001). This phenomenon, perceptual conflict, occurs when two incompatible representations of a sensory input are

simultaneously activated. Previous experimental work has yielded promising results showing that conflict occurring between stimuli triggers a similar activation in the theta band as conflict occurring between stimulus and response, in perceptual situations such as the McGurk effect (Morís Fernández et al. 2017, 2018), binocular rivalry (Drew et al. *forthcoming*) and surrealist imagery (Ruzzoli et al. 2020). In the present study, we propose that perceptual conflicts may be experienced, and play a role, in more complex contexts such as in film viewing.

Continuity editing, perceptual inference and conflict

Cinema provides an interesting case study for the framework laid out above, as a prime example of perceptual incongruities, most of which happen without our noticing. In fact, the most widely known form of cinematic montage consists of creating the illusion of temporal continuity across a collection of visually different takes and scenes (Bordwell, 1985). How does the brain manage to process perceptual discontinuities that would be striking in real life and why are these not visually unbearable for the viewer? According to the view where perceptual inference and perceptual conflict are interconnected, the brain should respond differently to the different types of continuity edits. If this were to be the case, different brain responses may indicate different events necessary for narrative organisation. In this study, we aim to draw on continuity editing theory and event-segmentation research in order to address how continuity edits used in cinematography come to be perceived. We believe theories of perceptual inference when combined with the cognitive conflict framework may help explain these perceptual phenomena, which we see as a step towards a more ecologically viable representation of real-world contexts. Both frameworks have been very influential in explaining different aspects of human information processing. Within the conflict monitoring system, the very idea of perceptual conflict can be linked to a breach in expectation or simultaneous activation of incompatible stimuli. Within the framework of predictive coding, the concept of conflict can be linked to unmet predictions, and the ensuing error signal (see Noppeney and Lee 2018 for an example in speech and music).

In particular, the present proposal is that breaks in low-level visual continuity (transients produced at edit points) may produce prediction errors and trigger mechanisms of cognitive conflict, corresponding to the visually abrupt change in sensory input. We further suggest that this process may have different outcomes depending on the predictiveness of the transitions at the higher level of the narrative. For example, at continuity edits or discontinuities in space or time within a scene, transitions are well accounted for by the overarching narrative, and therefore more expected. Stifling the detected perceptual conflict at these points benefits our understanding of the film. Furthermore, film theory holds that a successful continuity edit should barely be perceived at all. Instead, action discontinuities are necessarily salient for the purpose of narrative, as they often mark relevant boundaries between scenes. Here, a deeper narrative transition often marks the opening of a new scene, where the new visual image must be reinterpreted from scratch. These transitions will be characterised by large prediction errors the adjustment of which is more costly, hence require a major involvement of the conflict system and by implication are more noticeable. Based on this, we predict that different types of continuity edits will be processed differently. Low-level visual or spatial-temporal discontinuities will be reflected by no or little theta increase in fronto-medial areas, since it benefits the viewer's understanding of the film for these conflicts to be stifled and not invoke the requirement for particular attentional

adjustments in control. For major discontinuities in action, it is less likely for the visual system to enable this suppression process based on context, since the discontinuity should be perceived as such in order to maintain the narrative flow.

Types of edits used for analysis

The experiment will consist of fragments from different film sequences (see below: *Stimuli*). Each edit boundary, or transition between two different camera shots, will be marked. These edits will be placed manually by following the method used by Magliano & Zacks (2009; 2010) and classified in three categories, according to film practice (Bordwell, 1985). The different categories of edits marked will include:

- *Continuity edits* are transitions between shots, but are continuous in space, time and action (in other words simply involve a shift in camera position).
- *Spatial/temporal discontinuity edits* are discontinuous in space or time but continuous in action.
- *Action discontinuity edits* are discontinuous in action as well as space or time.

These edits are hierarchically nested: for instance, action and spatial-temporal discontinuity edits necessarily contain a continuity edit. Action discontinuities also always contain spatial-temporal discontinuities. Most edits discontinuous in time are also discontinuous in space, hence our reason for grouping them. To ensure sufficient statistical power for EEG analysis on the theta band, we will include a minimum of 30 events for each edit type (cf. *Sample estimation*) (Morís Fernández et al. 2017, 2018; Ruzzoli et al. 2020).

Hypotheses

Our hypothesis is based on measurements of theta-band responses to conflicting information from EEG electrodes. This is based on the conflict monitoring and cognitive control theory, and evidence for its neural correlates as discussed above (Botvinick et al 2001; Kerns et al 2004; Pastötter et al. 2013, Cohen and Ridderinkhof, 2013, Cavanagh and Frank, 2014). ACC activity, and the ensuing pattern of Theta oscillations as measured from fronto-medial areas, includes at least two types of dissociable response: conflict detection, and conflict resolution. We expect these to be conflated in the signal of time-frequency EEG analysis (Coderre et al. 2011). However, all film edits outlined above involve a sharp and abrupt visual transition and therefore, should these visual transitions induce an increase in theta power marking the detection of conflict, this effect should be seen for all three edits. With regard to conflict resolution, we expect that any transition will generate an attempt at resolving the conflict it creates. However, this will happen to varying degrees depending on the type of edit, as per our hypothesis. For low-level discontinuities that are visually abrupt but have no major bearing on the unfolding of the narrative and are embedded within a scene (i.e. continuity edits or spatial-temporal discontinuity edits), conflict resolution will be successful and relatively fast, in order to stifle the conflict as soon as possible. For edits that signal an important shift in narrative (i.e. action discontinuity edits), we

expect that the theta power reflecting an attempt to resolve the conflict will be greater than for the other edits since the conflict will be impossible to resolve completely, since it signals an important change in narrative.

Therefore, we expect neural mechanisms of perceptual conflict to be selectively engaged to different degrees when processing edits to bridge low-level discontinuities at the service of narrative. This will be reflected as:

1. an increase in fm-theta power around or following moments corresponding to action discontinuity edits relative to around or following moments corresponding to spatial-temporal discontinuity edits and continuity edits (in comparison, low or insignificant)

Methods

Participants

We will test healthy participants aged between 19-34 with an even ratio of female and male participants (when possible). Participants are required to have 'normal' or corrected-to 'normal' vision, not wear contact lenses or glasses and not be under any medication.

Sample estimation:

We calculated multiple sample estimations using G*Power. Given our time and financial possibilities, and aiming for reasonable power at $> .80$ and an effect size of $d=0.45$, the N returned by G*Power was of 32. Additionally, using equivalence testing on R, we were able to determine the minimal effect size we could reliably detect with the sample that we could afford to collect. For a sample of 30 participants for instance, the minimal effect size of interest must be outside $[-0.53, 0.53]$ (with 80% power). Taken together, we decided to collect a sample of 30 participants.

Stimuli and procedure

The experiment will consist in watching seven excerpts from chosen films, with a total duration of 19.497 min (mean 2.78 min, SD= 0.64 min). These clips are taken from the following films:

The Good, the Bad, and the Ugly, by Sergio Leone (1966) (training clip)
La Grande Belleza, Paolo Sorrentino (2013) (two clips used in the experiment)
Le Ballon Rouge, Albert Lamorisse (1956)
Le Quattro Volte, Michelangelo Frammartino (2010)
Somewhere, Sofia Coppola (2010)
Laurence Anyways, Xavier Dolan (2012)
Bin-jip, Kim Ki-duk (2004)

These film fragments were selected on the basis of whether they contained continuity edits of the three types described above (cf. *Types of edits used for analysis*), being of a short length, and containing no dialogue, with the exception of *Le Ballon Rouge*, which contains twice a vocal interjection in French, with subtitles. This decision was made in

order to avoid the interference of potential linguistic processes and conflating sources of activity in our signal.

Following the viewing of each clip, participants will be asked to respond to two 'yes/no' questions to serve as a reality check in order to gauge whether participants were paying attention to the films. The threshold for determining that candidates were not paying sufficient attention and are to be discarded will be decided during piloting.

Clips will be displayed on a screen placed at approximately 90 cm from participants' heads, with the soundtrack audible through speakers.

An EEG configuration with ActiCap and 64 electrodes placed in accordance with the 10-10 international system will be used to record cerebral activity while participants perform the experiment.

Additionally, electrodermal activity (EDA) will be recorded for each participant in parallel to EEG during the clip viewing, using BIOPAC, placing two electrodes on one hand of participants (on the index and middle fingers). While we have no hypothesis for this, it is well known that moments of heightened attention correlate with increases in electrodermal activity. These data will be analysed in an exploratory analysis (see below).

Analysis

Spectral bands

We will focus on the theta band since it is an established marker of cognitive conflict at the frequency of 5-7Hz. A filter will be applied to theta oscillations using a Fast Fourier Transform.

Regions of interest (ROI)

There will be one main ROI corresponding to the theta band which will consist of electrodes at Fz, FCz and Cz: the mid-frontal area supposedly involved in the cycle of cognitive conflict (detection, monitoring and resolution) mainly governed by the Anterior Cingulate Cortex (see for instance Botvinick et al 2001; Kerns et al 2004; Pastötter et al. 2013). Whilst theta-band power will be calculated for all electrodes, the analysis shall only be conducted over these electrodes (Fz, FCz and Cz).

Pre-processing

Pre-processing and analyses of EEG data will be done using Fieldtrip (Oostenveld et al. 2011) and custom-made code in Matlab. Data will be segmented into trials surrounding the moments of interest for the analysis using a pre-stimulus window (prior to the edit) of 500 ms and three post-stimulus windows (following the edit) of the same length (in total, 1.5 s). These segments thus encompass a short time window before and after the edit as well as the moment of the edit itself. Given the complexity of stimuli and that our hypotheses revolve around narrative understanding (a high-level cognitive process), we expect any potential effect to take place over time following the moment of interest, rather than in the immediate temporal vicinity of the edit. Trials will be defined as this

time window (0.5 ms before edit and 1.5 second after edit). Three different trials will be marked corresponding to the three different types of edits. Data will be inspected for artefacts using Independent Component Analysis (ICA) in order to manually reject components due to blinks, movements or noise upon visual inspection. The subsequent power analysis will be performed on the data cleaned from the ICA components representing artefacts.

Time-window and time-resolved analysis

Following pre-processing of data, the power in theta-band oscillations will be analysed during the first time window of 500 ms (in order to include three cycles at the central frequency of 6 Hz) prior to the edit and in the time period following the edit (three time windows of 500 ms each).

An initial statistical analysis will be done by means of a paired t-test in order to compare theta power following action edits against other edit types. This comparison will be done relative to each other. One power value for the time-window occurring after the edit (still to be defined, either from 0.5 s to 1 s post-edit or 1 s to 1.5 s post-edit) following the window of interest will be derived. Importantly, this power value will be acquired after subtracting the power value from the 0.5 s time-window prior to the edit, which will serve as a baseline. The resulting (baseline-corrected) power (dB) for the theta band at the ROI for this window will be compared in all 3 continuity edit types detailed above by means of a one-tailed paired t-test, as our overarching hypothesis has a directional prediction (theta following action edits should be higher than following spatial-temporal or continuity edits). We shall look for whether the difference in power in these windows is itself significant.

A second time-resolved analysis will be performed to look for within-trial changes in the theta band. Given our directional hypothesis, a one-tailed paired t-test will be performed comparing theta power in a time window post-edit (still to be defined, either from 0.5 s to 1 s post-edit or 1 s to 1.5 s post-edit) with theta power in the 0.5 s time window prior to the edit. In particular, we expect theta (conflict marker) around action discontinuities to be lower before the edit (when visual disruption has not yet occurred) than after the edit (when the disruption has been detected and processed as a conflict) so we expect this power difference to be insignificant for visual or spatial-temporal discontinuities. This expected increase of theta power should mirror the involvement of the conflict detection and monitoring system, in order to bridge the lower-level discontinuities in order to maintain a grasp on the narrative.

Exploratory analysis – Alpha power in sensory areas

One of the consequences of conflict detection is the recruitment of resources to resolve the conflict, usually facilitated by attention orienting. Because Alpha oscillations are known to have an inhibitory role in sensory areas, attention is often negatively correlated with Alpha power. Here, we will investigate whether there is a relationship between edits leading to high-conflict, and the power of Alpha oscillations in sensory areas.

Exploratory analysis – electrodermal activity

We will conduct an exploratory analysis using the data from EDA recordings around moments of edits. We have no directional hypothesis for this data but in the future, would like to link conflict responses at the neural level to physiological responses in order to probe the relationship between cognitive conflict and emotional response.

As a first step towards this exploratory hypothesis, we will extract the event-related skin-conductance responses (SCRs) around our different events (i.e. marked edits) to gauge which events correlated most to SCRs. As a future approach, we will conduct a reverse analysis around all non-event-related SCRs, in order to probe which features of the movie triggered them.

Future avenues: Source localisation

In the future, and after an initial overview of the evolution of theta power based on the two previous analyses, and depending on these findings, we will consider performing source localisation using a DICS filter in order to locate the origin of theta power as distributed over the ACC, where conflict is processed.

State of the project

To date, data has been collected from 25 participants. We have performed ICA decomposition and rejection on 3 participants (who will be included in the final analysis) and have not analysed further the data of these or any other participants. We ran three pilots (3 subjects) whose data will not be used for analysis.

References

- Bordwell, D. (1985). *Narration in the fiction film*. Madison: University of Wisconsin Press.
- Bordwell, D., & Thompson, K. (2006). *Film art: An introduction*. New York: McGraw-Hill.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108, 624–652.
- Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences*, 8(12), 539-546. <http://dx.doi.org/10.1016/j.tics.2004.10.003>
- Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in cognitive sciences*, 18(8), 414–421. doi:10.1016/j.tics.2014.04.012
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and brain sciences*, 36(3), 181-204.
- Coderre, E., Conklin, K., & van Heuven, W. J. (2011). Electrophysiological measures of conflict detection and resolution in the Stroop task. *Brain research*, 1413, 51-59.
- Cohen MX, Ridderinkhof KR (2013) EEG Source Reconstruction Reveals Frontal-Parietal Dynamics of Spatial Conflict Processing. *PLoS One*.
- Cutting, J. E. (2005). Perceiving scenes in film and in the world. In J. D. Anderson & B. F. Anderson (Eds.), *Moving image theory: Ecological considerations* (pp. 9–27). Carbondale, IL: Southern Illinois University Press.

- Friston, K. (2005). A theory of cortical responses. *Philosophical transactions of the Royal Society B: Biological sciences*, 360(1456), 815-836.
- Friston, K. (2012). Prediction, perception and agency. *International Journal of Psychophysiology*, 83(2), 248-252.
- Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., & Heeger, D. J. (2008). Neurocinematics: The neuroscience of film. *Projections*, 2(1), 1-26.
- Hasson, U., Malach, R., & Heeger, D. J. (2010). Reliability of cortical activity during natural stimulation. *Trends in cognitive sciences*, 14(1), 40-48.
- Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior Cingulate Conflict Monitoring and Adjustments in Control. *Science*, 303(5660), 1023-1026. DOI: 10.1126/science.1089910
- Magliano, J. P., & Zacks, J. M. (2011). The impact of continuity editing in narrative film on event segmentation. *Cognitive Science*, 35(8), 1489-1517.
- Morís Fernández L, Macaluso E, Soto-Faraco S (2017) Audiovisual integration as conflict resolution: The conflict of the McGurk illusion. *Hum Brain Mapp*.
- Morís Fernández, L., Torralba, M., & Soto-Faraco, S. (2018). Theta oscillations reflect conflict processing in the perception of the McGurk illusion. *European Journal of Neuroscience*, 48(7), 2630-2641.
- Noppeney, U., & Lee, H. L. (2018). Causal inference and temporal predictions in audiovisual perception of speech and music. *Annals of the New York Academy of Sciences*, 1423(1), 102-116.
- Oostenveld, R., Fries, P., Maris, E., Schoffelen, JM (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, Volume 2011 (2011), Article ID 156869, doi:10.1155/2011/156869
- Pastötter, Bernhard & Dreisbach, Gesine & Bäuml, Karl-Heinz. (2013). Dynamic adjustments of cognitive control: Oscillatory correlates of the conflict adaptation effect. *Journal of cognitive neuroscience*. 25. 2167-2178. 10.1162/jocn_a_00474.
- Press, C., Kok, P., & Yon, D. (2020). Learning to perceive and perceiving to learn. *Trends in cognitive sciences*, 24(4), 260-261.
- Ruzzoli, M., McGuinness, A., Morís Fernández, L., & Soto-Faraco, S. (2020). From cognitive control to visual incongruity: Conflict detection in surrealist images. *Plos one*, 15(6), e0224053.
- Schwan, S., & Ildirar, S. (2010). Watching film for the first time: How adult viewers interpret perceptual discontinuities in film. *Psychological Science*, 21, 1-7.
- Barry Smith & Roberto Casati (1994) Naive physics, *Philosophical Psychology*, 7:2, 227-247, DOI: [10.1080/09515089408573121](https://doi.org/10.1080/09515089408573121)
- Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic stimuli in neuroscience: critically acclaimed. *Trends in cognitive sciences*, 23(8), 699-714.

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in cognitive sciences, 13*(9), 403-409.)