

2018

COMMONALITIES AND DIFFERENCES IN VISUAL AND AUDITORY MULTISTABILITY

Taranu, Mihaela

<http://hdl.handle.net/10026.1/11983>

<http://dx.doi.org/10.24382/489>

University of Plymouth

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

COPYRIGHT STATEMENT

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the author's prior consent.

COMMONALITIES AND DIFFERENCES IN VISUAL AND AUDITORY MULTISTABILITY

By

Mihaela Taranu

Supervisors:

Prof. Sue Denham and Dr. Marina Wimmer

Thesis submitted to the University of Plymouth in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

School of Psychology | Faculty of Health and Human Science



January 2018

ABSTRACT

Perceptual bi/multi-stability—the phenomenon in which perceptual awareness switches between alternative interpretations of a stimulus—can be elicited by a large range of stimuli. The phenomenon is explored in vision, audition, touch, and even olfaction. The degree to which perceptual switching across visual and auditory bi/multi-stable paradigms depends on common or separate mechanisms remains unanswered. This main question was addressed in the current work by using four ambiguous tasks that give rise to bi/multi-stability and which are thought to involve rivalry at different levels of cognitive processing: auditory streaming and ambiguous-structure-from-motion (low-level tasks), and verbal transformations and ambiguous figures (high-level tasks). It was also investigated if individual differences in executive function (inhibitory control and set-shifting), creativity and personality traits have common relationships with perceptual switching in adults and children. A series of five experiments (four studies) were conducted.

In Study 1 (two experiments), perceptual switching behaviour of adult participants was examined in the four perceptual tasks mentioned above. In Experiment 1, participants reported higher switching rates for the ambiguous figure and verbal transformations than for ambiguous motion and auditory streaming. However, in Experiment 2 participants had a higher switching rate in verbal transformations than in auditory streaming, while the switching rates in the two visual tasks did not differ significantly. The correlations between visual and auditory switching rates were similarly inconclusive: in Experiment 1, no cross-modal correlations emerged, while in Experiment 2 there were correlations between ambiguous figure and verbal transformations and between ambiguous motion and verbal transformation. Furthermore, inhibitory control, set-shifting and creativity correlated with perceptual

switching rates in some of the perceptual tasks, although not in a consistent manner. In Study 2, the development of perceptual switching was investigated in children in the same four tasks used in Study 1. Findings showed that the number of switches increased with age in all four perceptual tasks, indicating general maturational developments. Executive functions and creativity were not associated with the ongoing perceptual switching, which was similar to what was found in adults. In Study 3, a neuroscientific perturbation approach was used to investigate whether the superior parietal cortex is causally involved in both visual and auditory multistability as a top-down mechanism. Transcranial magnetic stimulation on anterior and posterior superior parietal cortex did not increase or decrease the median phase durations in response to ambiguous motion and auditory streaming. These regions were not causally involved in either visual or auditory multistability. Perceptual switching across modalities correlated nevertheless, indicating common perceptual mechanisms. In Study 4, the effects of attentional control and instructions were further investigated in ambiguous motion and auditory streaming. There were strong correlations between perceptual switching in the two tasks, confirming that there are common mechanisms. However, the effects of voluntary attention did not explain the commonalities found. Possibly the commonalities found reflect similar functionalities at more low-level sensorial mechanisms.

In conclusion, perceptual switching in vision and audition share common mechanisms. These commonalities do not seem to be due to the same neural underpinning in parietal cortex. Moreover, attentional control does not explain the commonalities found, indicating a more low-level common mechanism or functionality. Perceptual switching across all ages is task specific, more than modality specific. No central influence of inhibitory control and creativity was constantly associated with perceptual switching regardless of task/modality, supporting the distributed mechanisms hypothesis.

TABLE OF CONTENTS

ABSTRACT.....	II
TABLE OF CONTENTS	IV
INDEX OF TABLES	VII
INDEX OF FIGURES	X
LIST OF ABBREVIATIONS	XII
ACKNOWLEDGEMENTS	XIII
AUTHOR'S DECLARATION	XV
1. GENERAL INTRODUCTION.....	1
1.1 RESEARCH ON VISUAL AND AUDITORY MULTISTABILITY	2
1.2 SEARCHING FOR COMMONALITIES: THE WHAT, WHERE AND HOW OF VISUAL AND AUDITORY MULTISTABILITY	5
1.3 FOCUS OF THE THESIS	6
1.4 COMMON INDIVIDUAL DIFFERENCES IN PERCEPTUAL SWITCHING ACROSS MODALITIES?.....	22
1.5 SUMMARY OF THE AIMS AND FINDINGS	38
1.6 STATISTICAL NOTE	40
2 STUDY 1: EFFECTS OF MODALITY AND LEVEL OF PROCESSING ON VISUAL AND AUDITORY MULTISTABILITY.....	48
1.1. INTRODUCTION EXPERIMENT 1	48
2.1 AIMS	48
2.2 METHOD	49
2.3 RESULTS	58
2.4 DISCUSSION EXPERIMENT 1	74
2.5 INTRODUCTION EXPERIMENT 2	78
2.6 METHOD	78
2.7 RESULTS	80
2.8 DISCUSSION EXPERIMENT 2	90
2.9 GENERAL CONCLUSION STUDY 1	92

3	STUDY 2: VISUAL AND AUDITORY MULTISTABILITY IN 6, 8, 10- YEAR OLD CHILDREN	96
3.1	INTRODUCTION	96
3.2	METHOD	98
3.3	RESULTS	99
3.4	DISCUSSION	113
4	STUDY 3: IS SUPERIOR PARIETAL CORTEX COMMONLY INVOLVED IN VISUAL AND AUDITORY MULTISTABILITY?	118
4.1	INTRODUCTION	118
4.2	METHOD	125
4.3	RESULTS	135
4.4	DISCUSSION	146
5	STUDY 4: DOMAIN GENERALITY/SPECIFICITY AND EFFECTS OF VOLUNTARY CONTROL ON VISUAL AND AUDITORY MULTISTABILITY	150
5.1	INTRODUCTION	150
5.2	METHODS	152
5.3	RESULTS	160
5.4	DISCUSSION	168
6	GENERAL DISCUSSION	172
6.1	CROSS MODAL COMMONALITIES	172
6.2	THEORETICAL IMPLICATIONS OF THE FINDINGS	176
6.3	PERCEPTUAL SWITCHING IN DEVELOPMENT	180
6.4	INDIVIDUAL DIFFERENCES IN EXECUTIVE FUNCTIONS, CREATIVITY AND PERSONALITY AND PERCEPTUAL SWITCHING.	182
6.5	CONCLUSION.....	186
7	SUMMARY CONTRIBUTIONS OF THIS THESIS.....	186
	APPENDICES.....	190
	APPENDIX A.....	190
	APPENDIX B.....	192
	APPENDIX C.....	194
	APPENDIX D.....	200

APPENDIX E 201

APPENDIX F 202

APPENDIX G..... 206

APPENDIX H..... 207

APPENDIX J..... 210

REFERENCES..... 211

INDEX OF TABLES

2. STUDY 1

Table 2.1. Means and Standard Errors for the Switching Rates in each Block	61
Table 2.2. The proportion of Time for each Percept (including None Presses), for Each of the Four Perceptual Tasks.	63
Table 2.3. Initial Reaction Times Performance across the Four Perceptual Tasks.....	63
Table 2.4. First Phase Durations Performance across the Four Perceptual Tasks.....	65
Table 2.5. Switching Rates Performance across the Four Perceptual Tasks	66
Table 2.6. The Overall Performance for Individual Differences Measures.....	68
Table 2.7. Commonalities between Perceptual Switching across the Four Perceptual Tasks.....	69
Table 2.8 Correlations for Perceptual Switching Variables (Initial Reaction Time, First Phase Duration, and Switching Rate) within Each Perceptual Task	70
Table 2.9. Perceptual Switching and Individual Differences in Ambiguous Figure	72
Table 2.10. Perceptual Switching and Individual Differences in Ambiguous Structure- from-motion.....	72
Table 2.11. Perceptual Switching and Individual Differences in Verbal Transformations	73
Table 2.12. Perceptual Switching and Individual Differences in Auditory Streaming ..	73
Table 2.13. Means and Standard Deviations for the Switching Rate in each Block	81
Table 2.14. Proportion of Time spent in Each Alternative Interpretation for each of the Four Perceptual Tasks	83
Table 2.15. Initial Reaction Times Performance across the Four Perceptual Tasks (N = 31).....	84
Table 2.16. First Phase Duration Performance across the Four Perceptual Tasks (N = 31).....	85
Table 2.17. Switching Rate Performance across the Four Perceptual Task (N = 31).....	86
Table 2.18. Correlations between the Initial Reaction Times across the Four Perceptual	89
Table 2.19. Correlations between the 1st Phase Durations across the Four Perceptual .	89

Table 2.20. Correlations between the Switching Rates across the Four Perceptual Tasks	90
---	----

3. STUDY 2

Table 3.1. The Proportion of Time Spent in Each Alternative Interpretation	100
Table 3.2. The Overall Performance on Stroop, Pattern Meanings Task, and Verbal Fluency	105
Table 3.3. Spearman Correlations between the Ability to Switch in the Ambiguous Tasks, Stroop, Pattern Meaning, and Verbal Fluency Tasks.....	108
Table 3.4. Spearman Partial Correlations between the Number of Switches in the Ambiguous Tasks, Stroop, Pattern Meaning, and Verbal Fluency Tasks, Controlling For the Effects of Age	109
Table 3.5. Spearman Correlations Between the Number of Switches in the Ambiguous Tasks, Stroop, Pattern Meaning, and Verbal Fluency Tasks.....	111
Table 3.6. Spearman Partial Correlations between The Number Of Switches In The Ambiguous Tasks, Stroop, Pattern Meaning, And Verbal Fluency Tasks, Controlling For The Effects Of Age.....	112

4. STUDY 3

Table 4.1. Summary of Four Previous TMS Studies that Investigated the Role of the Parietal Cortex in Perceptual Switching.....	123
Table 4.2. Summary Statistics for Each of the 12 conditions (2 Percepts).....	139
Table 4.3. Summary Statistics for Each of the 12 conditions (3 Percepts).....	139
Table 4.4. Spearman Correlations Matrix between the Median Phase Durations across the 12 conditions (2 interpretations).....	144
Table 4.5. Spearman Correlations Matrix between the Median Phase Durations across the 12 conditions (3 interpretations).....	145

5. STUDY 4

Table 5.1. Experimental Design Summary for Study 4	154
Table 5.2. Descriptive Statistics for the Number of Switches in Each of the Six Conditions in the Two-Alternative Response Group (N = 33).....	161
Table 5.3. Descriptive Statistics for the Number of Switches in Each of the Six Conditions in the Three-Alternative Response Group (N = 29).....	161
Table 5.4. Cross-modal Spearman Correlations between the Number of Switches across the Six Conditions (N = 33) in Two-Alternative Response Group	163
Table 5.5. Cross-modal Spearman Correlations between the Number of Switches across the Six Conditions (N = 29), in Three-Alternative Response Group	164
Table 5.6. Summary Statistics for Each of the Three Additional Measures, separately for the two groups tested	166
Table 5.7. Spearman Correlations between the Number of Switches across the Six Conditions and Individual Differences Measures for the Two-Alternative Response Group	167
Table 5.8. Spearman Correlations between the Number of Switches across the Six Conditions and Individual Differences Measures for the Three-Alternative Response Group.....	167

INDEX OF FIGURES

1. GENERAL INTRODUCTION

Figure 1.1. Duck-rabbit figure (Jastrow, 1899, p. 312).	1
Figure 1.2. Examples of the most known and investigated visual ambiguous paradigms.	3
Figure 1.3. A visual depiction of the main and secondary interpretations in auditory streaming paradigm.	4
Figure 1.4. A simplified illustration of the central vs. the distributed mechanisms in visual and auditory multistability. Adopted from Hupé et al., 2008.	10
Figure 1.5. Multistable paradigms used throughout the thesis.	18

2. STUDY 1

Figure 2.1. Training stages for each perceptual task.	53
Figure 2.2. Unambiguous picture of a girl (left side); Animation of a horse morphing into sheep (right side).	54
Figure 2.3. Images used in the Stroop task and the related memory control task (2.3.a). The corresponding colours for the buttons are in the second row (2.3.b).	56
Figure 2.4. An example of a pattern that participants were asked to describe in the Pattern meaning task.....	57
Figure 2.5. Example of a set of cards in the FIST task.....	58
Figure 2.6. Median first phase duration vs. the subsequent phase durations.....	60
Figure 2.7. Histograms of phase durations for each of the ambiguous stimuli.	62
Figure 2.8. Initial reaction times in each perceptual task.	64
Figure 2.9. First phase duration in each perceptual task.....	66
Figure 2.10. The switching rates in each perceptual task.	67
Figure 2.11. First phase durations versus subsequent phase durations.....	81
Figure 2.12. Histograms of phase durations for each task.....	82
Figure 2.13. Initial reaction times in each perceptual task.	85
Figure 2.14. First phase duration in each perceptual task.....	86
Figure 2.15. The switching rate in each perceptual task.....	87

3. STUDY 2

Figure 3.1. Average ability to switch (0 = no switch;1= switched) for each perceptual task across the three age groups.	101
Figure 3.2. The mean number of switches, for each task across the three blocks, separately for age group.	103

4. STUDY 3

Figure 4.1. Visualisation of the right (x = 38) anterior and posterior superior parietal lobe sites stimulated in Kanai et al. (2011).	121
Figure 4.2. Timeline for one experimental session. The TMS stimulation was administered between a pre- and a post-run.	126
Figure 4.3. The three interpretations for the ambiguous structure-from-motion (Left, Right and Combined).....	131
Figure 4.4. A visual illustration of the auditory streaming paradigm (from Farkas et al, 2016).....	133
Figure 4.5. Catch trial performance for the visual task (upper figure) and for the auditory task (lower figure).	136
Figure 4.6. The proportion of time spent in each interpretation.	137
Figure 4.7. Median change duration after TMS on each of the three brain regions stimulated (see agenda), only for the two main interpretations.....	138
Figure 4.8. Median change duration after TMS on each of the three brain regions stimulated (see agenda), for all three interpretations.....	140
Figure 4.9. Normalised median phase durations for each participant for the main two interpretations	141
Figure 4.10. Normalised median phase durations for each participant, for all three interpretations..	142

5. STUDY 4

Figure 5.1. Meta-analysis of the correlation coefficients from the previous studies that investigated perceptual switching across modalities.	151
Figure 5.2. The number of switches (y-axis) in the three voluntary conditions (on the x-axis).	163

LIST OF ABBREVIATIONS

AF = Ambiguous figures

AM = Ambiguous structure-from-motion (ambiguous motion)

VT = Verbal transformations

AS = Auditory streaming

fMRI = Functional Magnetic Stimulation Imaging

TMS = Transcranial Magnetic Stimulation

RTs = Reaction times

SD = Standard deviation

SEM = Standard error of the mean

N = Number of participants

M = Mean

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my PhD supervisors, Sue Denham and Marina Wimmer for giving me the chance to come to England to start this project under their supervision. It completely changed me, enriching both my research and personal skills and made me understand better the academic life and what it takes to be a good researcher. Sue, thank you for encouraging me to think interdisciplinarily and for igniting in me the wish to learn programming and always look for the bigger picture. Marina, thank you for teaching me the importance of being precise in analyses, language, speech, and writing. Thank you both for your immense patience and constant feedback. I have learned so much from you.

I would also like to thank István Winkler and Dávid Farkas for making my visit at the Institute of Cognitive Neurosciences and Psychology in Budapest a very intense learning period. István, I appreciate your dedication to science and how you support and care for your students to strive. Dávid, thank you for your constant help with statistics and with Matlab, and for showing me the beautiful Budapest. Thanks also to all the other lab members and especially to Júlia Simon who was such an amazing colleague while I was there.

Thank you, Raymond van Ee, for your constant friendly feedback and encouragements. Having you in this team was enriching and motivating. Thank you also for giving me the chance to visit Donders Institute and see how research is translated into business at Philips Research Laboratories.

Thank you, David Carmel, for welcoming me at Edinburgh University and for teaching me Transcranial Magnetic Stimulation. I will not forget how nervous I was when I had to mark the nasion and the tragus to make the right localisation on your skull. As the pointer was sharp, I guess you trusted me. Thank you also for taking good

care of me at the conference of the 20th ASSC in Buenos Aires. Your help was important for me and made my participation at this conference a memorable one. There I felt that consciousness research is truly amazing. Thank you, Cristel Devue, for being the postdoc I wanted so much to learn from and for an inspiring time spent in the lab in Edinburgh. You are such a good researcher and a great person.

I would like also to thank Prof. Rana Moyeed for answering my questions about mixed model analyses. Similarly, William Donalds (UC Davids), thank you for taking your time to check that my mixed model analyses approach was correct. Thanks also to the anonymous people online that share their knowledge constantly. Life in research is so much easier with you.

Thank you each and every Cognovian: Michael Sonne, Frank, Ilaria, Ivana, Vaibhav, Raluca, David, Katie, Diego, Michael Strauebig, Chun-Wei, Klara, Guy, Jack, Thomas, Jenny, Tara, Jacqui, François, Agi, Pinar, Joana, Christopher, Chris. I could write a long story about each of you and about how nice it was to have you around. I have learned a lot from all of you and I hope we are going to share our interdisciplinary experiences in the future as well. I hope our friendships will be long-lasting. Thank you Michael Sonne, Katie, and Vaibhav for proofreading this thesis. Special thanks to Michael Sonne. I met you in the perfect moment of my life. You are a truly inspiring person and your constant support was invaluable for me.

I am indebted to the examiners of my viva, Dr. Martin Doherty and Dr. Michaela Gummerum whose comments led to big improvements in this work after the viva.

Finally, thank you to the taxpayers of Europe and to the European Union who have contributed to my generous Marie Curie scholarship.

I dedicate this thesis to my mum who has given everything to my sisters and me.

AUTHOR'S DECLARATION

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award.

Work submitted for this research degree at the Plymouth University has not formed part of any other degree either at Plymouth University or at another establishment.

This study was financed with the aid of a Marie Curie Fellowship from the European Union (Marie Curie Initial Training Network FP7-PEOPLE-2013-ITN-604764) and carried out in collaboration with the University of Edinburgh and the Institute of Cognitive Neuroscience and Psychology (Hungarian Academy of Science).

Relevant scientific seminars and conferences were attended at which work was often presented; three external institutions were visited for training purposes and for data collection; two conference proceeding papers were published; two journal papers were submitted for publication. Some parts of this thesis (STUDY 4) has been published in Denham et al (2018) or will be submitted to publication (STUDY 2) in Taranu et al (in preparation).

External Collaborators:

Prof. István Winkler, Institute of Cognitive Neuroscience and Psychology, Research Centre of Natural Sciences, Hungarian Academy of Sciences, Hungary

Prof. Raymond van Ee, Radboud University, Donders Institute for Brain, Cognition and Behavior, Nijmegen, The Netherlands

Dr. David Carmel, University of Edinburgh, Department of Psychology, Edinburgh

Publications (journals):

Taranu, M., Wimmer C.M., Ross, J., Farkas, D., van Ee, R., Winkler, I., Denham, S., (in preparation). *Children's perception of visual and auditory ambiguity and its link to executive function and creativity*.

Denham, S. L., Farkas, D., van Ee, R., **Taranu, M.**, Kocsis, Z., Wimmer, M., Carmel, D., Winkler, I. (2018). Similar but separate systems underlie perceptual bistability in vision and audition. *Scientific Reports (Nature Publisher Group)*, 8, 1-10.

Taranu, M., Wimmer C.M., Farkas, D., van EE, R., Denham, S., Winkler, I. (in preparation). *Effects of modality and level of processing on visual and auditory bistability*.

Publications (conference proceedings):

Taranu, M., Loesche, F. (2017). *Spectres of ambiguity in divergent thinking and perceptual switching*. Avant, Vol. VIII, DOI: 10.26913/80s02017.0111.0012

Straeubig, M., Hsu, C.-W., Oztop, P., & **Taranu, M.** (2016). *(How) Does Play Matter? A Transdisciplinary Approach to Play and its Relation to Neurobiology, Creativity and Deception*. In M. Punt, S. Denham, & E. Doove (Eds.) (pp. 215–231). Off the Lip Conference - Transdisciplinary Approaches to Cognitive Innovation, Plymouth.

Conferences Attended:

Taranu, M. et al. (April 2017). *Executive functions and creativity effects on visual and auditory bistable perception in 6, 8, 10-year old children*. Society for Research in Child Development, Austin, Texas, USA. Poster

Taranu, M. et al. (June 2016). *Searching for the role of the parietal cortex in visual and auditory multi-stability*. 20th annual meeting of the Association for the Scientific Study of Consciousness. Buenos Aires, Argentina. Talk

Taranu, M. et al. (Sept 2015). *Effects of modality and processing level in visual and auditory perceptual bistability*. 19th conference of the European Society for Cognitive Psychology (ESCOP), Paphos, Cyprus. Poster

Taranu, M. et al. (July 2015). *Effects of modality and processing level in visual and auditory perceptual bistability*. School of Psychology Postgraduate Conference, Plymouth, UK. Talk

Taranu, M. et al. (Oct 2014) *Domain generality/specificity in the number of perceptual switches and phase durations in bistable visual and auditory stimuli*. Donders Discussion Conference, Donders Institute, Nijmegen, The Netherlands. Poster

Workshops:

CogNovo seminars/workshops 1-12 (April 2014-August 2017), University of Plymouth, UK.

Real World Eye Tracking Course (20-21 July 2016), Royal Holloway University of London, UK.

Mixed model analyses in R (4-6 January 2017), University of Plymouth, UK.

Word count: 46.511

Signed.....

Date.....

*We live in a world where almost everything we see can be construed in multiple ways.
As a result, we are constantly choosing between duck and rabbit. (Van Bavel, Nautilus)*

Why do we perceive things as we do? (Koffka, 1935, p.75)

1. GENERAL INTRODUCTION

Our visual and auditory perceptual experience of the world seems to be stable and unambiguous. However, in many situations the information received by our eyes and ears is ambiguous. For example, the image in Figure 1.1 will sometimes appear as representing a duck and sometimes a rabbit.

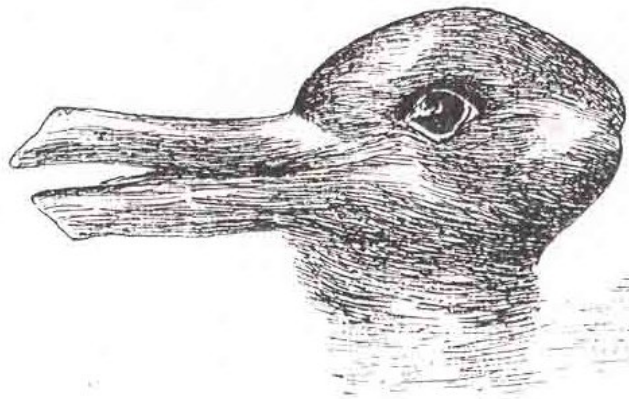


Figure 1.1. Duck-rabbit figure (Jastrow, 1900, p. 312) is an ambiguous figure, also known as reversible or bistable figure. After prolonged exposure to the figure, the perception switches between a duck and a rabbit.

The experience of sensory awareness switching between different interpretations of a stimulus in the absence of physical stimulus change is known as *multistable perception* (see Leopold & Logothetis, 1999; Schwartz, Grimault, Hupé, Moore, & Pressnitzer, 2012; Sterzer, Kleinschmidt, & Rees, 2009). This thesis presents a study of multistable perception with a special focus on identifying to what extent perceptual

switching in the visual and auditory domain is related or not. This chapter introduces the background, focus, research questions, and methodology of the thesis.

1.1 Research on Visual and Auditory Multistability

Many visual stimuli evoke multistability. Some of the most well-known and most often investigated stimuli are:

- The *Duck-rabbit figure* (Jastrow, 1971), presented previously (Figure 1.1), which is an image that can be perceived either as a duck or as a rabbit;
- The *Necker cube* (Necker, 1832), which can be experienced as a cube with two alternative orientations (Figure 1.2 (a));
- The *Rubin's vase-face* (E. Rubin, 1915), which after prolonged viewing leads to spontaneous switches between the perception of a vase or a face (Figure 1.2. (b));
- The *Ambiguous structure-from-motion* (H. Wallach & O'Connell, 1953), which can be interpreted as a sphere moving clockwise or anticlockwise (and sometimes moving both clockwise and anticlockwise) (Figure 1.2. (c));
- *Binocular rivalry* (Wheatstone, 1838), which arises when different images are presented to each eye with a mirror stereoscope. Any significant difference in orientation, colour, texture, movement, etc. will trigger the switch between different interpretations (Figure 1.2. (d)).

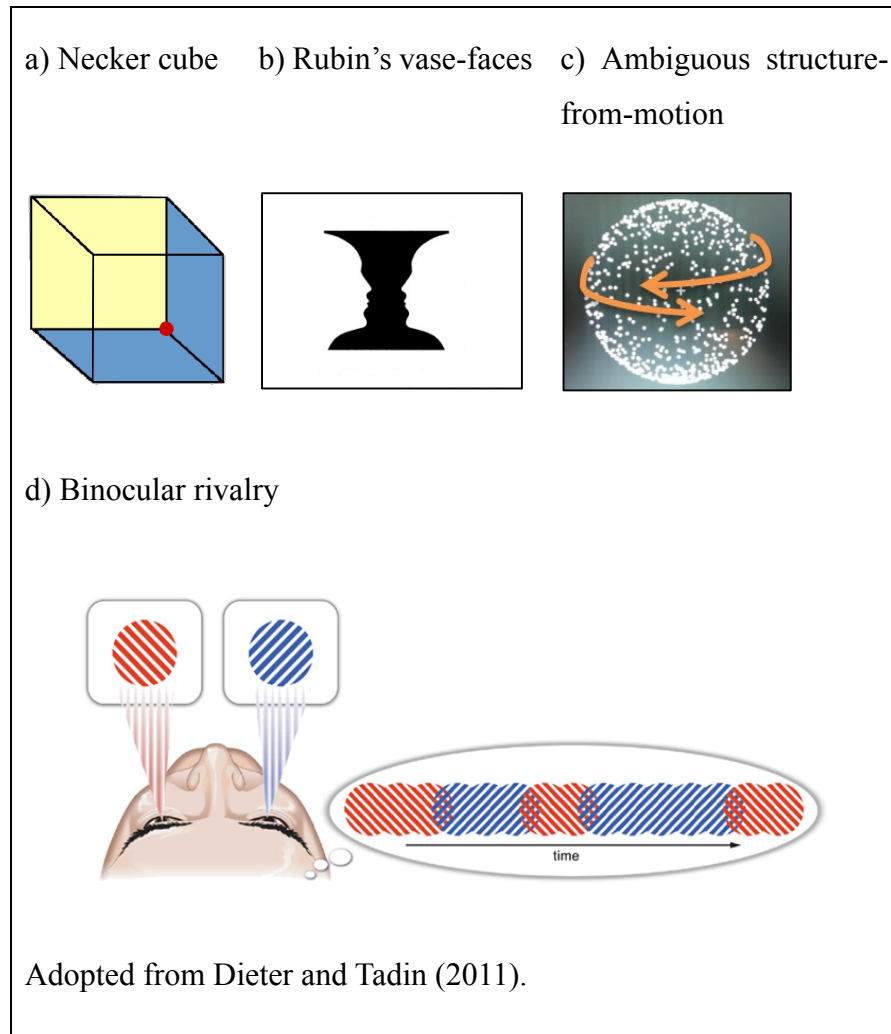


Figure 1.2. Examples of the most known and investigated visual ambiguous paradigms.

The study of auditory multistability has recently increased (Denham et al., 2012; Denham et al., 2014; Denham, Gyimesi, Stefanics, & Winkler, 2013; Denham & Winkler, 2006; Farkas et al., 2016; Hupé, Joffo, & Pressnitzer, 2008; Pressnitzer & Hupé, 2006). Investigations of auditory multistability have mostly used the paradigms of *auditory streaming* (van Noorden, 1975) and *verbal transformations* (Warren & Gregory, 1958).

The auditory streaming paradigm consists of a sequence of a repeating sound triplet of the form “LHL—LHL—”, where ‘L’ and ‘H’ denote two different sounds and

‘—’ is a silent interval with the same duration as the two sounds. The two most predominant interpretations are the *integrated* percept (grouping all sounds into a galloping sound sequence “LHL—”) and the *segregated* percept (grouping the ‘L’ and ‘H’ sounds into separate repeating “L_L_L” and “H__H__H” streams, one appearing in the background and the other in the foreground). Secondary interpretations are also possible, although they are not considered to appear with the same frequency as the main interpretations (Denham et al., 2012; Denham et al., 2014). Figure 1.3 offers a visualisation of the main and secondary interpretations.

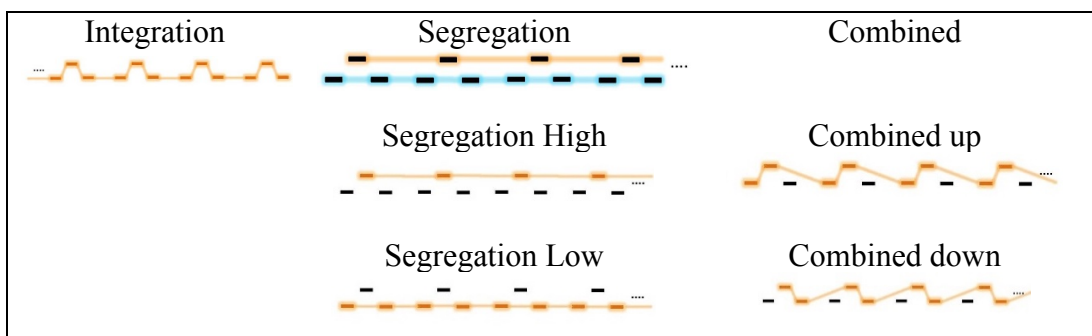


Figure 1.3. A visual depiction of the main and secondary interpretations in auditory streaming paradigm.

The verbal transformations effect (Kondo & Kashino, 2007; Warren & Gregory, 1958) is obtained when a word, repeated without a pause (e.g., “life”) switches into a different word (e.g., “fly”).

Due to the unique properties of these stimuli to give rise to different interpretations in the absence of stimulus changes (Leopold & Logothetis, 1999), they are used in research for many purposes. For instance, they have been utilised in research to identify: the determinants of perception and perceptual organisation (Alais & Blake, 2015; Pomerantz & Kubovy, 1981), the transition from sensation to perception (Blake & Logothetis, 2002), and to determine the neural mechanisms of awareness and consciousness (Crick & Koch, 1998; Dehaene & Changeux, 2011; Schwarzkopf & Rees, 2015). Moreover, their use has increased in research on autism (Van de Cruys, de-

Wit, Evers, Boets, & Wagemans, 2013), on bilingualism (Wimmer & Marx, 2014) and on emotions (Gray, Adams, & Garner, 2009; Heenan & Troje, 2015). Research on multistability in other modalities (e.g., audition, touch, or olfaction) has also increased (Carter, Konkle, Wang, Hayward, & Moore, 2008; Denham et al., 2012; Denham et al., 2014; Pressnitzer & Hupé, 2006; Schwartz et al., 2012; W. Zhou & Chen, 2009). Furthermore, attention has been given to study the extent to which perceptual switching in visual and auditory modalities share common mechanisms (see Hupé et al., 2008; Kondo et al., 2012; Pressnitzer & Hupé, 2006). However, as will be discussed further, the answer to this question is not unequivocal.

1.2 Searching For Commonalities: The What, Where And How Of Visual And Auditory Multistability

Over the last few years, it has been debated in the literature whether the processes underlying perceptual switching in the visual and auditory modality is domain specific or domain general (Hupé et al., 2008; Kondo et al., 2012; Pressnitzer & Hupé, 2006). The work presented in this thesis seeks to advance the current understanding of this question. To delimit the purpose of the thesis, distinctions must first be made between three predominant questions addressed by research on this topic: the HOW, WHAT and WHERE of visual and auditory multistability.

As suggested by Hupé et al. (2008) and Schwartz et al. (2012), a primary distinction is between HOW competition takes place in multistability and WHAT competes. HOW deals with the principles and mechanisms of functioning and with the means by which perceptual switching occurs. HOW is mainly focused on models that can explain the way in which multistability occurs. In this thesis, HOW is approached in respect to the central versus the distributed hypothesis as discussed by Hupé et al. (2008) (see description at page 9). WHAT, on the other hand, refers to what competes; the

characteristics of the stimuli or the type of representations that compete for awareness. For example, WHAT is addressed in studies investigating whether or not perceptual switching is dependent on the stimuli characteristics. One example is whether the semantic content of the stimuli influence perceptual switching across modalities (see page 14). The third important question is the WHERE of multistability, which focuses on identifying the neural basis the perceptual switching.

To fully understand the nature of the domain generality/specificity in multistable perception, this thesis seeks answers to all these three questions as will be detailed below.

1.3 Focus of The Thesis

The thesis comprises four studies that include five experiments. The HOW question is examined in all four studies with respect to the central versus distributed hypothesis detailed in Section 1.3.1, as first examined by Pressnitzer and Hupé (2006). The studies investigate to what degree perceptual switching in vision and audition is associated. The WHAT question is investigated in two studies: one using an adult population (Study 1) and one using a child population (Study 2). The aim is to determine, on the one hand, to what extent perceptual switching between stimuli at different levels of processing is similar and, on the other hand, whether perceptual switching across modalities relates only at similar levels of processing. In Study 2, both the HOW question and the WHAT question are investigated in a child population to understand how perceptual switching manifests during development. Study 3 tests the degree to which the same brain region is commonly involved in perceptual switching in response to both visual and auditory tasks. Study 3 presents a perturbation approach using Transcranial Magnetic Stimulation (TMS) to see whether the parietal cortex is causally involved in both visual and auditory multistability. Finally, Study 4 examines

to what degree asking participants to report two versus three interpretations has an impact on the number of perceptual switches, and whether voluntary control is a common influence on perceptual switching across modalities.

1.3.1 HOW: The Central versus the Distributed Hypothesis in Visual Multistability

Research on the commonalities between perceptual switching across modalities has its roots in visual research, where it is debated whether perceptual switching in various paradigms (ambiguous figures, binocular or monocular rivalry¹) reflects common or separate principles of functioning (Taddei-Ferretti et al., 2008; van Ee, 2005).

One school of thought hypothesises that perceptual switching in different visual paradigms reflects a common cause (Andrews & Purves, 1997; Carter & Pettigrew, 2003; Leopold & Logothetis, 1999). This viewpoint proposes the *central mechanism hypothesis* (or supra-modal) which suggests that perceptual switching in various paradigms has a common cause and neural basis. Evidence to support this hypothesis comes primarily from findings showing that the perceptual switching rate in different forms of visual paradigms is consistent at an individual level (i.e., correlates). For instance, an individual who reports slow/fast perceptual changes for one paradigm, is likely to report slow/fast perceptual switches in another paradigm and vice versa (Carter & Pettigrew, 2003; Shannon, Patrick, Jiang, Bernat, & He, 2011; Sheppard & Pettigrew, 2006). A second stream of evidence to support the central mechanism hypothesis explains why perceptual switching in various multistable paradigms manifests similar statistical characteristics. For example, the phase durations (i.e., the length of time

¹ Occurs when two different images are optically superimposed during prolonged viewing, one image becomes clearer than the other for a few moments then the other image becomes clearer.

during which one interpretation is sustained) in different visual paradigms can be statistically summarised according to a gamma or lognormal-shaped distribution (Leopold & Logothetis, 1999; Mamassian & Goutcher, 2005; van Ee, van Dam, & Brouwer, 2005). A lognormal distribution of the phase durations is a statistical description of the data characterised by many short phase durations and a long tail of few long phase durations; the majority of the durations are intermediate, producing a distribution with a marked right skew. Neuroscientists and computational modelers use these distribution characteristics to understand the intrinsic stochastic dynamics of brain activity (Leopold & Logothetis, 1999; Pastukhov et al., 2013; Y. H. Zhou, Gao, White, Merk, & Yao, 2004). Opposing views consider that this is weak evidence for a common causal link because many other unrelated phenomena seem to conform to a lognormal/gamma distribution (see distribution of rainfall over time; Sharma & Singh, 2010). However, in recent reports, the distributions of the phase duration across multistable paradigms showed task specific characteristics as they differed from the phase durations of other time-based tasks (Cao, Pastukhov, Mattia, & Braun, 2016). Further evidence for a common mechanism is assumed to be manifested in the inability to fully control the perceptual switching frequency in studies where participants are asked to exert voluntary control (i.e., by being instructed to maintain or switch faster between the alternative interpretations). The results show that this control is only possible to a certain degree, which can lead to the interpretation that a common mechanism/functionality is responsible for this inability for total control. Leopold and Logothetis (1999) suggested that this inevitability might suggest that our cognitive system needs to rely on persistently changing perceptions of a particular sensory pattern to interact successfully with the environment. However, these are only assumptions and no concrete evidence of how this is implemented has been shown so far. The central mechanism hypothesis is further supported by findings showing that brain structures

such as frontal or parietal cortex are necessary to initiate percept changes (Sterzer & Kleinschmidt, 2007), subsequently sending signals to sensory cortices (Cosmelli et al., 2004), acting as a common control mechanism.

As an alternative explanation to the common hypothesis, the *distributed mechanism hypothesis* suggests that perceptual switching occurs because of a distributed competition throughout many levels of processing in different brain regions, in a stimulus-specific manner. That is, perceptual switching across paradigms occurs because of a variety of causes and neural substrates and is not driven by a single central mechanism (Meng & Tong, 2004; Quinn & Arnold, 2010). This hypothesis is supported by findings showing no correlations between the perceptual switches in different visual paradigms (Gallagher & Arnold, 2014) or between multistability and binocular rivalry (Meng & Tong, 2004). Moreover, the role of attentional control in a range of paradigms is not clear. Although the ambiguous stimuli are amenable (to some degree) to attentional control, the precise level of control is specific to the stimulus used (Strüder & Stadler, 1999; van Ee et al., 2005). Similarly, neuroscientific studies exploring the neural correlates of perceptual switching in different paradigms show that no isolated cortical area selectively correlates with the participant's current percept, regardless of the paradigm (Moreno-Bote, Shpiro, Rinzel, & Rubin, 2010; Tong, Meng, & Blake, 2006). Rather, it seems that the object representation is a result of a complex interplay between early visual and several higher brain areas across the cortex, which is specific for each stimulus (Kornmeier, Hein, & Bach, 2009).

1.3.2 HOW: The Central versus the Distributed Hypothesis in Visual and Auditory Multistability

The debate between the central and the distributed hypothesis has extended into investigations of both visual and auditory multistability, which have been typically

studied separately (Hupé et al., 2008; Kondo et al., 2012; Pressnitzer & Hupé, 2006). Studying perceptual switching in both modalities has an advantage over studies focusing only on one of the modalities, as it can facilitate the understanding of the general principles and mechanisms of perceptual organisation, regardless of modality.

Hupé et al. (2008) summarised in Figure 1.4 the two possibilities regarding the central versus the distributed mechanisms hypothesis while exploring the extent to which visual and auditory multistability share commonalities. In the case of the distributed hypothesis, perceptual switching occurs independently across modalities (Figure 1.4, left side). According to this view, the representations for the competing interpretations are implemented separately within the auditory and visual pathways, possibly at various cortical and subcortical processing stages. The green arrow represents the contextual effects that are expected to occur but which should be independent of the switching mechanism. On the other side, in case of the central hypothesis (Figure 1.4, right side), the perceptual switches in the two modalities should reveal some interactions because of a common influence depicted with blue arrows (most likely arising from higher order brain regions).

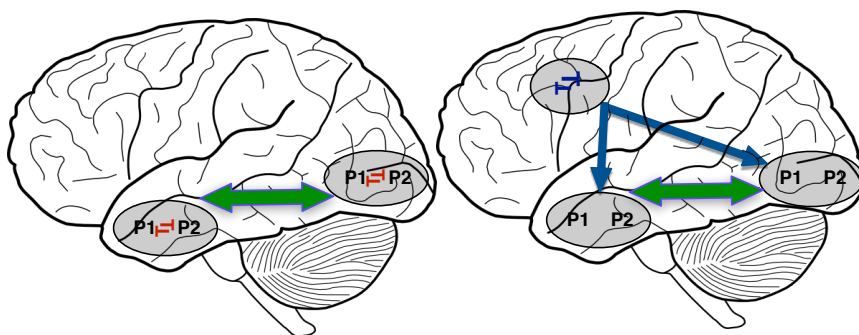


Figure 1.4. A simplified illustration of the central vs. the distributed mechanisms in visual and auditory multistability. The left image illustrates the brain mechanisms for the distributed hypothesis; the right image depicts the central hypothesis brain mechanisms. Adopted from Hupé et al., 2008.

Based on what was discussed above, the common hypothesis would be supported by a correlation between perceptual switching in the two modalities, while the distributed hypothesis would be supported by results showing no correlations. Similarly, common higher brain regions should be causally involved in both modalities to support the common hypothesis.

So far, three studies investigated the relationships between visual and auditory multistability and explored whether perceptual switching in the two modalities have common principles of functioning (Hupé et al., 2008; Kondo et al., 2012; Pressnitzer & Hupé, 2006). In the first study of Pressnitzer and Hupé (2006), participants had to report their perception in response to auditory streaming and to moving plaids². Each task was presented three times for four minutes; the first time in a neutral condition and subsequently in a grouped condition (instruction: *try to hear Integrated/see a single plaid*) or in a split condition (instruction: *try to hear Segregated/see two gratings*). In the neutral condition, the distributions of the phase durations in the two tasks were similar and followed a gamma or log-normal distribution (Leopold & Logothetis, 1999), suggesting commonalities between perceptual switching in the two tasks. In both tasks, volitional control did not increase the targeted phase durations but reduced the unwanted phase durations. This was formulated in the seminal work of Levelt (1967) and is considered a robust finding across different multistable paradigms (Klink, van Ee, & van Wezel, 2008). Participants were not able to hold onto a specific interpretation for the entire duration of the stimulus for any of the tasks. These similarities between perceptual switching across modalities were interpreted to show that perceptual switching shares common functional principles. However, participants who had firmer

² H. Wallach and O'Connell (1953) describe moving plaids as crossing lines seen through a circular aperture which could be perceived either as a single plaid moving vertically or as two gratings sliding horizontally across each other in opposite directions.

control over perceptual switching in the auditory task (i.e., switched faster or held onto one of the interpretations for longer time) could not necessarily do the same in the visual task, interpreted as showing domain specific mechanisms. Moreover, no correlation was found between the number of perceptual switches between modalities. That is, there was no individual bias to switch quickly or slowly across the two modalities, which supports the distributed mechanisms hypothesis. Overall, the fact that the distributions of phase durations are similar and that it is not possible to fully control perceptual switching in both tasks speaks in favour of domain general mechanisms. However, as no correlations were found between perceptual switching across tasks, it was argued that perceptual switching in response to ambiguous stimuli is caused by separate mechanisms in the two sensory modalities, although they can have similar functional principles.

In another study, Hupé et al. (2008) presented visual, auditory, and audiovisual tasks (visual and auditory stimuli presented simultaneously). In Experiment 1, participants were shown LHL- tones and visual plaids simultaneously (bimodal condition) or separately (unimodal condition) and asked to report their perceptions in each modality. In the second experiment, they presented LHL- tones and apparent motion patterns with a strong cross-modal coherence (spatial and temporal coincidence between the two modalities) to increase the likelihood of cross-modal interactions in perception. They hypothesised that if there is a central structure responsible for perceptual switching in both modalities, a degree of facilitation or interference is expected in the dynamics of the bimodal condition compared with the unimodal condition. Their results showed that there was no difference between the number of perceptual switches and the percentage of *grouped* interpretations between the unimodal and bimodal condition, indicating no cross-modal interaction between their overall dynamics. This result suggests that there is no supra-modal or central mechanism to

control the perceptual switching in both modalities and that perceptual switching in each modality is experienced independently of each other. In sum, the results from Pressnitzer and Hupé (2006) and Hupé et al. (2008) speak against the assumption of common mechanisms being responsible for perceptual switching across visual and auditory multistability and support the distributed hypothesis.

The third study of Kondo et al. (2012) contradicts in some regards the results found in the two previously described studies. The authors investigated the relationship between visual and auditory multistability using four different paradigms: auditory streaming, verbal transformations, visual plaids, and reversible figures. Each task was presented for five 90-second blocks. In contrast to the previous two studies it was found that the number of switches in the auditory and visual tasks correlated positively. By using exploratory and confirmatory factor analysis, they found that a three-factor model provided the best fit for the number of switches derived from the four paradigms investigated. The two auditory tasks shared a common factor, while the two visual tasks did not share a common factor. However, although these three factors appeared to account for distinct variables, they still correlated with each other, indicating commonalities between perceptual switching in the two modalities. The positive correlation between the numbers of switches across perceptual modalities and different paradigms contradicts the findings of Pressnitzer and Hupé (2006) and Hupé et al. (2008), suggesting the existence of a common functional mechanism.

Recently, Kondo and Kochiyama (2017) investigated perceptual switching in auditory streaming and visual plaids, using the same stimuli as Pressnitzer and Hupé (2006). The aim of the study was to see whether perceptual switching in the two tasks is different between adults in their twenties ($N = 11$), thirties ($N = 12$), forties ($N = 9$) and fifties ($N = 6$). It was found that the number of switches in both tasks decreased with

age. The study did not specifically investigate the correlations between perceptual switching between modalities. Kondo has subsequently stated (personal communication, July 4th, 2017) that there was a significant correlation between the number of switches in the auditory streaming and visual plaids ($r = 0.329$, $p = 0.044$, $N = 38$). However, when the effect of age was controlled with partial correlations, results did not reach statistical significance ($r = 0.234$, $p = 0.164$). These results support the findings of Pressnitzer and Hupé (2006) and contradict those of Kondo et al. (2012).

In conclusion, the correlational analyses in the three published studies are contradictory, which leaves the question of domain generality/specificity unanswered. This dispute will be addressed throughout this thesis alongside other aims. The correlational analyses should be clarified as it can make a difference between different types of models that could explain how multistability occurs: 1) models where switching happens because of a competition distributed throughout many levels of processing in a task-specific manner (Tong et al., 2006); 2) models where perceptual switching is the result of a supramodal mechanism (i.e., frontal cortex) that “weights” different inputs received by the sensory modalities in order to make a perceptual decision (Cosmelli et al., 2004); 3) models where switching happens because of competition distributed throughout many levels of processing but with a common neuronal architecture in charge of resolving conflicting information (Pressnitzer & Hupé, 2006).

1.3.3 WHAT in Visual and Auditory Multistability

One reason why Pressnitzer and Hupé (2006) did not find correlations between perceptual switching in vision and audition might be that the association is less robust when stimuli are abstract and have no semantic content.

There is substantial evidence to show that perceptual switching is influenced by the content of the stimuli. Early findings of Walker and Powell (1979) suggested that factors such as affective content and familiarity influence perceptual switching dynamics in visual ambiguous stimuli. In other studies, visual ambiguous stimuli with conflicting emotional or symbolic content were presented to different religious groups, and percept predominance (i.e., which interpretation was maintained more) was measured. For example, when perceiving the David star, Christian participants saw it more often as a cross, while Jewish participants saw it more often like a star (Losciuto & Hartley, 1963). Similarly, a figure of a person that was seen before tended to predominate over a figure never seen before in a binocular rivalry task (Goryo, 1960). Familiar and emotional content stimuli can involve a greater attentional load, which is known to modulate perceptual switching durations (Paffen, Alais, & Verstraten, 2006).

While investigating different kinds of visual ambiguous stimuli, Strüber and Stadler (1999) argued that the content of the stimuli has a major top-down influence on perceptual switching. Their study had three experiments where perceptual switching of 10 participants was tested in three conditions: *neutral* (participants were instructed to gaze at the figures in a passive manner and report the spontaneous switches, without trying to control the rate of the alternations), *speed* (participant were instructed to try to alternate as fast as possible) or *hold* (participants were instructed to try to keep the present interpretation for as long as possible). In Experiment 1, participants saw the duck-rabbit figure and the Necker cube for 3 minutes each. Results showed that the number of switches for the Necker cube was significantly larger than for the duck-rabbit in the *hold* condition and lower in the *speed* condition. There was no difference in the number of switches between the two tasks in the *neutral* condition. This result indicates that participants have greater control over the switching rate for ambiguous images with meaningful content (e.g., the duck-rabbit figure) as compared to images with more

abstract content (e.g., the Necker cube). In Experiment 2, participants saw Schroder's staircase figure and the chef-dog figure (Rock, 1956) and their perception was assessed as in Experiment 1. In the hold condition, the results had the same pattern as in Experiment 1; the number of switches for the staircase figure was higher than for the chef/dog figure. In the neutral condition, the number of switches was higher for the Schroder's staircase figure than for the chef/dog figure, while in the switch condition there was no difference in the number of switches between the two figures. In Experiment 3, participants saw the Maltese cross and the Rubin's vase-faces, and the results suggested similar trends as in Experiment 1, however these results were not statistically significant. Overall, the authors suggest that the content of the stimuli is an important factor for top-down influences, which is stronger for images such as duck-rabbit compared to more abstract stimuli such as the Necker cube. Overall, their findings indicate that the content of the stimuli has an impact on the number of switches, but that this impact depends on the type and characteristics of the stimuli and on the instructions (i.e., voluntary conditions).

In another study, Wolf and Hochstein (2011) manipulated the semantic content of the stimuli in a binocular rivalry paradigm. A word was presented to one of the eyes and a non-word to the other eye, and results showed longer phase durations for non-words than for words. They also found longer phase durations for line drawings of impossible figures than for possible objects. These results show that content-based stimuli have an impact on perceptual switching. Similarly, van Ee et al. (2005) reported that in binocular rivalry, participants had greater attentional control over the perception of stimuli representing complex objects, such as houses and faces, than over simpler stimuli, such as sinewave gratings. Furthermore, it was also found that in a neutral condition, the phase durations were larger for the face/house than for orthogonal

gratings. Overall, these findings suggest that the content of the stimuli has a major impact on perceptual switching in visual tasks.

In the auditory modality, the stimuli's content influences also the perceptual switching. Specifically, Kashino and Kondo (2012) described the auditory streaming phenomenon as a process requiring simple acoustic feature transformations (with no prior knowledge and experience dependence), distinct from the verbal transformations effect, which involves schema-based processes (with knowledge and prior experience dependence). They also found a larger number of switches in verbal transformations than in an auditory streaming task indicating that the content of the auditory tasks has an impact on perceptual switching. Thus, content-based stimuli are cognitively easier to access and might therefore give rise to an increased number of switches. One way to explore this possibility more systematically is to present both “high-level” stimuli that have semantic content in the visual and auditory domain and “low-level” stimuli without semantic meaning.

Studies 1 and 2 of this thesis compared perceptual switching in the auditory and visual modality as well as between low-level and high-level stimuli (Figure 1.5). Four multistable paradigms were used: two auditory (verbal transformations and auditory streaming) and two visual (the duck-rabbit figure and the ambiguous structure-from-motion cylinder). The duck-rabbit figure and verbal transformations are regarded as *high-level multistable phenomena* because the alternative perceptions involve different objects with semantic content. The ambiguous structure-from-motion and auditory streaming are considered to be *low-level multistable phenomena* because the alternative interpretations differ in object features as opposed to full objects with semantic content (see Figure 1.5 for the general classification of the perceptual tasks used in this thesis). The effect of the level of processing was investigated in Study 1 (Experiment 1 and 2)

and Study 2, while the correlation between visual and auditory perceptual switching was addressed in all the studies (Study 1 to 4).

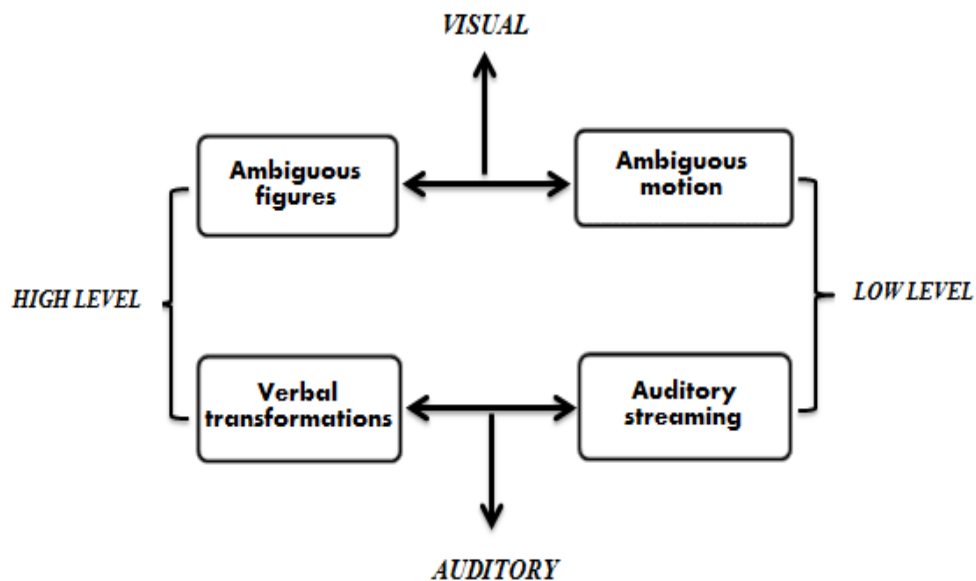


Figure 1.5. Multistable paradigms used throughout the thesis.

These paradigms were selected for both practical and empirical reasons. Firstly, ambiguous figures, verbal transformations, and auditory streaming have been investigated previously in Kondo et al. (2012). An important decisional factor in choosing these paradigms was whether they could be easily disambiguated and verbalised. As Study 2 examined children, another criterion for selection of these paradigms was to be able to make child-friendly tasks. For the high-level visual stimulus, the duck-rabbit figure was used in previous work with children (Doherty & Wimmer, 2005; Gopnik & Rosati, 2001; Mitroff, Sobel, & Gopnik, 2006; Rock, Gopnik, & Hall, 1994; Wimmer & Doherty, 2011; Wimmer & Marx, 2014). The duck-rabbit figure can also be easily disambiguated (Doherty & Wimmer, 2005). For the low-level visual task, ambiguous-structure-from motion was selected instead of other paradigms used before (such as the Necker cube) because its interpretations can be verbalised easier than any abstract stimulus. Left and right movement was assumed

easier to report than perspective changes in images such as the Necker cube. Other low-level visual stimuli would have been difficult either to disambiguate or to be made children friendly. For the auditory modality, there were only two paradigms available and previously investigated: auditory streaming and verbal transformations (Kondo et al., 2012). Therefore, they were further investigated here. For auditory streaming, the pure tones used in adult studies were replaced in the first two studies by more realistic sounds (water dripping sounds). These were synchronised with an animation that helped children understand easier how to report what they hear.

If the processing of ambiguous stimuli is domain general (Kondo et al., 2012), the number of switches in visual and auditory tasks should be positively correlated. If perceptual switching is domain specific (Pressnitzer & Hupé, 2006), then no correlations should emerge. Moreover, based on the previous findings (Strüber & Stadler, 1999) a larger number of switches may be found for high-level stimuli (ambiguous figure and verbal transformations) than for lower-level stimuli (ambiguous motion and auditory streaming). In addition, there might be correlations between the high-level stimuli (ambiguous figure and verbal transformations) but not between the low-level stimuli (ambiguous motion and auditory streaming) (Pressnitzer & Hupé, 2006).

1.3.4 WHERE: Brain Regions Involved in both Visual and Auditory Multistability

A step further in understanding the mechanisms of multistable phenomena and the question of domain generality/specificity is to investigate the potential brain areas that could be commonly involved in visual and auditory multistability. As each modality has distinct low-level brain regions for processing information (i.e., visual cortex for visual information and auditory cortex for auditory information), if there is a common site for

perceptual switching across modalities, then it has to be located in the higher-level brain regions (Kleinschmidt, Sterzer, & Rees, 2012; Sterzer et al., 2009). As will be argued below, a likely common site is the parietal cortex (Cusack, 2005; Kanai, Carmel, Bahrami, & Rees, 2011).

So far, there are no neuroscientific investigations of the neural basis of multistability between modalities in a within-design study. However, there is neurophysiological research within each modality, which supports the hypothesis that perceptual multistability has neuro-correlates in both sensory-specific brain areas and non-sensory brain areas (Long & Toppino, 2004; Cusack, 2005). With regards to visual multistability, most studies acknowledge that the lower-level brain regions have an essential role in the competition between the interpretations (Tong et al., 2006). However, high-level brain regions are also considered to be involved in perceptual switching (Leopold & Logothetis, 1999; Sterzer & Kleinschmidt, 2007; Weinhhammer, Ludwig, Hesselmann, & Sterzer, 2013). The idea that high-level brain regions trigger perceptual transition by selecting one of the possible interpretations, was initially supported by Lumer, Friston, and Rees (1998). In a functional magnetic resonance imaging (fMRI) study on binocular rivalry, participants were tested under two viewing conditions. In the first condition (rivalry condition), participants viewed stimuli consisting of a red-coloured drifting grating shown to one eye and a green-coloured face shown to the other eye. In the second condition, participants were exposed to a replay of their response perception during rivalry (control condition). In both conditions, there was activity recorded in the extra-striate cortex, whereas fronto-parietal cortex was associated with perceptual switching only during the rivalry condition. This finding was interpreted to show that fronto-parietal regions trigger the process of perceptual selection and then promote the selection via feedback to early visual areas (Leopold & Logothetis, 1999). Later on, further evidence using ambiguous structure-from-motion

and the Necker cube showed that the activity in the frontal (Sterzer & Kleinschmidt, 2007) and parietal (Britz, Landis, & Michel, 2008) precedes the neural activity from the occipital cortex.

For auditory multistability, the views are similar. Most of the studies conducted in the last two decades to determine the neural basis of auditory streaming, have predominantly focused on the auditory cortex. In a fMRI study, listeners showed more activity in the auditory cortex when hearing two streams as opposed to one stream (Hill, Bishop, Yadav, & Miller, 2011). However, new studies show that non-primary sensory cortices are also involved. For example, Kondo and Kashino (2009) examined neural correlates of switching between the integrated and segregated percept in an auditory streaming paradigm and found activation in both the non-primary auditory cortex and thalamus. This suggests that thalamo-cortical interactions are important for auditory perceptual switching. In another fMRI study, Cusack (2005) recorded fMRI blood oxygen-level dependent (BOLD) signals from intraparietal sulcus and auditory cortex and found that only the former co-varied with the changes in the listeners' interpretations. This suggests that cortical areas beyond the auditory cortex and thalamus are associated with perceptual switching in auditory multistability.

As parietal cortex was found to be associated with perceptual switching in both visual (Britz et al., 2008) and auditory multistability (Cusack, 2005), it seems reasonable to infer that a potential non-modality specific area involved in multistability could be the parietal cortex. Since fMRI studies allow only correlational interpretations of the data, it is important to be cautious in interpreting the precise role of the brain areas measured in the neurophysiological studies. One of the most reliable demonstrations of a causal role for the parietal cortex in visual multistability comes from studies using transcranial magnetic stimulation (TMS). Previous studies on visual multistability have

shown that stimulation of specific regions of the parietal cortex leads to an increase or decrease of the phase durations (Carmel, Walsh, Lavie, & Rees, 2010; Kanai, Bahrami, & Rees, 2010; Zaretskaya, Thielscher, Logothetis, & Bartels, 2010). From auditory research, there are indications that the intraparietal sulcus (which is located in close proximity to the areas found in visual multistability) is involved in auditory multistability (Cusack, 2005). However, the same parietal regions have not been investigated for both modalities in a within-subject experiment. One way to approach the question of a common or separate neural mechanism in multistable perception is using TMS, which disrupts the normal activity of a specific part of the brain. If parietal cortex is involved in both visual and auditory multistability, we should see an effect on perceptual switching in both modalities after perturbing its activity. If parietal cortex is involved only in visual multistability, then we should see the effect only in visual multistability as previously found. This was investigated in Study 3.

1.4 Common Individual Differences in Perceptual Switching across Modalities?

It is well known that participants display large inter-individual differences in their reports to multistable paradigms (Kleinschmidt et al., 2012; Scocchia, Valsecchi, & Triesch, 2014). When participants are asked to report their perception in multiple experimental stimuli blocks, their perceptual switching between different blocks tends to be more similar at an individual level than when compared to other participants (Denham et al., 2014; Farkas et al., 2016). There is no clear answer to why these individual differences occur. Evidence suggests that genetic factors as well as sensory and cognitive factors are involved (see Scocchia et al., 2014 for a comprehensive review). For example, genetic studies found that heritable factors contribute to these individual differences. However, the influence of heritable factors was not found across all multistable paradigms. While heritable factors were found for binocular rivalry,

perceptual switching in the Necker cube is less genetically determined (Shannon et al., 2011). Patients with different clinical conditions, some with heritable factors involved (e.g., schizophrenia, bipolar disorder, autism), had distinctive switching patterns in various multistable paradigms. When exposed to images with emotional valance, patients with social anxiety report the fearful image as the first percept in a binocular rivalry task. They also display shorter phase durations for negative valance images compared to neutral ones (Singer, Eapen, Grillon, Ungerleider, & Hendler, 2012). Bipolar disorder patients have a lower switching rate in binocular rivalry than non-patients (Pettigrew & Miller, 1998; Vierck et al., 2013). The reduced alternation rate is so consistent that it is proposed as a possible endophenotype of the pathology (Ngo, Mitchell, Martin, & Miller, 2011). However, the decrease in the switching rate is less prominent when patients with bipolar disorder are tested on ambiguous structure-from-motion (Krug, Brunskill, Scarna, Goodwin, & Parker, 2008). Thus, it seems that perceptual switching in different visual paradigms might be related to different heritable factors and that not all paradigms have been identified to be related to the same genetic factors.

Different aspects of brain functioning have also been found to influence the individual differences in perceptual switching rates (Kleinschmidt et al., 2012). For instance, Strüber, Basar-Eroglu, Hoff, and Stadler (2000) found that participants with a large number of switches in an apparent motion paradigm³ show an enhancement of gamma band activity, indicating higher arousal/vigilance. Nakatani and van Leeuwen (2005) observed a characteristic pattern of sequential occipital alpha and frontal theta band activity in frequent switchers, which is assumed to be related to the attentional effort necessary for perceptual switching. The number of perceptual switches also correlates

³ Two brief stationary stimuli are presented in succession at two different locations which gives an impression of movement.

with individual differences in the structure of the parietal cortex (Kanai et al., 2010), indicating a direct relation between higher order brain structures and individual differences in perceptual switching. This evidence shows that the higher-order brain regions play a role in the individual difference manifested in multistable phenomena. Given there are large individual differences in how people respond to perceptual switching and that cognitive factors were found to partly explain the idiosyncrasies in some perceptual tasks, it was questioned whether these individual differences at the perceptual level across modalities are linked in a consistent way with individual differences in cognitive functioning. Because in previous research there was contradictory evidence across modalities and tasks, the individual differences investigated here were executive functions, creativity, and personality. Investigating the individual differences in these functions in relationship with perceptual switching contributes in particular to the understanding of HOW and WHAT questions. In case these individual differences are related consistently across tasks/modalities this will bring evidence for a domain general role for perceptual switching (HOW question). In case the individual differences relate only with the high or low-level tasks this will indicate there is level specificity (WHAT question).

1.4.1 Executive Function and Perceptual Switching

Support for a relation between executive function (i.e., inhibition and set-shifting), creativity and perceptual switching comes from research with children and adults. For instance, Wimmer and Doherty (2011) reported a relationship between the *ability to perceptually switch* between interpretations and inhibitory control (i.e., inhibition of a prepotent response) measured using a Stroop task (Gerstadt, Hong, & Diamond, 1994). Note here the difference between two terms: *the ability to switch* measured once (whether people can change to the second interpretation) and ongoing

perceptual switching (the switching back and forth between interpretations) which is measured continuously for a certain time. In their study, children were asked after 0, 30, and 60 seconds to report specific features of ambiguous figures such as the duck-rabbit. Children with better performance in the Stroop task had better abilities to switch, which was argued to show that inhibitory control plays an important role for the ability to switch. In the same study, mental set-shifting abilities were measured using the Dimensional Change Card Sort task (Frye, Zelazo, & Palfai, 1995) but they were not related to the ability to switch.

Another source of support for a relationship between executive function and perceptual switching comes from neuroscientific evidence showing that patients with lesions in prefrontal cortex (thus with inhibitory control impairments) were less able than healthy subjects to recognise and intentionally switch between the two possible interpretations of ambiguous figures (Windmann, Wehrmann, Calabrese, & Gunturkun, 2006). Ricci and Blundo (1990) also showed that in patients with frontal damage, there was a negative correlation between the ability to switch and the number of errors in the Wisconsin Card Sorting Test, which is a measure of cognitive flexibility (Grant & Berg, 1948). This finding suggests that in patients with frontal lobe damage, a poor performance on an executive task relates to a weak ability to see other interpretations in visual multistable paradigms.

Indirect support for the relation between inhibition and perceptual switching comes from studies showing that bilinguals aged 4 to 5 years—which are thought to have superior inhibitory control (Bialystok & Martin, 2004; Carlson & Meltzoff, 2008; Wimmer & Marx, 2014; see also Paap & Greenberg, 2013)—are more likely to perceive alternative interpretations of ambiguous figures than monolinguals (Wimmer & Marx, 2014; see also Bialystok & Shapero, 2005).

Although inhibition and/or set-shifting may be necessary for the ontogenetic onset of perceptual switching (Wimmer & Doherty, 2011; Wimmer & Marx, 2014; Windmann et al., 2006), neither of these appear to affect the number of switches experienced by adults for ambiguous visual stimuli. For example, art students produce a greater number of switches for the ambiguous rotating cylinder stimulus than non-art students (Chamberlain, Swinnen, Heeren, & Wagemans, 2017). However, their switching rate is not correlated with individual differences in inhibition or cognitive flexibility measured with a set-shifting task (Diamond, 2013). Similarly, inhibition, working memory, and cognitive flexibility was found to be unrelated to perceptual switching produced in response to the Necker cube in adults (Díaz-Santos et al., 2017).

The relationship between executive functions (inhibition and set-shifting) and auditory multistability has been investigated in adults by Farkas et al. (2016). They found that participants with better inhibitory abilities—measured using a Stroop task (Lansbergen, Kenemans, & van Engeland, 2007)—had longer average segregated phase durations and a lower switching rate. That is, longer response time in a Stroop task (less inhibition) was positively associated with the number of switches in auditory streaming. Therefore, in the auditory domain, inhibition appears to play a negative role in the number of switches between alternatives.

In the same study, it was also found that there was a strong correlation between set-shifting—measured with a Verbal fluency task (Troyer, Moscovitch, & Winocur, 1997)—and the proportion of the integration interpretation (the non-dominant interpretation) in auditory streaming. Set-shifting seems to facilitate switching in the auditory domain. Notice that in the study on children by Wimmer and Doherty (2011) the focus was on executive function and the *ability to switch* in response to a visual stimulus, while in the study by Farkas et al. (2016) the focus was on the relationship

between the executive function and *switching rate* in auditory streaming. If inhibition and set-shifting have the same effects on both visual and auditory multistability, this will support the hypothesis that executive function is commonly involved in perceptual switching regardless of modality. However, if the relationship between inhibition, set-shifting and perceptual switching is found only in one of the modalities, this might indicate modality specific relationships. It is also possible that executive function relates with perceptual switching in a task-specific way. These aspects will be explored in Study 1, Study 2 and Study 4 of this thesis.

1.4.2 Creativity and Perceptual Switching

Individual differences in creativity—measured with divergent thinking tasks or creative potential measures—have also been found to relate to perceptual switching, especially in the visual modality. In two experiments, Wiseman, Watt, Gilhooly, and Georgiou (2011) examined the relationship between creative potential and the ease of switching to the second interpretation of an ambiguous figure. In a first online experiment, participants viewed the duck-rabbit figure after they were told about the two interpretations (in a disambiguation phase). In a questionnaire, participants were asked about which interpretation they saw first, how easy it was to see the other interpretation after their first interpretation (*very easily, easily, not at all easily, cannot see the other animal*), and whether they would describe themselves as artistically creative, and as a creative problem solvers (*definitely yes, yes, uncertain, no, definitely no*). The results showed that participants who found it easier to switch to the second interpretation would rate themselves as being more creative. A second experiment investigated whether the results found in Experiment 1 would replicate by use of a standard creative problem-solving task. Specifically, participants were asked to complete one of two shortened versions of Guilford's Alternative Uses task (Guilford,

Christensen, Merrifield, & Wilson, 1978) and asked to name as many possible uses for either a brick or a paperclip. Participants were shown the duck-rabbit figure for approximately 30 seconds and then requested to complete the same kind of questionnaire used in Experiment 1. The results indicated that the production of unusual uses for familiar objects was positively related to the ease of perceptual switching, suggesting an association between the two phenomena.

In another study, Doherty and Mair (2012) explored the relationship between creativity and the ease of perceptual switching in three visual paradigms (duck-rabbit, vase-face and Necker cube). Specifically, participants were asked to view each of the ambiguous stimuli for 60 seconds, and to indicate the changes in their interpretations by making a dash in a test booklet for each subsequent reversal. Creativity was measured with the M. A. Wallach and Kogan (1965) Pattern meanings test which consists of abstract visual patterns that can be interpreted in multiple ways. Results showed a positive correlation between the creativity score and the number of switches in response to the ambiguous figures. This supports a relation between perceptual switching in visual tasks and creativity. Nevertheless, the relation was stronger for the two figures involving switching of perspective (Necker cube and vase-face) than for the figure involving the switching of the meaning/content (duck-rabbit), suggesting that distinct visual paradigms might relate differently to creativity.

A few studies argue that being exposed to ambiguous images facilitates divergent thinking. For instance, Wu, Gu, and Zhang (2016) presented ambiguous and non-ambiguous stimuli before an alternative uses task (participants had to report as many unusual uses for a visual stimulus as they could) or a general uses task (participants had to report common uses for a stimulus). It was found that participants who saw the ambiguous figure first were able to generate significantly more solutions in

the alternative uses task, compared to participants who were shown first an unambiguous image. It was argued that seeing ambiguous figures facilitates creative ideas.

Recently, Laukkonen and Tangen (2017) investigated the relationship between insightful problem solving and perceptual switching and found a correlation between the ability to identify the two interpretations of an ambiguous image and solving verbal insightful problems. They proposed that one of the commonalities between the two phenomena comes from the fact that both involve resolving conflicting information. In a different experiment it was investigated whether perceptual switching and the ability to solve insightful tasks rely on the same neural networks activity involved in conflict solving (Botvinick, Braver, Barch, Carter, & Cohen, 2001). Participants were asked to solve an insight problem after they were shown either an ambiguous Necker cube or a disambiguated 3D drawing of a cube. The results showed that more insight problems were solved after participants viewed the Necker cube than after the non-ambiguous figure. It was argued that conflicting information from the ambiguous figure facilitated insightful problem solving as it activated the cognitive control mechanisms required to overcome the informational conflict, which is speculated to be involved in both phenomena.

In the auditory domain, one single study has looked at the relationship between creativity and perceptual switching in an auditory streaming paradigm, but it failed to find any relationship (Farkas et al., 2016). In this study, participants completed two divergent thinking tasks and an index of creativity was computed from the two. In the first task (i.e., use of objects; Lezak, Howieson, Bigler, & Tranel, 2012), participants were instructed to produce as many novel uses as they could for three common objects (brick, paperclip, and newspaper). In the second task (Caption generation; Jung, Grazioplene, Caprihan, Chavez, & Haier, 2010), participants were instructed to write as

many captions as they could for three *New Yorker Magazine* cartoons. The creativity index was not associated with the switching rates in auditory streaming. One reason why no relation was found could be the type of creativity tasks or the bistable paradigms used. A further exploration with another auditory multistable task (i.e., verbal transformations) or measures of creativity (i.e., divergent thinking tasks) would be necessary to show whether or not the relation between perceptual switching and creativity is task-specific.

In conclusion, it is important to determine whether creativity is related only with perceptual switching in vision or whether it is also associated with perceptual switching in the auditory domain. This will be investigated in Study 1 and Study 2.

1.4.3 Personality and Perceptual Switching

Personality characteristics were also linked to perceptual switching in vision and audition. Bosten et al. (2015) aimed to find the genetic substrate of binocular function which is the ability to coordinate the activity of the eyes so that the visual fields of both eyes are merged into one image. Between many other measures, personality was assessed using the mini IPIP Big Five scale (Donnell, Maki, Pliskin, & Kraus, 2004) and correlated with perceptual switching in a binocular rivalry task. Results showed that personality was not related to median phase durations in the binocular rivalry task. This study is inconsistent with old findings showing that introverts have fewer number of perceptual switches in figure-ground stimuli (e.g., vase-faces) than extroverts (Lindauer & Reukauf, 1971). However, this discrepancy could be also due to the different personality measures and stimuli used.

In the auditory domain, Farkas et al. (2016) found a positive relationship between switching rates in auditory streaming and a measure of personality (i.e., ego-resiliency). Ego-resiliency is an individual characteristic that reflects adaptability to environmental

stress and change (Block & Kremen, 1996). It is thought that people with low scores on ego-resiliency do not respond adaptively and flexibly to the environment, remaining trapped in the same pattern of actions. Farkas et al. (2016) found that ego-resiliency (measured with ER89 questionnaire; Block & Kremen, 1996) is related to the individual differences observed in perceptual switching in an auditory streaming task.

While in the visual modality there are contradictory findings regarding the relation between perceptual switching and personality, there is only one study in the auditory domain to show a relationship between personality and perceptual switching. The question is to what extent measures of personality relate to perceptual switching regardless of modality and level of processing. In case the relationship is only with perceptual switching in the auditory modality, the results will speak for a domain specific relationship. If there is a relationship with visual multistability as well, the results will favour the modality generality hypothesis. There is also a possibility that the relationship is task-specific, not being dependent on modality but on the type of multistable paradigm used.

1.4.4 Perceptual Switching in Development

An important approach for the understanding the domain generality/specificity question is to investigate it during ontogenetic development. So far, there were no attempts to investigate this question in children. A developmental trajectory investigation can reveal the mechanisms by which perceptual switching occurs and develops. Moreover, a developmental approach to the question of domain generality/specificity helps to understand how the relationship between perceptual switching across modalities and tasks manifests in development. Because both perception and cognition are still developing during childhood, one such investigation could help to understand whether: 1) there are modality specific developments; 2) there

are task demands effects on perceptual switching that manifest differently across ages, 3) developments in executive functions or creativity underline these changes. As multistability has been reported in vision, audition, and even olfaction, this suggests that whatever causes perceptual switching, it must be an intrinsic part of normal perception (Kleinschmidt et al., 2012; Leopold & Logothetis, 1999; Sterzer & Kleinschmidt, 2007). This conclusion would be strengthened by the demonstration of perceptual switching in childhood across a range of tasks and modalities. It is known that myelination of connections between sensory areas and prefrontal cortex takes place quite late in development, even into early adulthood (Fuster, 2002). If prefrontal cortex has an important role in switching as suggested by some studies (Brascamp, Sterzer, Blake, & Knapen, 2018; Kleinschmidt et al., 2012; Leopold & Logothetis, 1999; Sterzer & Kleinschmidt, 2007; Sterzer et al., 2009) then we might expect to see differences in perceptual switching behaviour at different ages. Moreover, the activity in prefrontal cortex has been associated with the activity in executive functions (Diamond, 2013) and executive functions have been suggested to be a requirement for the ability to switch (Wimmer & Doherty, 2011). If executive functions have a role in perceptual switching, we would expect that the number of switches relates with measures of executive functions.

From a developmental perspective, children's basic visual and auditory abilities develop in the first year of life (Hainline, 1998; Litovsky, 2015). However, children are not thought to perceive ambiguity in their environment until around the age of 5, when the ability to switch between two competing interpretations of a stimulus develops. Development of perceptual switching has been investigated only in a few studies in the visual modality (Doherty & Wimmer, 2005; Gopnik & Rosati, 2001; Mitroff et al., 2006; Rock et al., 1994; Wimmer & Doherty, 2011) and even fewer in the auditory domain (Sussman, Wong, Horvath, Winkler, & Wang, 2007; Warren & Warren, 1966).

In one of the first developmental studies of perceptual switching, Rock et al. (1994) showed that uninformed 3- to 5-year-old children did not switch to a second interpretation when presented with visual ambiguous figures such as the duck-rabbit, the vase-faces, and the man-mouse. Furthermore, even when instructed about the two interpretations (through a disambiguated pre-testing phase), the 3-year-olds were not able to switch their perception. In the visual domain, research suggests that the ability to switch between interpretations of an ambiguous figure develops between the ages of 4 and 5, provided the child is first instructed about the two interpretations (Doherty & Wimmer, 2005; Gopnik & Rosati, 2001; Wimmer & Doherty, 2011). For instance, Wimmer and Doherty (2011) assessed with a feature identification task the ability to switch interpretations in ambiguous figures in children between 3 and 5 years of age. After children were told about the two alternative interpretations, they had to produce the contrary interpretation of what was offered by the experimenter (e.g., *"I say it's a duck, what else can it be?"*) and then indicate on request specific features of the alternative interpretation they offered (*"Can you point to the mouth of the rabbit?"*). 3-year-olds could not switch interpretations. At the age of 4, children's performance increased significantly, reaching ceiling by the age of 5.

In the auditory domain, Warren and Warren (1966) investigated whether 5-year olds reported hearing different words when exposed to verbal transformations. Children were asked to report all of the interpretations that they could hear while still listening to the sounds. Children listened to repeating words such as "trees", "see" or non-words like "flime" in the form of verbal transformations. The results showed that they perceived on average 1 to 1.6 changes for "trees" and "see" and 5 switches for "flime". This suggests that the ability to perceive new interpretations in verbal transformations develops around the age of 5. However, in this experiment children were not instructed about the alternative interpretations, which is known from visual research to be essential

for the ability to switch in children before the age of 5 (Gopnik & Rosati, 2001; Rock et al., 1994). Therefore, it is unclear whether the results will remain unchanged when children are instructed about the alternative interpretations.

In another study, Sussman et al. (2007) investigated the minimum frequency difference between “L” and “H” tones needed for stream segregation to occur. Children aged 5–8, 9–11 years, and young adults had to listen to a sequence of the auditory streaming paradigm and report whether they heard the sounds as integrated or segregated. The “L” tone had a fixed frequency, and “H” tones were presented at different frequency separation thresholds (1, 7, 11, 15, 19, 23, 27 semitones: ST). The results showed that adults hear segregation at much lower frequency differences (5ST) than older children (7ST). The younger children only heard segregation consistently from 23ST frequency differences upwards. These data suggest that the ability to switch in auditory streaming depends on the difference between low and high tones and that there is an inverse relationship between age and the minimum frequency difference needed for stream segregation. The younger the participants, the larger the frequency differences need to be for participants to hear segregation. As the study did not assess perceptual switching rates, it is unknown how perceptual switching in auditory streaming manifests in children once they can hear segregation. Overall, it is suggested that the ability to switch interpretations develops over preschool age but the specific developmental trajectory appears to be stimulus and modality specific. However, this remains an open question as task methodologies differed across studies and no comparison across tasks was conducted with the same developmental sample, which will be addressed in the current research.

Warren and Warren (1966) in a second experiment of the study described above examined 6-, 8- and 10-year-old children using the same methods as in Experiment 1.

The number of switches increased significantly between 6 and 10 years and from 10 years to young adulthood (Warren & Warren, 1966). An increase in the number of switches from 10 years to young adulthood has also been shown for an ambiguous motion task (Ehlers, Struber, & Basar-Eroglu, 2015), suggesting that the mechanisms underlying switching rate undergo a maturational process. There is evidence from the visual perceptual domain showing that although important visual functions emerge in the first year of life (Kovács, 2000) there are ongoing developments in some other functions. For instance Nayar, Franchak, Adolph, and Kiorpes (2015) found that between the age of 4 and 7 there are developmental shifts from local processing styles to increasingly global processing styles. Moreover, between the age of 4 and 10 children become more susceptible to visual illusions (Bremner et al., 2016; Doherty, Campbell, Tsuji, & Phillips, 2010) suggesting a shift from more local processing styles to more global processing styles. Thus, given children's processing style changes one might also expect developmental changes in switching rate with age after the ability to switch has developed which was investigated in the current research.

It is still unknown what mechanisms underlie perceptual switching rate. Inhibitory control has shown to predict the ability to switch (Wimmer & Doherty, 2011) whereas it might be unrelated to the number of switches experienced (Díaz-Santos et al., 2017; Farkas et al., 2016). Wimmer and Doherty (2011) suggest that to switch between alternative interpretations of an ambiguous stimulus, one may need to inhibit the current, dominant interpretation. Thus, increased inhibition may be needed for the ability to switch (Wimmer & Doherty, 2011). In support of this, and as mentioned previously, Wimmer and Doherty (2011) found that perceptual switching is predicted by inhibitory control measured with a Stroop task (Stroop, 1935). Additional support for the relation between inhibitory capacity and ability to switch comes from neuroscientific evidence showing that patients with lesions of the prefrontal cortex (and

thus with impaired inhibitory control) were less able to recognise and intentionally switch between interpretations of ambiguous figures than healthy subjects (Windmann et al., 2006). In indirect support of this claim, bilinguals (aged 4 to 5 years)—which are thought to have superior inhibitory control (Bialystok & Martin, 2004; Carlson & Meltzoff, 2008; Wimmer & Marx, 2014; see also Paap & Greenberg, 2013)—are more able to perceive the alternative interpretation of ambiguous figures than monolinguals (Wimmer & Marx, 2014; see also Bialystok & Shapero, 2005). This advantage might also be explained by bilinguals' advantage in set-shifting (Prior & MacWhinney, 2010). Both set-shifting and perceptual switching requires switching between two competing interpretations or responses. Perceptual switching was also related to mental imagery (as an important component of creative thinking; Jankowska & Karwowski, 2015; Palmiero et al., 2016). For instance, Wimmer and Doherty (2011) found that mental imagery—the ability to understand that a picture is a representation of something which can be manipulated in our own minds—was a key component skill that facilitated perceptual switching. One way to interpret this result is to consider that imagery or/and creativity may be necessary for children to construct alternative interpretations of ambiguous stimuli. Both perceptual switching and creative thinking require the formation of an alternative interpretation and imposition of the imagined structure onto the figure, in order to come up with a different interpretation (Schooler & Melcher, 1995).

To my knowledge, there are no studies in the auditory modality that have investigated the cognitive abilities necessary for perceptual switching to occur. In adults both inhibition and set-shifting have been associated with the switching rate in auditory streaming (Farkas et al., 2016). Set-shifting ability, measured with a verbal fluency task (Troyer et al., 1997), correlated with the proportion of hearing *integrated* (i.e., the non-dominant interpretation) in auditory streaming (Farkas et al., 2016). Thus, set-shifting

facilitates switching in the auditory domain. However, a longer response time in a Stroop task (less inhibition) was positively associated with the number of switches in auditory streaming. Therefore, in the auditory domain, inhibition appears to play a negative role in the number of switches (Farkas et al., 2016).

In adults studies, it is also assumed that creativity is related to the ability to switch in visual ambiguous stimuli (Doherty & Mair, 2012; Laukkonen & Tangen, 2017; Wiseman et al., 2011; Wu et al., 2016) (see Section 1.4.2). In the auditory modality, this relation has not been found (Farkas et al., 2016).

Given the mixed results, an important theoretical question is whether in children, the processes underlying perceptual switching across modalities are domain specific or domain general. As discussed, only a few studies have investigated the onset of perceptual switching in children (Doherty & Wimmer, 2005; Gopnik & Rosati, 2001; Rock et al., 1994; Wimmer & Doherty, 2011; Wimmer & Marx, 2014) and no studies have investigated both visual and auditory bistability in the same sample of children. While it is known that children develop the ability to switch by the age of 5 (Warren & Warren, 1966; Wimmer & Doherty, 2011), it is unknown how the number of switches changes with age and whether the number of switches in visual and auditory modalities follows the same developmental schedule and/or draws on common cognitive mechanisms. Despite lively debates in the adult literature concerning the domain specificity of perceptual switching abilities, there has been no attempt to observe the development of these skills; that is, to determine if they emerge to the same schedule. This is important, as the common hypothesis would predict that visual and auditory switching abilities should emerge together and associate with the same cognitive skills. Producing the first systematic developmental picture of this phenomenon, Study 2 tracks the visual and auditory bistability of 6-, 8-, and 10-year-old children. Producing

the first systematic developmental study of this phenomenon, Study 2 tracks the visual and auditory bistability of 6-, 8-, and 10-year-old children. These age groups were selected to investigate the ongoing perceptual switching after the ability to switch develops, thus after the age of 5. Previous findings on verbal transformations (Warren & Warren, 1966) showed an increase in the number of switches in the same range of age groups. In the visual modality, it has also been shown that there are changes in perceptual switching styles between the age of 4 and 10 (Bremner et al., 2016; Doherty et al., 2010). Given these changes were found in separate studies it was reasoned that 6 to 10 years of age was the appropriate age range to use in order to observe whether the developments in visual and auditory perceptual switches are modality/task specific or general.

1.5 Summary of the Aims and Findings

The current research reports four studies (five experiments) that examine visual and auditory multistability in adults and children.

The general aim was to investigate to what extent perceptual switching in visual and auditory multistability share common or separate mechanisms (HOW). These questions were explored using four ambiguous tasks that give rise to multistability and which are thought to involve competition at different levels of cognitive processing (WHAT): auditory streaming and ambiguous-structure-from-motion (low-level multistability), and verbal transformations and duck-rabbit ambiguous figure (high-level multistability). Perceptual switching of adults and children was also investigated in relationship to individual differences in executive function (inhibition and set-shifting), creativity, and personality traits.

Specifically, in the first two experiments (Study 1) visual and auditory ambiguous stimuli with different levels of complexity were investigated to clarify to

what extent perceptual switching manifests domain- and level-specificity or -generality. Moreover, it was explored whether perceptual switching across modalities and levels of processing relates consistently to individual differences in inhibition, set-shifting, creativity, and ego-resiliency. In Experiment 1 (Study 1) an effect of level was found, with participants reporting higher switching rates in the high-level tasks than the low-level tasks, in both modalities. However, in Experiment 2 (Study 1) the level effect was significant only for the auditory tasks. Participants experienced more perceptual switching in verbal transformations than in auditory streaming, while perceptual switching rates in the two visual tasks did not differ significantly. An exploration of the correlations between visual and auditory switching rates was inconclusive; in Experiment 1, perceptual switching rates in visual and auditory tasks did not correlate, while in Experiment 2 the perceptual switching rate in the two high-level tasks correlated. Inhibition, set-shifting, and creativity correlated with perceptual switching in some of the tasks, although not in a consistent manner.

In Study 2 (Experiment 3), a developmental approach was employed to examine the perceptual switching in visual and auditory bistability in children. It was studied whether perceptual switching increases with age and whether developments in creativity and inhibition could explain these differences. It was found that perceptual switching rates increased with age and that inhibition control and creativity were not associated with the number of perceptual switches.

In Study 3 (Experiment 4), the role of superior parietal cortex was investigated in both visual and auditory multistability (WHERE). Disruption of normal brain activity by use of TMS showed that superior and anterior superior parietal cortices were not causally involved in either visual or auditory multistability, indicating that these regions are not commonly involved in perceptual switching across modalities.

Finally, in Study 4 (Experiment 5), the aim was to clarify whether there is a difference in perceptual switching when participants have to report two main alternative interpretations versus three interpretations, and whether attentional control affects the relationship between perceptual switching across modalities. The results showed that perceptual switching in vision and audition correlates in both the two- and three-alternative groups. This supports the idea that there are common mechanisms involved in perceptual switching across modalities. Moreover, participants switched similarly across modalities and the effect of attentional bias was similar for the two modalities, suggesting that voluntary control cannot explain the commonalities found.

Overall, the results from these studies suggest that perceptual switching in vision and audition follows common principles of functioning. Although in Study 1 the correlation analyses were not consistent, the positive correlations between perceptual switching in visual and auditory tasks in Study 3 and Study 4 suggest that the switching rates across modalities manifest domain general mechanisms. As attentional control did not influence the correlations, this suggests that more low-level sensorial mechanisms are at the cause of the commonalities. Both in adults and children, there is no strong evidence that individual differences in executive function, creativity, or personality commonly relate to perceptual switching regardless of modality and task. Task differences across stimuli found in adults and children suggest that perceptual switching manifest task specific characteristics, supporting the distributed mechanisms hypothesis.

1.6 Statistical Note

Throughout the thesis, three main kinds of analyses were conducted: linear mixed models, repeated measures ANOVA (rANOVA) and Spearman correlations (plus partial correlations to control for age effects in Study 2). The differences across the four perceptual tasks were investigated with linear mixed models in Studies 1 and Study 3 or

with rANOVA in Study 2 and Study 4. The relationships between perceptual switching across perceptual tasks and the measures of individual differences were assessed with Spearman correlational analyses.

In Study 1 and Study 2, *no reports* (durations in which no interpretation was selected) are treated differently, consistently with the traditional practices in children and adult research. In children research no reports are treated as zero because this allows to investigate when perceptual switch occurs (Gopnik & Rosati, 2011; Mitroff et al., 2006; Rock et al., 1994; Wimmer & Doherty, 2011). On the other hand, in the adult literature it is difficult to tell the reason for a no-switch because it is assumed that typical-developed adults are able to switch if enough time is given (Denham et al., 2012; Leopold & Logothetis, 1999), especially after participants are instructed about the different interpretations (Mitroff et al., 2006). Moreover, when perceptual switching is recorded in more than one block, if participants do not switch in some of the blocks but switch in others it is difficult to tell the reason for these no switches. A zero value could in principle distort the natural distribution of the number of switches, while missing values would not do this. For these reasons, in the main text of the thesis the no switches are treated as missing values for Study 1 on adults and as zero in Study 2 on children.

1.6.1 Perceptual Tasks

The main variables of interest across the studies were: the number of switches (or the phase durations in the TMS experiment), the initial reaction times, and the first phase durations. Note that phase durations are inversely proportional to switching rates; the larger the phase durations, the smaller the switching rates. The switching rate represents the number of perceptual switches per time limit. The initial reaction time represents the time when the first perceptual decision was made. The first phase

duration is the time in which the first interpretation was maintained until the first switch occurred.

The design of each study involved repeated measurements of perceptual switching across several experimental blocks (five in Study 1, three in Study 2, four in Study 3 and Study 4). The traditional analyses for this kind of data are repeated measures ANOVA (rANOVA). In the first two experiments, some characteristics of the experimental design and preliminary analyses suggested that they were not suitable for this traditional analysis and mixed model analyses were used instead. Firstly, the first experiment had two tasks nested within one of the four groups, which made it unsuitable for the classical rANOVA (Gelman & Hill, 2007; Zuur, Ieno, Walker, Saveliev, & Smith, 2009). Secondly, in Study 1 there were missing values, which rANOVA does not handle well as it requires balanced data sets with no missing cells (Gelman & Hill, 2007; Zuur et al., 2009). Mixed models are considered more suitable than rANOVAs in these instances, as they allow to use a more appropriate covariance structure of the random effects (Gelman, 2005; Wang, 2013).

The mixed model analyses were conducted in RStudio (Team, 2015) using *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2015) as a front end to the package *lme4* (Bates et al., 2015), as it provides p-values. *P*-values were obtained with Kenward-Roger approximations for degrees of freedom, as it is considered to achieve the best approximation of an F-distribution (Spilke, Piepho, & Hu, 2005). The degrees of freedom that were not integers (because of the missing values) were reported with full numbers, rounded to the nearest integer.

1.6.2 Random Effects

In research on multistability, it is almost unanimously accepted that perceptual switching varies significantly between participants. Their overall perceptual switching

rates might vary for different reasons. For instance, some participants may have, on average, a relatively larger switching rate than others (irrespective of the task), while other participants may have, on average, lower switching rates (this could have been triggered by random external factors such as a bad headache or tiredness). Additionally, participants may differ in their responses to the experimental manipulations. Some participants might have lower switching rates in a visual task than in an auditory task, while others might show the opposite trend. Various types of random effects are required to model these kinds of random variation (Barr, Levy, Scheepers, & Tily, 2013). A random intercept takes into account how each participant's average switching rates (irrespective of task) may differ, while random slopes are necessary to address any variability in the repeated measures of the perceptual switching across tasks. Many authors stress the importance of random slopes in a repeated measures design, as omitting them in the presence of variability among random slopes can lead to increased Type I error rates (Barr et al., 2013; Schielzeth & Forstmeier, 2008). In Experiment 2, the random structure was the following: $(1 + \text{Modality} * \text{Level} | \text{Participant})$ ⁴ while in Experiment 1, only Participant ID was set as a random intercept $(1 | \text{Participant})$ ⁵ as the data structure did not allow a random slope for Modality*Level at the individual level. In Study 1 Modality and Level were specified as fixed effects, including their interaction. To ease interpretation, only the significance of the main effects and their

⁴ Model \leftarrow lmerTest(log (DV) ~ Modality * Level + (1 + Modality * Level |ID), data = dataSetFile, na.action = "na.omit")

Note. DV is the dependent variable, the Modality*Level specifies the main effects of each fixed factor, plus their interaction (*), and (1+Modality*Level |ID) is the random effect of Modality*Level at the individual level (ID), na.action = "na.omit" is for specifying that the missing data are omitted from the analyses.

⁵ Model \leftarrow lmerTest(log (DV) ~ Modality * Level + (1 |ID), data = dataSetFile, na.action = "na.omit")

interaction are reported. Interactions were explored with *lsmeans* package (Lenth, 2017) using Holm-Bonferroni pairwise comparisons tests for p-values adjustments for multiple testing.

The visual exploration of the residuals (i.e. the difference between the observed value of the dependent variable and the predicted value) for the dependent variables showed non-normality of residuals (but lognormal distributions) and heteroscedasticity. Thus, responses for each dependent variable were logarithmically transformed to normalise the distribution of residuals and reduce the influence of extreme values (Zuur et al., 2009). The assumptions of normality, homogeneity and collinearity of variance, were met after the log-transformation of the response variable as displayed in visual inspections of the data (see Appendix A for an example of assumptions violation and their correction after lognormal transformation). To make the understanding of the results easier, descriptive statistics (in tables, figures and averages reported in the Results sections) are reported on the untransformed data.

1.6.3 Relationships between Individual Differences Tasks and Perceptual Tasks

Correlations were calculated firstly to investigate whether perceptual switching across the four perceptual tasks relate to each other, and secondly to examine whether the perceptual switching across the four perceptual tasks relate consistently with the individual differences measures. Because the Pearson test is not recommended when there are extreme values or skewed distributions (Kowalski, 1972), Spearman rank correlations were performed instead. The α level was .05. The correlational and rANOVA analyses were conducted in IBM SPSS Statistics V22.

1.6.4 Short Phase Durations

Throughout the thesis, perceptual phases shorter than 300ms were discarded because they might reflect inaccurate switching between the responses keys (Moreno-

Bote et al., 2010). The total amount of short phase durations excluded was between .04 and .05% in Study 1, 2, and 4, and 0.8% in Study 3.

1.6.5 1st vs. Subsequent Phase Durations.

It is known that the first phase duration tends to be longer than subsequent phases, which reflects different processes, and that its removal stabilises the phase durations (Denham et al., 2014; Hupé & Pressnitzer, 2012). In Study 1, 2, and 3, the first phase duration was significantly different from the subsequent phase durations and therefore removed from the analyses conducted on the switching rates.

1.6.6 Sample size justification

The sample sizes for each of the studies was chosen based on the available information about the effect sizes from the existing literature.

Study 1. In the previous study of Pressnitzer and Hupé (2006) the correlations between the number of switches in the visual and auditory tasks was $r = .40$, which corresponds to a Cohen $d' = 0.872$. For this large effect size, we need 21 participants to assure a power of 80% to find a correlation between perceptual switching across modalities. In all adult studies reported in this thesis, the number of participants is around this number (Experiment 1) or slightly higher (Experiment 2). Statistical power increases also when the same observation is made across a series of blocks. Thus, Study 3 and Study 4 should have even higher statistical power, as the number of blocks was increased substantially compared to the previous study of Pressnitzer and Hupé (2006), which used only a single four-minute block.

Study 2. No published research with children has used the same approach to test the number of switching across four ambiguous tasks, with more than one experimental block. As previous research on children has focused predominantly on the ability to

switch across ages, the effect size for Study 2 was based on the effect size from these lines of studies. Previous findings do not explicitly report effect sizes, however, from my own calculations, the age differences for the prompted ambiguous figures switching reported by Wimmer and Doherty (2011) at page 56 reveal a rather large effect size; Cohen's $d = .80$. For the magnitude of this effect size, we would need 26 participants to assure we have a power of 80%. As in Study 2 the final sample size for each group was around 21 participants, the statistical power is slightly smaller but still good enough to assure the results reflect a true effect.

Study 3. To estimate how many participants were needed for the TMS study, the previous size effects from Carmel et al. (2010) were used. Specifically, a Cohen's $d = 1.12$ was reported for the difference between the phase durations in the posterior superior parietal cortex and a no TMS control condition. For such a large effect size, we would only need six participants in a within study to have a statistical power of 80%. As Study 3 examined two perceptual tasks compared to Carmel et al. (2010), the aim was to have at least double the number of participants compared to what Carmel et al (2010) had. Other studies also use a relatively small number of participants: 12 participants in Kanai et al. (2010) or 15 participants in Zaretskaya et al. (2010). Thus, it was reasoned that a total number of 30 participants would provide enough statistical power to detect true effects.

1.6.7 Numbers

All reported numbers were rounded to the nearest full number if the decimal was .05 or higher and rounded down to the nearest number if the decimal was below .05.

All the units of measurement for reaction times (RTs) were reported in seconds.

The null hypothesis was tested for 2-tailed tests.

The significance levels were labelled accordingly:

$p < .05$ is labelled with *

$p < .01$ is labelled with **

$p < .001$ is labelled with ***

2 STUDY 1: EFFECTS OF MODALITY AND LEVEL OF PROCESSING ON VISUAL AND AUDITORY MULTISTABILITY

1.1. Introduction Experiment 1

Study 1 investigates to what degree visual and auditory multistability manifest domain general or domain specific characteristics. This question has been previously explored in the three studies detailed in the general introduction (Hupé et al., 2008; Kondo et al., 2012; Pressnitzer & Hupé, 2006). Although all these studies appear to argue for the same thing—that perceptual switching in visual and auditory multistability manifest both common and distinct principles of functioning—the correlational results reported are contradictory. While Pressnitzer and Hupé (2006) did not find correlations between the number of switches in auditory streaming and visual plaids, Kondo et al. (2012) reported positive correlations between the number of switches in auditory streaming, verbal transformations, ambiguous figures (Necker cube and vase-faces), and visual plaids. As discussed in the general introduction, one reason why Pressnitzer and Hupé (2006) did not find correlations might be that the association is less robust when stimuli are abstract and without semantic content (Kondo et al., 2012; Strüber & Stadler, 1999; van Ee et al., 2005; Wolf & Hochstein, 2011). The question of domain generality/specificity was thus addressed in this study by comparing perceptual switching in visual and auditory multistability as well as between lower- and higher-level stimuli.

2.1 Aims

The main aim of this study was to examine whether perceptual switching in visual and auditory multistability and across levels of processing are related or not. The second aim was to see whether executive functions, creativity, and personality relate to

perceptual switching in both modalities. Both aims are explored in two experiments in Study 1.

If the processing of ambiguous stimuli is domain-general, similar switching patterns should be observed in vision and audition, at least separately within each level. Based on the previous findings (Kashino & Kondo, 2012; Struber & Stadler, 1999; van Ee, van Dam, & Brouwer, 2005; Wolf & Hochstein, 2011), higher switching rates are expected for the higher-level stimuli than for the lower-level ones. Secondly, if perceptual switching is related to the same individual differences, better performance on inhibition, set-shifting, creativity or personality should correlate in the same direction with the switching rates in both the visual and auditory tasks.

Experiment 1

2.2 Method

2.2.1 Participants

One hundred and eleven English-speaking students (26 male, $M_{\text{age}} = 21.7$, $SD = 5$) were recruited online and received course credit or remuneration. All participants reported corrected-to-normal vision and normal hearing. Informed written consent was obtained from all participants, and the University of Plymouth Ethics Committee approved the study.

2.2.2 Design

Participants completed a series of eight tasks in a quiet laboratory room for around 65 minutes. To be able to include all the tasks in the experiment, and with consideration for time constraints, an unbalanced design with different observations across participants was chosen. Divided into four experimental groups, participants carried out a visual and an auditory task in the following way: Group 1 was assigned to

ambiguous figure and verbal transformations; Group 2 was assigned to ambiguous motion and auditory streaming; Group 3 was assigned to ambiguous motion and verbal transformations; and Group 4 was assigned to ambiguous figure and auditory streaming. Participants also carried out: a Stroop task (Simpson & Riggs, 2005), FIST (Dick, 2014), Pattern meaning task (M. A. Wallach & Kogan, 1965), Verbal fluency (Troyer et al., 1997), and two tasks aimed to control for false responses (see Section 2.2.3.3). Participants were assigned randomly to one of the four groups. Eye-tracking signals for the visual tasks were recorded for piloting purposes and will not be reported here.

2.2.3 Materials and Procedure

All perceptual tasks, the Stroop task, and the control tasks were presented with Visual Basics on a Dell Latitude E6520 computer (15" monitor with 1600 x 900 resolution). Participants were seated in front of the computer, their head approximately 60 cm from the screen. The sounds were presented binaurally (sounds transmitted at both ears) through headphones (Sennheiser, HD 518). Each participant initially adjusted the loudness of the sounds to a comfortable level, which was maintained constant during the experiment. The experimenter sat next to the participants during the entire duration of the experiment to start the presentation of each trial for each task. If needed, the experimenter reminded participants not to move their head during the tasks.

2.2.3.1 Perceptual tasks.

In the *Ambiguous figure (AF)* task (high-level visual task; Figure 2.1, upper left panel), the stimulus was an ambiguous line drawing (7.7 x 5 cm) which could be interpreted as either a duck or a rabbit. The stimulus subtended 4.81×7.2 (v x h) degrees of visual angle. For disambiguated versions, the body of the duck on a lake with another duck in the background or the body of a rabbit with a carrot was added to the ambiguous head.

In the *Ambiguous motion (AM)* task (low-level visual task; Figure 2.1, lower left panel), the stimulus was a revolving cylinder composed of two transparent layers of 200 randomly positioned white dots over a black background, moving in opposite directions with a sinusoidal speed profile (Klink, van Ee, Nijs, et al., 2008). The cylinder subtended 5.11×6.15 (v \times h) degrees of visual angle. A single dot was 12.29×12.35 arcmin (v \times h) and moved with a peak angular speed of 6.96 degrees/second. Disambiguated versions were created by dimming the luminance of half of the dots on the back of the sphere.

In the *Verbal transformations (VT)* task (high-level auditory task; Figure 2.1, upper right panel) the stimulus was the word “fly”, recorded by a native female English speaker (26 years old). Word duration was 540 ms, mean pitch 191.3 Hz (see stimulus in the Appendix B). To create the ambiguous stimulus, the word was looped without pause 112 times. In the disambiguated versions of the sounds, the word “life” and the word “fly” were presented one time each.

In the *Auditory streaming (AS)* task (low-level auditory task; Figure 2.1, lower right panel), the stimulus was a sequence of a repeating low-high-low_ pattern (LHL—), where “L” and “H” were complex sounds and “—” indicate a silence with the same duration as the sounds (van Noorden, 1975). The sound “L” was a recording of a water droplet hitting a glass (a wine glass), and the sound “H” was a recording of a water droplet hitting ceramic (chinese cup). The frequency difference between the two sounds was 30Hz, with a mean pitch of L at 402 Hz (min pitch: 397 Hz, max pitch: 408 Hz) and that of H set at 430 Hz (min pitch: 428 Hz, max pitch: 431Hz). The stimulus onset asynchrony (SOA, onset to onset time interval) was 150 milliseconds. The sounds were adjusted for each participant to a comfortable level and kept constant throughout

the experiment (range 30-45dB). The sound sequences were presented in five testing blocks of 60 seconds.

Participants were instructed to respond *Integrated* if they perceived all the tones as belonging together or *Segregated* if they perceived the sound as separating into two streams of sound. These interpretations were demonstrated to participants using disambiguated examples. For the *Integrated* example, participants saw a visualisation depicting one dripping tap with three water droplets falling from the tap in synchrony with the “LHL—” sounds. The *Segregated* percept was demonstrated with a visualisation of two dripping taps while a 19dB intensity difference was created between the “L” and “H” sounds (L =76 dB, H =57dB).

The spectrograms of the stimuli used for the verbal transformations and auditory streaming task can be found in Appendix A and Appendix B.

2.2.3.2 Experimental phases.

For each perceptual task, there was a training and a testing phase (see Figure 2.1).

Training phase. Training started with a 15-second presentation of the ambiguous stimulus (1), after which participants were asked about what they saw/heard. Next, participants were presented with the disambiguated versions (2) followed by the presentation of the ambiguous stimulus (3) and indications about what to report in the testing phase. In the Auditory streaming task there was an extra training phase after (3) where participants had to close their eyes and report what they heard (i.e., one or two taps).

Testing phase. Each perceptual task was presented five times (in five blocks) for 60 seconds. Participants were instructed to keep the buttons pressed down as long as

they perceived one of the interpretations and to switch to the other button as soon as their perception changed. The state of the two buttons was recorded continuously. Reporting was made through two interface button boxes with attached pictures of the disambiguated interpretations: one box on the right and another on the left side of the table. The two interface boxes had attached pictures representing the disambiguated interpretations. The position of the two boxes with their assigned interpretation of the stimulus was randomised between participants except the ambiguous motion task that was kept constant (left/right). For each stimulus block, the *initial reaction time*, the *1st phase duration*, and the *number of switches* were recorded.

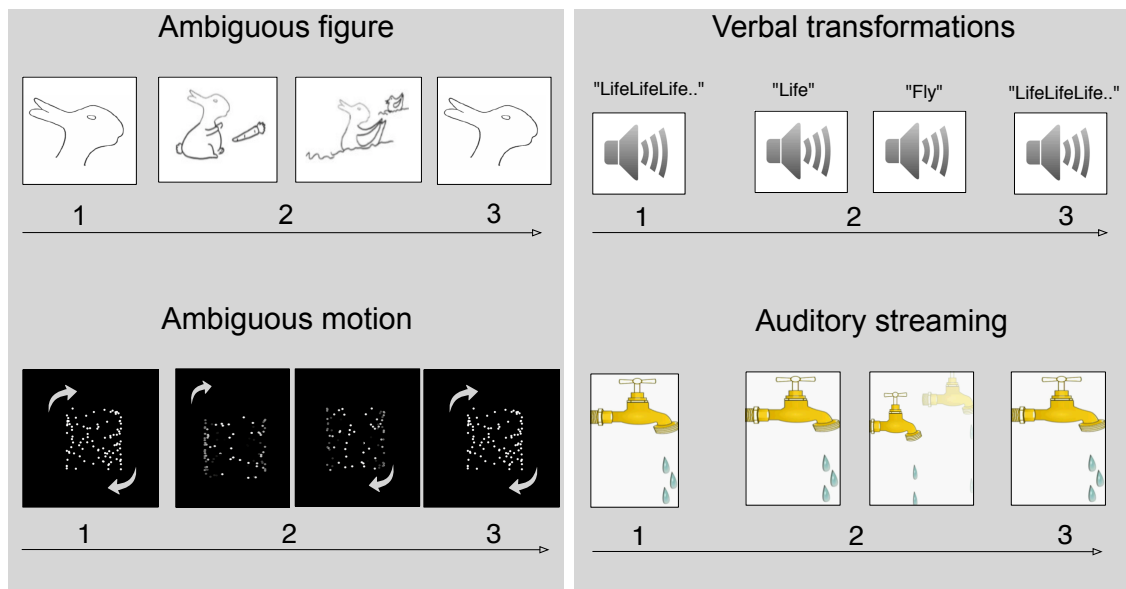


Figure 2.1. Training stages for each perceptual task. 1) initial interpretation: participants were instructed to tell what they saw/heard after an initial presentation of the ambiguous stimulus; 2) disambiguation phase: the alternative interpretation of the stimulus was introduced to the participants; 3) the ambiguous stimulus was presented again and participants were instructed that they needed to report their perception continuously for 60 seconds in Experiment 1 and 120 seconds in Experiment 2.

2.2.3.3 Control tasks.

One visual and one auditory control task were administered to ensure that participants understood the tasks and followed instructions.

Visual control task. Participants were instructed to look at an unambiguous picture (Figure 2.2 left side) for 60 seconds and a morphing animation (Figure 2.2 right side) for another 60 seconds. They were instructed to press a key on the keyboard whenever they saw the picture changing. Participants were excluded from the analysis if they reported that they saw a change in the picture with the girl (there was no physical change here) or if they did not report any change in the morphing animation (which was changing). The order of the two control tasks was randomised across participants. Two participants were removed from the final analyses for failing this task (see Results).

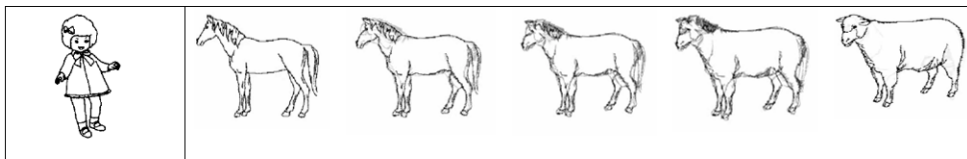


Figure 2.2. Unambiguous picture of a girl (left side); Animation of a horse morphing into sheep (right side).

Auditory control. Participants listened to a sequence of sounds for 60 seconds. One 30-second segment displayed the sound “L” described above, while the second 30-second part displayed the sounds “L” and “H”. Participants had to indicate whether they perceived one sequence of drips (as one dripping tap) or two sequences of drips (as two dripping taps). The identity and time of all button presses were recorded. Participants had to report the change in the sound sequence and choose the right patterns to be included in the analyses.

2.2.3.4 Individual differences measures.

Stroop. Inhibition was measured using a Day-Night Stroop task (Simpson & Riggs, 2005). Two pictures with a day and a night scenario were presented. Participants were instructed to press a dark blue button in response to the day scenario and a yellow button in response to the night scenario (see Figure 2.3.a). Participants had a pre-test phase where they were shown each of the two pictures once and asked to complete the task as quick as possible. During the test phase, 16 pictures were presented in a pseudo-random order -DNNDNDDNNDNDDNDN (8 day and 8 night pictures), and participants were asked to respond as accurately and quickly as possible. Each picture was presented for a maximum of eight seconds with a two-second inter-trial interval. Accuracy and reaction times were recorded. In the analysis, only the reaction times (Stroop RTs) for the correct answers were analysed.

Stroop Memory Control. A neutral task was used to measure basic working memory components of the Stroop task (i.e., remembering what button to press for which picture). The procedure was the same as for the Day-Night Stroop task, but instead of the day and night pictures, participants were presented with two abstract images (Figure 2.3a, last two pictures on the left side) and asked to press the associated dark blue or yellow button. To compare the performance from Stroop and memory control, an inhibition effect (Inhibition RTs) was computed by subtracting memory control RT from Stroop RT.

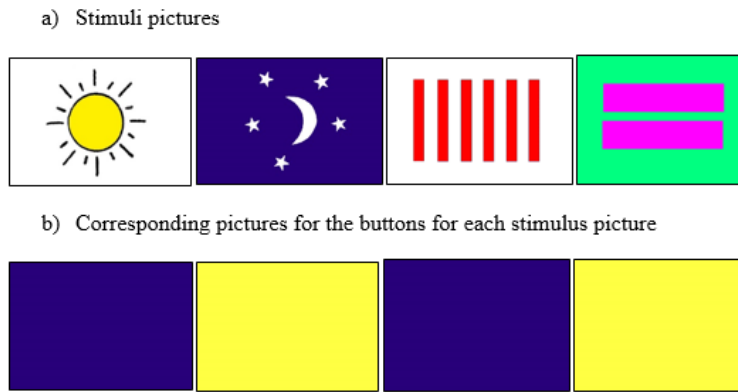


Figure 2.3. Images used in the Stroop task and the related memory control task (2.3.a). The corresponding colours for the buttons are in the second row (2.3.b).

Verbal Fluency. Set-shifting was measured with a verbal fluency task with the same procedure from Troyer et al. (1997). Participants had to name either as many animals they could think of (for semantic fluency) in 60 seconds or words that start with the letters F, A, and S (for phonetic fluency). The four blocks were randomised. Responses were recorded using a digital voice recorder. The total number of words produced was listed, and clusters of words belonging to the same category were determined. Switches between clusters were then determined. The main variables of interest were the total numbers of words produced (VF total) and the number of switches between clusters (VF switches).

Pattern meanings task. Creativity was assessed using the Pattern meaning task of M. A. Wallach and Kogan (1965), which is a divergent thinking test. The test comprised eight abstract patterns presented individually on 20×14 cm laminated cards (see example Figure 2.4). Participants were asked to describe “all the things you think it could be or that it reminds you of”. The main variables of interest were: the number of unique responses (cUnique) defined as responses given by less than 1% of the participants; the number of unusual responses (cUnusual) provided by less than 20% of the participants, and the total number of responses produced (cTotal).



Figure 2.4. An example of a pattern that participants were asked to describe in the Pattern meaning task.

Flexible Item Selection Task (FIST). Cognitive flexibility was measured with the FIST task. This task was adapted from Dick (2014) and Jacques and Zelazo (2001). The stimuli were presented on a touch screen laptop (Lenovo ThinkPad X220 Tablet, 12.5-inch, 1366 x 768). The task lasted approximately 20 minutes. Stimulus objects were contained in “cards” on a white background with three or four stimulus cards being presented at a time, one under another in the middle of the screen. Each set of cards contained objects that were derived from the combination of four dimensions (colour, number, size, and shape). Each dimension varied along three attributes: Colour (red, green, blue); Number (one, two, three); Size (large, 2.2 x 2.4 cm, medium, 0.83 x 1.2 cm, small, 0.83 x 0.6 cm); Shape (dog, flower, boat). See Figure 2.5 for an example of a set of cards. Participants were told they would see some cards with pictures, which they needed to match according to some given criteria. Each participant received in total four sets of six trials each. In each set, the number of possible matching selections was different; two, three, four, or six matches were possible from a set of three or four cards. Within each trial, a correct response consisted of selecting all the possible matching cards. The percentage of correct responses and median reaction times of the correct responses for each trial were recorded. From these measures, the variables of interest were: the overall accuracy (proportion of correct responses), the overall reaction time, and the total number of errors.

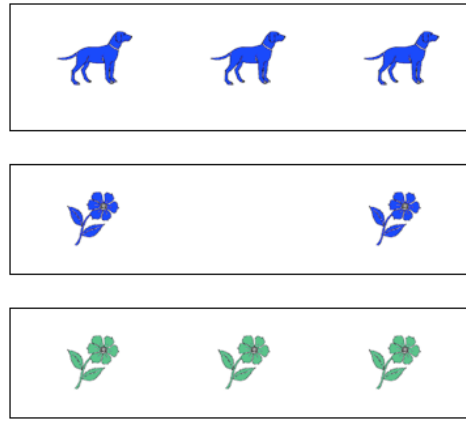


Figure 2.5. Example of a set of cards in the FIST task. The three cards contain objects in two shapes (flowers and dogs) and two colours (green and blue) and similar size (large). There are maximum three correct matches in this example: first and last card have the same number, first and second have similar colours, second and last have the same object.

2.3 Results

2.3.1 Data Analyses

The main dependent variables analysed for the perceptual tasks were the initial reaction time, the first phase durations, and the switching rates. The initial reaction time was analysed to see how fast participants made their initial perceptual decision, while the first phase duration investigated the length of the first interpretation and how much time it took to switch to the second interpretation. This is referred to as inertia of the first phase duration (Hupé & Pressnitzer, 2012). To examine the perceptual dynamics in the four tasks, switching rates were computed for each block by dividing the number of switches by the length of a block (60 seconds) after the first perceptual phase was excluded. Mixed model analyses were conducted (see the full rationale in Section 1.6) with Modality and Level as fixed factors and participants as a random factor on each of the three dependent variables.

Spearman rank-order correlations were calculated separately for the three dependent variables to assess whether perceptual switching in the four tasks relate to each other and to the individual differences in executive function, cognitive flexibility and creativity.

2.3.2 Participants Exclusion

To control for false positive/negative answers, participants that failed in the control tasks were excluded from further analyses. In total, there were 20 participants excluded. In the first group, eight participants were excluded (seven failed the visual control task and one participant that never switched in any of the two perceptual tasks). In the second group, nine participants were excluded (six of them failed the visual control task and three failed the auditory control task). In the third group, two participants were excluded because they failed the visual control task. In the fourth group, one participant failed the auditory control task and was excluded. The final sample size was 91 participants (from 111 tested).

2.3.3 Preliminary Analyses

Thirty percent of the switching rates had values of zero as 34 participants did not switch in some of the blocks and tasks (predominantly in the auditory streaming task; see Table 2.2). These values were treated as missing values and omitted from the analyses. The exact same analyses we present here, but with the missing values considered as zero, can be found in Appendix C. The results from Appendix C are very similar to what is presented in this chapter.

Before the main analyses were conducted there were two separate analyses investigating: (1) whether the first and subsequent phase durations are different, and (2) whether there is an effect of block number (i.e., the order of the viewing trials, from 1 to 5 blocks) on the switching rates.

2.3.3.1 1st vs. subsequent phase durations.

It was explored whether the first phase duration and the subsequent durations were significantly different. A Linear mixed effects model (LMEM) on the phase durations showed that there is a significant difference between 1st phase durations and the median subsequent phase durations. The first phase duration ($M = 19.91$, $SD = 17.54$) was significantly higher than the subsequent phase durations ($M = 5.50$, $SD = 6.77$), $F(1, 316) = 184.94$, $p < .001$.

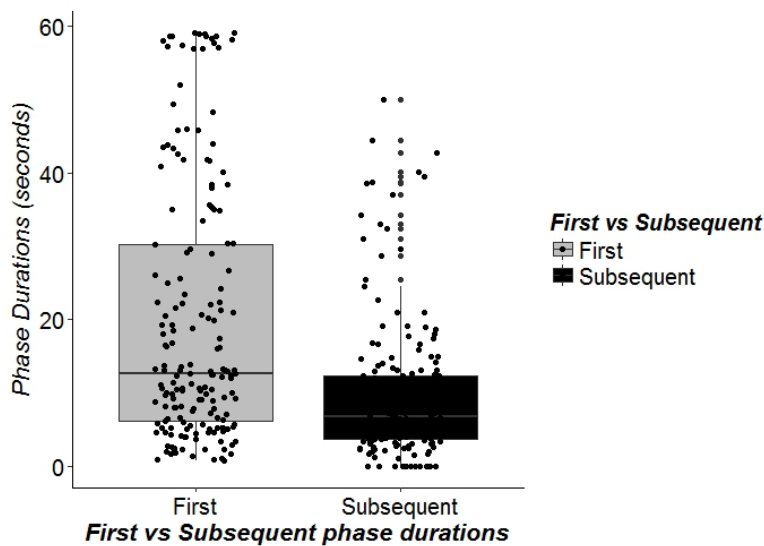


Figure 2.6. Median first phase duration vs. the subsequent phase durations. The box represents the 25th and 75th percentile; data points over the boxes indicate data range.

2.3.3.2 Switching rates across blocks.

Recent findings show that perceptual switching in the first block is significantly different from the rest of the blocks, and that excluding it from analyses increases the intra-individual consistency (Farkas et al., 2016). Before the main analyses were conducted, a LMEM was performed to assess whether switching rates differ across blocks (see descriptive statistics in Table 2.1). The results showed that there was an effect of Block on the switching rates, $F(4, 437) = 2.98$, $p = 0.019$. Planned comparisons

using Dunnett test⁶ between block 1 and each of the subsequent blocks revealed that there was a significant difference between block 1 and 3 ($p = 0.04$) and between block 1 and block 5 ($p = 0.03$). These results support previous findings by showing that the switching rates in the first block are significantly different from the subsequent blocks. Because of these differences in switching rates between blocks, block 1 was excluded from the main analyses.

Table 2.1

Means and Standard Errors for the Switching Rates in each Block

	Block1	Block2	Block3	Block4	Block5
Mean	.10	.11	.13	.13	.13
SD	.07	.10	.10	.11	.11

2.3.4 Distributions of Subsequent Switching Rates

The visual inspection of the phase durations in the histograms in Figure 2.7 displays the expected lognormal distributions for what was previously reported as a critical characteristic of multistability (Leopold & Logothetis, 1999; Pressnitzer & Hupé, 2006). Note however that the distribution of the switching rates in auditory streaming is flatter than the rest of the tasks, showing a slightly different distribution. In order to characterise more accurately the distributions of phase durations, a larger number of observations is needed than we had collected here (Pressnitzer & Hupé, 2006).

⁶ This test was used because it is more suitable for testing planned comparisons than Bonferroni.

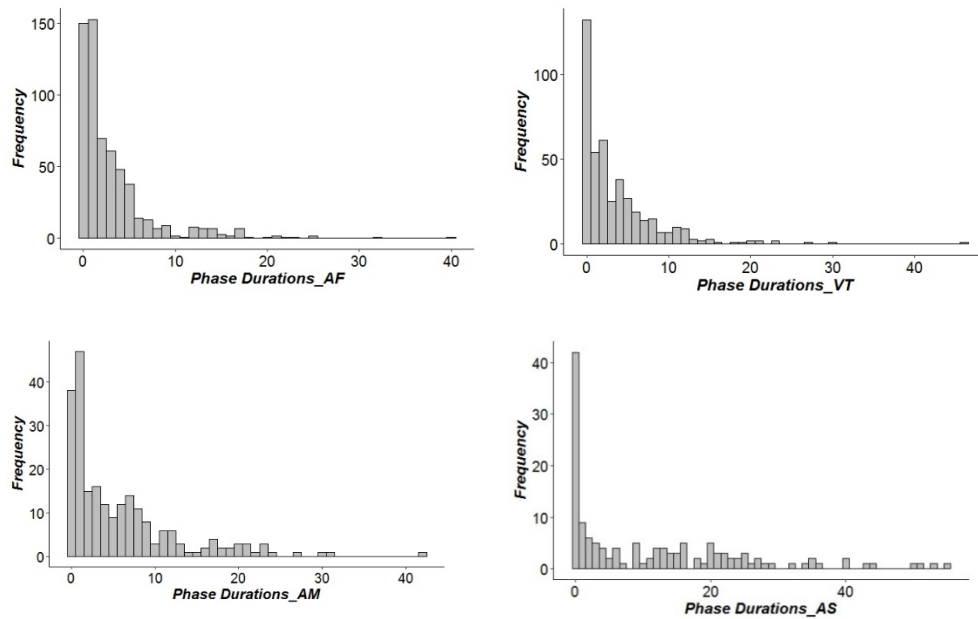


Figure 2.7. Histograms of phase durations for each of the ambiguous stimuli. The x-axis displays the phase durations. The y-axis displays the frequency of the phase durations that fall within a specific interval.

2.3.5 Balance between the Alternative Interpretations

The switching rate is known to be dependent on the balance between the two perceptual alternatives (Moreno-Bote et al., 2010). Therefore, this was assessed using the following formula: $|\text{proportion1} - \text{proportion2}|$, where proportion1 and proportion2 is the percentage of time each interpretation is perceived. A complete balance would result in zero, while perceiving one alternative for the entire block duration would yield one. The results showed that the reports were reasonably balanced for ambiguous figure ($= 0.19$), ambiguous motion ($= .09$), and auditory streaming ($= .15$), while less balanced for verbal transformations ($= .55$). Table 2.2 summarises the average percentage of time each of the alternative interpretations were maintained, plus the “none” response (i.e., when no key was pressed). The percentage of participants that reported one of the interpretations as their first answer is also summarised in Table 2.2. Overall, the

percentage of time for “none” responses was below 9% for each of the tasks, showing that participants had perceived one or the other alternative most of the time.

Table 2.2

The proportion of Time for each Percept (including None Presses), for Each of the Four Perceptual Tasks. Percentage of Participants that selected Percept 1 as a first Interpretation is summarised in the last column

Task	Percept 1 Mean (SD)	Percept 2 Mean (SD)	None Mean (SD)	Percept 1 as first interpretation
AF	.36 (.13)	.55 (.15)	.09 (.07)	36.9%
AM	.43 (.25)	.52 (.25)	.06 (.03)	49%
VT	.74 (.15)	.19 (.13)	.07 (.03)	100%
AS	.38 (.18)	.54 (.19)	.08 (.05)	81.25%

Note. Percept 1 and Percept 2 in each of the 4 tasks were as follow: AF: Percept 1 = duck, Percept 2 = rabbit; AM: Percept 1 = Clockwise, Percept 2 = anticlockwise; VT: Percept 1 = fly, Percept 2 = life, AS: Percept 1 = Integrated; Percept 2 = Segregated.

2.3.6 Effects of Modality and Level of Processing on Perceptual Switching

2.3.6.1 Initial reaction time.

The initial reaction time performance across the four perceptual tasks is summarised in Table 2.3 and in Figure 2.8.

Table 2.3

Initial Reaction Times Performance across the Four Perceptual Tasks

	Task	MaxObs	FinalObs	Min	Max	Mean	SD	SEM
Initial RTs	AF	184	182	0.05	9.55	2.02	1.21	.090
	AM	180	178	0.04	9.88	1.80	1.09	.082
	VT	172	172	0.12	3.30	1.45	0.68	.052
	AS	192	192	0.76	20.98	3.93	3.36	.236

Note. MaxObs: the maximum number of observations per task; FinalObs: the final number of observations after omitting missing values.

Participants responded faster in the visual tasks compared to the auditory tasks, as indicated by the effect of Modality, $F(1, 631) = 42.38, p < .001$. The initial reaction

time in the high-level tasks was shorter than in the low-level tasks, as evidenced by an effect of Level, $F(1, 714) = 46.36, p < .001$. There was an interaction between Modality and Level, $F(1, 704) = 104.97, p < .001$. Pairwise comparisons showed that in the visual modality, there was a longer reaction time in the ambiguous figure (high-level task) compared to ambiguous motion (low-level task), $t(714) = 2.59, p = 0.011$; while in the auditory modality, participants responded faster to the verbal transformations (high-level task) than to the auditory streaming (low-level task), $t(712) = 12.18, p < .001$. Secondly, the impact of the level of processing was investigated between modalities (high-level vision vs. high-level auditory and low-level vision vs. low-level auditory). The results showed that the initial reaction time in the ambiguous figure was longer than in verbal transformations, $t(722) = 4.174, p < .001$, while the initial reaction time was significantly longer in the auditory streaming compared to ambiguous motion, $t(722) = -12.32, p < .001$.

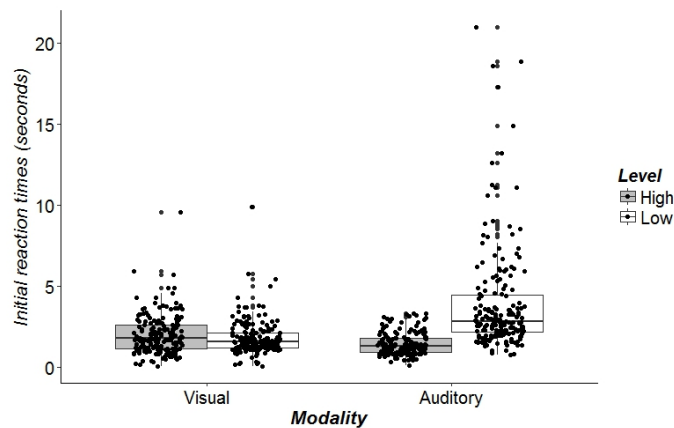


Figure 2.8. Initial reaction times in each perceptual task. Box represents the 25th and 75th percentile; data points over the boxes indicate data range.

2.3.6.2 First phase duration.

The first phase duration performance across the four perceptual tasks is summarised in Table 2.4 and in Figure 2.9.

Table 2.4

First Phase Durations Performance across the Four Perceptual Tasks

	Task	MaxObs	FinalObs	Min	Max	Mean	SD	SEM
1 st Phase Duration	AF	184	153	0.02	43.10	7.10	7.19	0.58
	AM	180	133	0.01	59.10	15.43	13.08	1.13
	VT	172	136	0.14	59.87	19.37	17.92	1.53
	AS	192	87	0.02	41.61	8.84	8.29	0.89

Note. MaxObs: the maximum number of observations per task; FinalObs: the final number of observations after omitting missing values.

Participants' first phase duration was longer in the auditory tasks than in the visual tasks, as indicated by the effect of Modality, $F(1, 461) = 4.54, p = .034$. There was no effect of Level, $F(1, 478) = 1.06, p = .302$. However, there was an interaction, $F(1, 449) = 35.01, p < .001$. Firstly, there was a significant difference between the two levels within each modality (i.e., high-level vision vs. low-level vision and high-level auditory vs. low-level auditory). In the auditory modality, the first phase duration was longer for verbal transformations (high-level task) than for auditory streaming (low-level task), $t(489) = 3.389, p = .002$. In the visual modality, the first phase duration was longer in ambiguous motion (low-level task) than in ambiguous figure (high-level task), $t(416) = -5.257, p < .001$. Secondly, the impact of Level was investigated between modalities (high-level vision vs. high-level auditory and low-level vision vs. low-level auditory). Results showed that the first phase duration in the ambiguous figure was shorter than in the verbal transformations, $t(492) = -6.399, p < .001$, while the first phase duration in the ambiguous motion was longer than for the auditory streaming, $t(504) = 3.109, p = .003$.

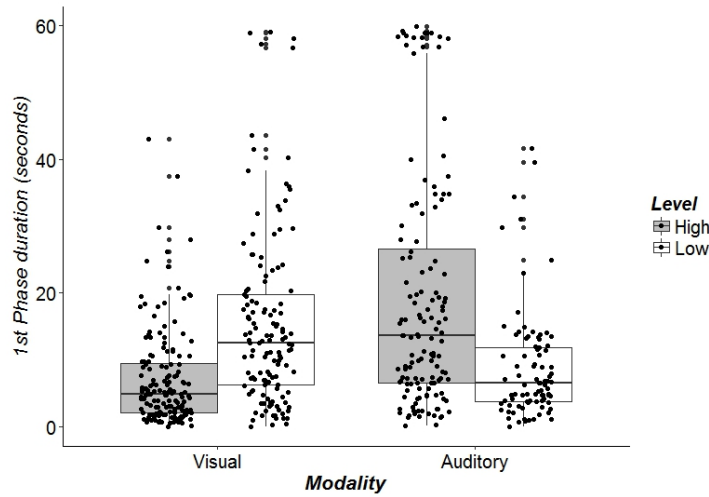


Figure 2.9. First phase duration in each perceptual task. Box represents the 25th and 75th percentile; data points over the boxes indicate data range.

2.3.6.3 Switching rates.

The switching rates across the four perceptual tasks is summarised in Table 2.5 and in Figure 2.10.

Table 2.5

Switching Rates Performance across the Four Perceptual Tasks

	Task	MaxObs	FinalObs	Min	Max	Mean	SD	SEM
Switching Rate	AF	184	153	0.02	0.72	0.16	0.13	.01
	AM	180	133	0.02	0.47	0.10	0.07	.01
	VT	172	136	0.02	0.55	0.17	0.10	.01
	AS	192	87	0.02	0.11	0.04	0.02	.002

Note. MaxObs: the maximum number of observations per task; FinalObs: the final number of observations after omitting missing values.

Participants had a higher switching rate in the auditory tasks than in the visual tasks as indicated by the effect of Modality, $F(1,454) = 20.05, p < .001$. There was a significant effect of Level, $F(1, 500) = 154.88, p < .001$. The switching rate was higher in the high-level tasks than in the low-level tasks. However, there was an interaction between Modality and Level, $F(1, 485) = 30.62, p < .001$. Pairwise tests indicated a difference between the two levels within each modality (i.e., high vision vs. low vision and high auditory vs. low auditory). In the auditory modality, the switching rate was

higher in verbal transformations (high-level task) than in the auditory streaming (low-level task), $t(504) = 12.187, p < .001$. In the visual modality, the switching rate was higher in ambiguous figure (high-level task) than in ambiguous motion (low-level task), $t(462) = 4.48, p < .001$. The results showed a Level effect across modalities (high vision vs. high auditory and low vision vs. low auditory). The switching rates between the visual and auditory high-level tasks were not significantly different $t(504) = -1.748, p = 0.081$. On the other hand, the switching rate in the ambiguous motion (low-level visual task) was significantly higher than in auditory streaming (the low-level auditory task), $t(504) = 6.742, p < .001$.

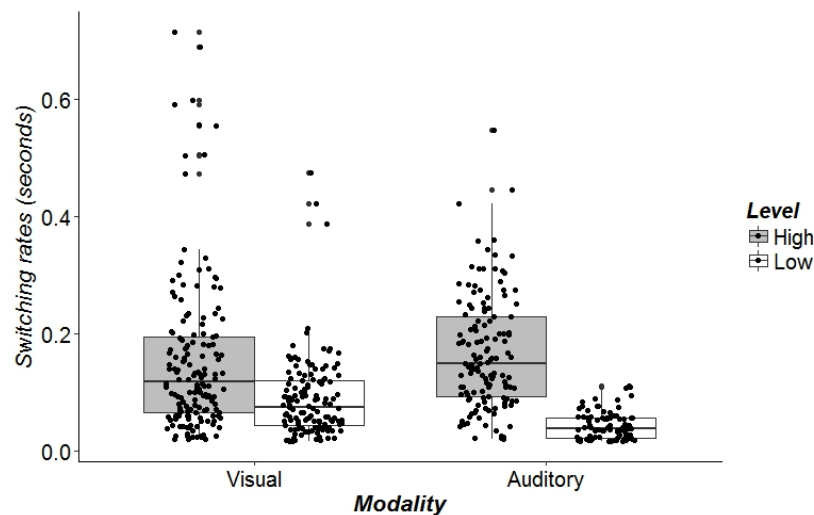


Figure 2.10. The switching rates in each perceptual task. Box represents the 25th and 75th percentile; data range is indicated by data points over the boxes.

2.3.7 Inhibition, Set-shifting, Cognitive Flexibility and Creativity

A summary of performances for Day-Night Stroop, FIST, Pattern Meanings test, and Verbal fluency is presented in Table 2.6.

Table 2.6

The Overall Performance for Individual Differences Measures

		N	Minimum	Maximum	Mean	SD
Stroop	Stroop RTs	91	.51	1.45	.82	.23
	Memory control RTs	91	.44	1.56	.66	.19
	Inhibition RTs	91	-.44	.61	.15	.15
FIST	Accuracy	91	.49	.97	.81	.09
	Reaction times	91	4.16	12.58	6.57	1.43
	No. of errors	91	3	52	21.95	9.98
Pattern meanings	cTotal	91	12	65	32.44	10.99
	cUnique	91	0	34	8.66	5.99
	cUnusual	91	0	25	11.10	5.10
Verbal fluency	No. of responses	90	5.17	19.50	10.36	2.47
	No. of switches	90	8.67	27.83	17.91	3.47

Note. Reaction times (RT) are displayed in seconds; FIST accuracy shows the percentage of correct responses on all trials. Inhibition effect (Inhibition RT) was calculated by subtracting reaction times in the Stroop memory control task from Stroop task.

2.3.7.1 Inhibition effect.

Performance accuracy (percentage of correct responses) on Stroop was very high (98.03% correct). Performance accuracy on Stroop memory control was also very high (average 99.64% correct). As performance accuracy was so high, the analyses were focused on reaction times (RTs) on the following variables: Stroop RTs, Stroop memory control RTs and inhibition RTs (Stroop RTs minus Stroop memory control RTs). A paired t-test revealed that memory control RTs were significantly shorter ($M = 0.57s$, $SD = .13$) than Stroop RTs ($M = .65s$, $SD = .17$), $t(110) = 10.78$, $p < .001$).

2.3.7.2 Pattern meaning tasks and Verbal fluency reliability

Subsets of data from the Pattern meaning tasks and verbal fluency task were coded by two independent raters.

Pattern meaning task. One of the pictures selected randomly from group 3 was coded by two raters. A two-way mixed intraclass correlation analysis on the unique

number of responses showed that there was a good agreement between the two raters, $ICC = .882$ (95% CI , .727 to .949), $p < .001$. A substantial agreement between the two raters was also for the number of unusual responses, $ICC = .548$ (95% CI , -.044 to .805), $p = .031$.

Verbal fluency. A two-way mixed intraclass correlation analysis on the number of switches showed that there was an excellent agreement between the two raters, $ICC = .896$ (95% CI , .842 to .931), $p < .001$.

2.3.8 Relationships between Perceptual Switching Variables across Perceptual Tasks

Correlations were investigated between perceptual switching (initial reaction time, 1st phase duration, and switching rates) across the two perceptual tasks in each experimental group (Table 2.7). The initial reaction times in ambiguous figure and verbal transformations correlated positively. Similarly, the initial reaction times in ambiguous motion and verbal transformations correlated positively. No other associations were found (Table 2.7).

Table 2.7

Commonalities between Perceptual Switching across the Four Perceptual Tasks

Group	AF & VT	AM & AS	AM & VT	AF & AS
Initial RT				
S rho	.439*	.051	.426*	.357
N	21	23	22	25
1stPhase Durations				
S rho	.397	-.300	.177	.188
N	16	15	20	16
Switching rate				
S rho	.444	.425	.355	-.053
N	16	17	20	16

Note. S rho = Spearman test rho.

2.3.9 Relationships between Perceptual Switching Variables within the Same Perceptual Task

Data were further explored to investigate the relationships between initial reaction time, first phase duration, and switching rate within each task. Two possible scenarios were possible: 1) the initial reaction time is positively related to the first phase duration and negatively related to the switching rate; or 2) initial reaction time correlates negatively with 1st phase duration and positively with the switching rate. The results (Table 2.8) support the former scenario. In all tasks, the first phase duration correlated negatively with the switching rate. The initial reaction time in auditory streaming correlated positively with the first phase duration.

Table 2.8

Correlations for Perceptual Switching Variables (Initial Reaction Time, First Phase Duration, and Switching Rate) within Each Perceptual Task

		AF		AM		VT		AS	
		1PhD	SR	1PhD	SR	1PhD	SR	1PhD	SR
Initial	S rho	.129	-.253	.101	-.231	.068	-.270	.413*	-.148
RT	N	43	43	42	42	37	37	33	33
1stPhD	S rho	--	-.856***	--	-.482***	--	-.500**	--	-.596***
	N	--	43	--	42	--	37	--	33

Note. S rho = Spearman test rho, Initial RT = Initial reaction time, 1PhD = first phase duration, SR = Switching rates.

2.3.10 Relationships Between Perceptual Switching, Inhibition, Cognitive Flexibility, Set-Shifting, And Creativity.

Correlations between the perceptual variables and the individual differences scores were run separately for each perceptual task (Tables 2.9-2.12).

The initial reaction time in ambiguous figure correlated with Stroop RTs and memory control RTs. The first phase duration in ambiguous motion correlated negatively with

the accuracy in the FIST task and positively with number of errors in the FIST task. The switching rate in ambiguous motion correlated positively with inhibition RTs.

The switching rates in verbal transformations correlated negatively with the number of solutions in the creativity task. The first phase duration in verbal transformations also correlated positively with both the number of creative solutions and the number of unique responses in the creativity task. Perceptual switching variables in auditory streaming did not correlate with any of the individual measures tasks.

Table 2.9

Perceptual Switching and Individual Differences in Ambiguous Figure

		Stroop			FIST			Creativity			Verbal Fluency	
		Stroop RTs	Memory Control RTs	Inhibition RTs	Accuracy	RT	Errors	cTotal	cUnique	cUnusual	VFswitching	VFtotal
Initial	rho	.401**	.352*	-.035	-.237	.207	.239	-.022	.017	-.025	-.137	-.080
Reaction	N	46	46	46	46	46	46	46	46	46	46	46
time												
1 st Phase	rho	.228	.138	-.014	-.054	-.020	-.028	.049	-.251	.011	.057	-.019
Duration	N	43	43	43	43	43	43	43	43	43	43	43
Switching	rho	-.294	-.279	.037	-.051	-.089	.114	.015	.283	-.025	-.066	-.056
Rate	N	43	43	43	43	43	43	43	43	43	43	43

Table 2.10

Perceptual Switching and Individual Differences in Ambiguous Structure-from-Motion

		Stroop			FIST			Creativity			Verbal Fluency	
		Stroop RT	Memory RT	Inhibition RT	Accuracy	RT	Errors	cTotal	cUnique	cUnusual	VFswitching	VFtotal
Initial	rho	.271	.115	.364*	-.056	.244	.048	-.177	-.188	-.107	-.013	-.142
Reaction	N	45	45	45	45	45	45	45	45	45	44	44
time												
1 st Phase	rho	-.078	-.088	.056	-.313*	.035	.398*	-.217	-.086	-.231	-.071	-.123
Duration	N	42	42	42	42	42	42	42	42	42	41	41
Switching	rho	-.125	-.096	-.139	.135	.014	-.158	-.179	-.100	-.290	-.179	-.045
Rate	N	42	42	42	42	42	42	42	42	42	41	41

Table 2.11

Perceptual Switching and Individual Differences in Verbal Transformations

		Stroop			FIST			Creativity			Verbal Fluency	
		Stroop RT	Memory RT	Inhibition RT	Accuracy	RT	Errors	cTotal	cUnique	cUnusual	VFswitching	VFtotal
Initial	rho	.371*	.338*	.214	.113	.146	-.068	-.019	.117	-.158	-.039	.095
Reaction	N	43	43	43	43	43	43	43	43	43	43	43
time												
1 st Phase	rho	-.159	-.224	-.158	-.042	-.057	.029	.427**	.351*	.254	-.049	-.038
Duration	N	37	37	37	37	37	37	37	37	37	37	37
Switching	rho	-.078	-.031	-.102	.265	.182	-.355*	-.144	-.021	-.157	.181	.045
Rate	N	37	37	37	37	37	37	37	37	37	37	37

Table 2.12

Perceptual Switching and Individual Differences in Auditory Streaming

		Stroop			FIST			Creativity			Verbal Fluency	
		Stroop RT	Memory RT	Inhibition RT	Accuracy	RT	Errors	cTotal	cUnique	cUnusual	VFswitching	VFtotal
Initial	rho	.206	.188	-.070	.111	.184	-.157	-.087	-.100	-.017	.065	-.053
Reaction	N	48	48	48	48	48	48	48	48	48	47	47
time												
1 st Phase	rho	.080	.032	.005	.228	.194	-.203	-.037	-.091	.062	.187	.303
Duration	N	33	33	33	33	33	33	33	33	33	32	32
Switching	rho	.104	.055	.071	-.209	-.219	.208	.076	-.001	.053	-.253	-.102
Rate	N	33	33	33	33	33	33	33	33	33	32	32

2.4 Discussion Experiment 1

2.4.1 Effects Of Modality and Level Of Processing on Perceptual Switching

The main aim of this study was to investigate to what extent the perception of multistable stimuli configurations has modality- and representation-level-specific or general features. For this purpose, participants' perception was recorded continuously for two auditory and two visual ambiguous tasks with one low- and one high-level representation stimulus in each modality.

Firstly, there was a Modality effect across all three perceptual switching variables. The initial reaction time in the visual tasks was shorter than in the auditory tasks, suggesting that it was much easier to make an initial perceptual decision in the visual than in the auditory modality. There was also an effect of Modality on the first phase duration; once the first perceptual interpretation is chosen, it takes much longer to change to the subsequent interpretation in the auditory tasks than in the visual tasks. This is consistent with a previous finding of Hupé and Pressnitzer (2012) who compared the phase durations in auditory streaming and visual plaids. An effect of Modality was also found for the switching rates; the switching rate was smaller in the visual tasks compared to the auditory tasks. This result is in contradiction to results of Kondo et al. (2012) who found that the number of switches was larger in the visual tasks (Necker cube and visual plaids) than in the auditory tasks (verbal transformations and auditory streaming) (see page 1917, Table 1). However, the contradictory results might be due to the stimuli's parameters.

Secondly, there was a main effect of Level on the initial reaction time and the switching rate. The first perceptual decision was made faster in the high-level tasks than in the low-level tasks. Similarly, the switching rate was much higher in the high-level tasks than in the low-level tasks. This result supports the findings from visual research of Strüber and

Stadler (1999), van Ee et al. (2005), and Wolf and Hochstein (2011) who have found larger switching rates for stimuli representing complex objects and with semantic content than for less complex stimuli and without semantic content.

There were interactions between Modality and Level for all the perceptual switching variables. The initial reaction time is shorter in verbal transformations than auditory streaming, while in ambiguous figure and ambiguous motion it did not differ significantly. The first phase durations were shorter in ambiguous figure than ambiguous motion, and longer in verbal transformations than auditory streaming. The switching rate was higher in the high-level tasks in both visual and auditory modalities. Ambiguous figure had higher switching rates compared to ambiguous motion, and verbal transformations had higher switching rates than auditory streaming. Previous literature that compared the reaction times of visual and auditory stimuli suggests that reaction times for simple stimuli are faster for the auditory than the visual stimuli. For example, Thompson et al. (1992) showed that it takes approximately 180–200 ms to detect visual stimuli, whereas for sounds it is around 140–160 ms. Similarly, early research of Kemp (1973) shows that an auditory stimulus input takes 8–10 ms to reach the cortex while a visual stimulus input takes 20–40 ms. Based on these findings, it is unsurprising that first perceptual decision is much faster for verbal transformations than for ambiguous figure. The finding that perceptual decision in verbal transformations is the fastest could also be due to the fact that our cognitive system has extensive training with verbal content (i.e., people are engaged extensively in verbal communications). The finding that the perceptual decision in auditory streaming is the slowest could be explained by the physical characteristics of the stimulus. At least a few LHL triplets (one repetition lasts 450ms) must be experienced to reach an interpretation (Bregman, 1994), which lasts longer than the 540ms it takes for the verbal transformations to be played once, and which is enough for participants to discriminate the meaning of the word.

2.4.2 Commonalities Between Perceptual Switching In The Four Perceptual Tasks

The present study brings weak support for the central hypothesis theory (Carter & Pettigrew, 2003; Kondo et al., 2012) which states that perceptual switching in different multistable paradigms is modulated by a central brain mechanism. The initial reaction time in ambiguous figure and ambiguous motion correlated positively with the initial reaction time in verbal transformations. However, the correlations were not very strong, and their interpretation should be made with caution. Moreover, the correlations between the main variable of interest (i.e., the switching rates) across perceptual tasks were not significant. The results in the current study are in line with the findings of Pressnitzer and Hupé (2006) who did not find any correlations between visual and auditory switching rate.

2.4.3 Relationships Between Perceptual Switching And Individual Differences

The correlational analyses between the individual differences measures and perceptual switching in the four perceptual tasks did not show a consistent pattern of results. However, some isolated relations partially align with previous findings. For instance, the longer the initial reaction time in the ambiguous figure, the shorter was the Stroop RTs. This indicates that inhibition impacts the first perceptual decision in ambiguous figure. After showing that people need to inhibit the first interpretation to select the second interpretation, Wimmer and Doherty (2011) suggested that inhibition is necessary for the ability to switch. The current result suggests that inhibition might be required for the first perceptual interpretation in ambiguous figures.

The first phase duration in the ambiguous motion task was related negatively to cognitive flexibility measured with FIST, meaning that less cognitive flexibility is associated with a longer time to switch to the second interpretation in ambiguous motion. This shows that cognitive flexibility might be necessary for perceptual switching in ambiguous motion.

The switching rate in verbal transformations was related negatively to the number of errors in FIST. This result suggests that the more errors participants make in the FIST task (i.e., less cognitive flexibility), the fewer switching rates are reported in verbal transformations. Finally, the longer it takes to make the first perceptual decision in verbal transformations, the weaker the memory and the inhibition. Perceptual switching in auditory streaming did not correlate with any of the individual differences measures tasks.

2.4.4 Limitations

Moreno-Bote et al. (2010) showed that the switching rates reach a maximum at equi-dominance (where the alternative interpretations of the stimuli are equally perceived). The balance analyses indicate that the verbal transformations task was biased towards the “life” interpretation. This could have had an impact on the results presented here. Overall, verbal transformations had the largest switching rate, followed closely by ambiguous figure. The switching rate is expected to increase further in case the stimulus is more balanced. This aspect was further investigated in Experiment 2 of this study. A possible other limitation of this study might arise from the fact that the experimental design was not fully contra-balanced. This could have reduced the power effects (Maxwell & Delaney, 2004). Therefore, in the second experiment, the question of domain/level generality/specificity was investigated further in a repeated measure design study, where each participant carried out all the perceptual switching tasks.

Another aspect that might have influenced the results of the present study is the length of each experimental block. In both auditory research (see Denham et al., 2012; Denham et al., 2014) and visual research (see van Ee et al., 2005) block durations were larger than three minutes. Considering that longer duration time is required for perceptual switching to stabilise (Suzuki & Grabowecky, 2007), this aspect will be addressed in Experiment 2.

Experiment 2

2.5 Introduction Experiment 2

Experiment 2 investigates further the perceptual switching in visual and auditory modalities as well as perceptual switching between lower- and higher-cognitive level demands. Contrary to the previous experiment, the current one has a fully crossed design; each participant received all the four perceptual tasks described in Experiment 1 in a single session. Because the balance between the two interpretations in verbal transformations was not very good in Experiment 1, the stimulus was modified. The stimulation time was increased to 120 seconds, as a higher viewing/listening time was found necessary for stabilising the phase durations (Denham et al., 2013). Farkas et al. (2016) reported for the first time a relationship between perceptual switching in an auditory streaming and a personality measure, namely ego-resiliency (see Section 1.4.3). For exploratory reasons, we included this measure to see to what extent the results generalise to different multistable paradigms.

2.6 Method

2.6.1 Participants

Thirty-five fluent English speakers (eight males) were recruited online via the university recruiting system and received course credits or £10 for participation in the experiment, $M_{age} = 21.57$, $SD = 5.13$. All participants reported corrected-to-normal vision and normal hearing. Informed written consent was obtained from each participant, and the University of Plymouth Ethics Committee approved the study.

2.6.2 Design

Each participant carried out the same four perceptual tasks described in Experiment 1 (see Section 2.2.3). Only the Stroop task was included from the individual differences tasks described in Experiment 1, plus a short ego-resiliency personality questionnaire. Eye-tracking signals for the visual tasks were recorded for piloting purposes and will not be reported here.

2.6.3 Materials and Procedure

The procedure was identical with Experiment 1 with two exceptions: 1) Each test block had a duration of 120 seconds (compared to 60 seconds in the first experiment); 2) To prevent habituation with the perceptual tasks, after each test block, participants were asked to tell in 60 seconds as many words as possible that start with random letters or are animals. In contrast, in Experiment 1 the additional tasks were administered between each block.

2.6.3.1 Perceptual tasks.

The perceptual tasks were the same as in Experiment 1, except for the verbal transformations task that was changed in this experiment. The stimulus was the word “life”, recorded by a native female English speaker (26 years old). The word duration was 390 ms, with a mean pitch of 203 Hz. The word was repeated without pause for 307 times to create the ambiguous sounds. In the disambiguated versions of the sounds, the word “life” and the word “fly” were presented once. See the spectrograms of the sounds in Appendix D.

2.6.3.2 Individual differences tasks.

Day-Night Stroop Task. The procedure was identical to the one described in Experiment 1 (see Section 2.2.3.4).

Ego-Resiliency. Ego resiliency was measured using a 14-item questionnaire (with a range of responses from 1: *Does not apply to me* to 4: *Applies very strongly*). The items were

formulated in the following way: "*I like to do new and different things*" or "*My life is full of things that keep me interested*". The questionnaire can be found in Appendix E. The total scoring is calculated by adding the responses to all the 14 items.

2.7 Results

2.7.1 Participants Exclusion

Four participants were excluded from the final analyses: one due to technical problems and three due to failing in the control tasks (two in the visual control task and one in the auditory control task) that were designed to test whether participants understood the instructions and were able to mark their perception as instructed (see Section 2.2.3.3). Final sample, $N = 31$ (seven males), $M_{age} = 21.2$, $SD = 3.69$).

2.7.2 Data Analysis

Data analyses were performed similarly to Experiment 1. The switching rates were computed for each block by dividing the number of switches by 120 (the block duration in seconds, from which the first perceptual phase was excluded).

Spearman rank-order correlations were calculated separately for the perceptual switching variables across the four tasks to assess whether they correlate. Correlations between perceptual switching variables in each perceptual task and the individual differences tasks were evaluated to investigate the relationships between perceptual switching and individual differences in inhibition and ego-resiliency.

2.7.3 Preliminary Analyses

There was a 20% switching rate with zero values as most participants (81%) did not switch in some of the blocks and tasks (predominantly in the ambiguous motion and auditory

stimuli, see Table 2.14). These values were treated as missing values and omitted from the analyses. The analyses with zero values instead of missing values are reported in Appendix F.

2.7.3.1 First phase durations vs. subsequent phase durations.

A LMEM on the phase duration showed the first phase duration ($M = 28.35$, $SD = 33.68$) was significantly longer than the subsequent phase durations ($M = 8.68$, $SD = 11.65$), $F(1, 211) = 68.18$, $p < .001$. See Figure 2.11.

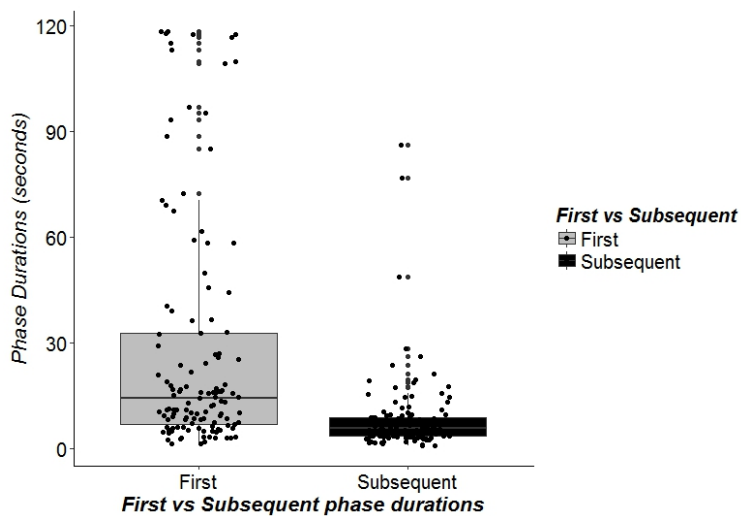


Figure 2.11. First phase durations versus subsequent phase durations.

2.7.3.2 Switching rates across blocks.

A LMEM was performed to assess whether there is a difference in the switching rates between blocks (See descriptive statistics in Table 2.13). Results showed no effect of Block on the switching rates $F(4, 195) = 0.734$, $p = 0.569$. This indicates that in this experiment the switching rates across the five blocks were not statistically different. However, to maintain consistency with Experiment 1, the first block was excluded from subsequent analyses.

Table 2.13

Means and Standard Deviations for the Switching Rate in each Block

	Block1	Block2	Block3	Block4	Block5
Mean	.089	.090	.089	.087	.084
SD	.080	.093	.088	.108	.102

2.7.3.3 Distribution of phase durations.

The visual inspection of the switching rates in the histograms below (see Figure 2.12) display the expected trends of what was previously reported: fewer switches that have a long phase duration and many shorter phase durations (Leopold & Logothetis, 1999; Pressnitzer & Hupé, 2006). There are no abnormalities in participants' perceptual switching behaviour in any of the tasks. It can be observed, however, that verbal transformations does not have as many long phase durations compared to the other tasks, which is also reflected in the higher number of switches.

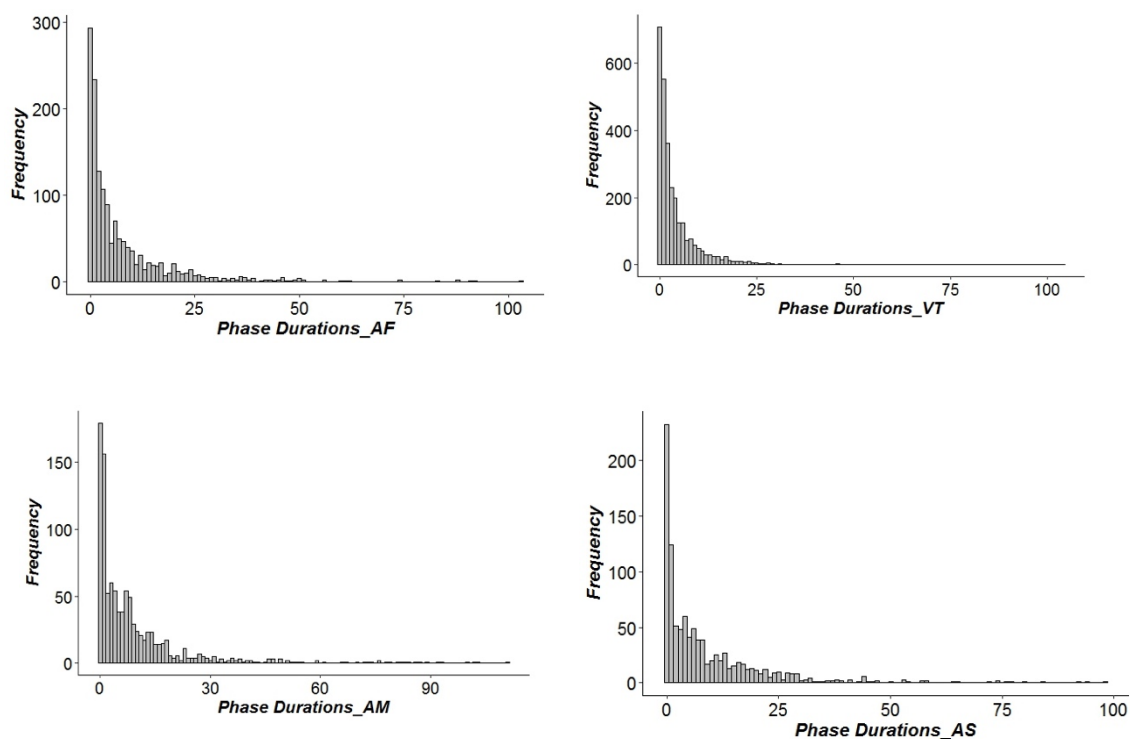


Figure 2.12. Histograms of phase durations for each task. Phase durations are displayed on the x-axis and the frequency of the phase durations that fall within a specific time interval on the y-axis. *Note.* The length of the axes is individually adapted for each data set.

2.7.4 Balance Between Different Interpretations

The balance between the two perceptual alternatives was assessed using the same procedure as described in Section 2.3.5. The analysis showed that the reports were more balanced than in Experiment 1 (ambiguous figure = .25, ambiguous motion = .08, verbal transformations = .06, auditory streaming = .16). Table 2.14 shows the average percentage of time spent in each of the alternative interpretations (first perceptual phase included) and for the “none” response (i.e., when no key was depressed), together with the percentage of participants reporting percept 1 as their first perception of the stimulus. The percentage of time during which neither response button was pressed down is below 10% in each of the tasks, showing that participants had clearly perceived one or the other alternative most of the time. It is of some surprise that the highest rate of no-responses was for the verbal transformations. One reason for this could be that the stimulus triggered a very large number of switches (the most over the four tasks as will be shown later). Because of this, participants might be reluctant to change buttons with such a great frequency of change, and they made decisions to release both buttons unless they heard a stable interpretation for longer time.

Table 2.14

Proportion of Time spent in Each Alternative Interpretation for each of the Four Perceptual Tasks

Task	Percept 1 Mean (SD)	Percept 2 Mean (SD)	None Mean (SD)	Percept 1 as first interpretation
AF	.35 (.15)	.60 (.16)	.05 (.03)	37%
AM	.43 (.14)	.51 (.13)	.05 (.06)	48%
VT	.42 (.11)	.48 (.12)	.10 (.07)	39%
AS	.39 (.17)	.55 (.17)	.06 (.05)	45%

Note. Percept 1 and Percept 2 in each of the 4 tasks were as follow: AF: Percept 1 = duck, Percept 2 = rabbit; AM = Percept 1 = CW, Percept 2 = ACW; VT: Percept 1 = fly, Percept 2 = life, AS: Percept 1 = Integrated; Percept 2 = Segregated.

2.7.5 Effects of Modality and Level of Processing on Perceptual Switching

2.7.5.1 Initial reaction time.

The overall performance of the initial reaction times across the four perceptual tasks is summarised in Table 2.15.

Table 2.15

Initial Reaction Times Performance across the Four Perceptual Tasks (N = 31)

	Tasks	No. Obs	Min	Max	Mean	SD	SEM
Initial RT	AF	124	0.29	13.74	2.69	1.84	.17
	AM	124	0.16	8.92	2.31	1.47	.13
	VT	124	0.67	7.59	2.13	0.98	.09
	AS	124	0.83	34.33	4.67	5.53	.05

Note. Maximum number of observations was 124.

The initial reaction times for each task (Modality x Level) is displayed in Figure 2.12. The initial reaction times were longer for auditory than visual tasks as indicated by the effect of Modality, $F(1, 34) = 6.92, p = 0.02$. There was an effect of Level, $F(1, 41) = 12.20, p = .001$, the initial reaction time for the high-level tasks was shorter than for the low-level tasks. The main effects were qualified with an interaction between Modality and Level, $F(1, 44) = 23.21, p < .001$. Pairwise contrasts showed that the difference between the initial reaction time in ambiguous figure (high-level visual) and ambiguous motion (low-level visual) did not differ $t(30) = 1.69, p = 0.122$. In the auditory modality, participants responded faster to verbal transformations (high-level) than to auditory streaming (low-level), $t(30) = -5.166, p < .001$. Further, results showed that the initial reaction time in ambiguous figure (high-level visual) was not significantly different from verbal transformations (high-level auditory), $t(30) = 1.715, p = 0.122$, while the initial reaction time in ambiguous motion (visual low-level) was significantly shorter than in auditory streaming (auditory low-level), $t(30) = -4.139, p < 0.001$.

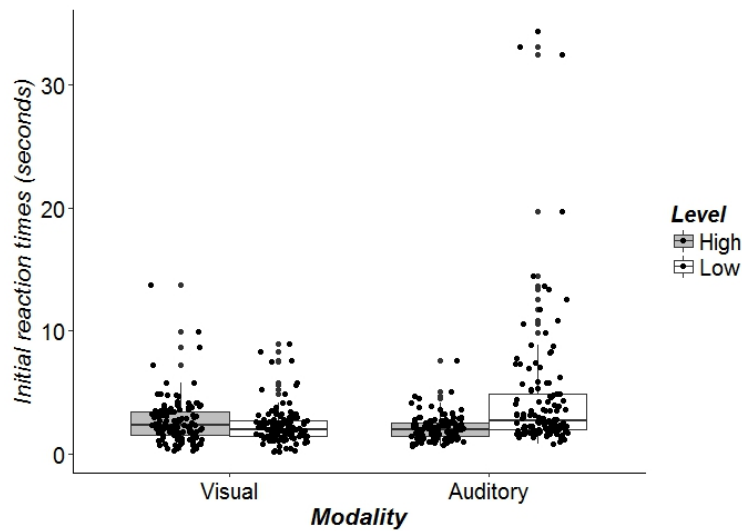


Figure 2.13. Initial reaction times in each perceptual task. Box represents the 25th and 75th percentile; data points over the boxes indicate data range.

2.7.5.2 First phase duration.

The overall performance of first phase duration across the four perceptual tasks is summarised in Table 2.16 and in Figure 2.14.

Table 2.16

First Phase Duration Performance across the Four Perceptual Tasks (N = 31)

	Tasks	No. Obs	Min	Max	Mean	SD	SEM
1 st Phase Duration	AF	104	1.09	66.05	15.41	14.00	1.51
	AM	84	0.04	106.9	20.09	22.67	2.47
	VT	118	0.02	44.93	8.92	8.38	0.82
	AS	89	0.41	89.97	19.89	17.69	1.88

Note. Maximum number of observations was 124.

The first phase duration for each task (Modality x Level) is displayed in Figure 2.13. There was no effect of Modality, $F(1, 37) = 2.06, p = .160$. There was a main effect of Level, $F(1, 29) = 10.37, p = .003$, the first phase duration in the high-level tasks were larger than in the low-level tasks. However, there was an interaction, $F(1, 34) = 11.63, p = .002$. Pairwise comparisons showed that only in the auditory modality there was a significant difference between the first phase duration in verbal transformations (high-level) and auditory streaming (low-level), $t(26) = -5.198, p < .001$. In the visual tasks, the first phase durations in

ambiguous figure (high-level) and ambiguous motion (low-level) did not differ significantly, $t(24) = 0.03, p = 0.976$. The results showed further that there was an impact of Level across modalities (high vision vs. high auditory and low vision vs. low auditory). In ambiguous figure (high-level visual task) there was a significantly larger first phase duration than in verbal transformations (auditory high-level), $t(27) = 4.46, p < 0.001$, while the first phase duration in auditory streaming and ambiguous motion did not differ significantly, $t(25) = -0.99, p = 0.397$.

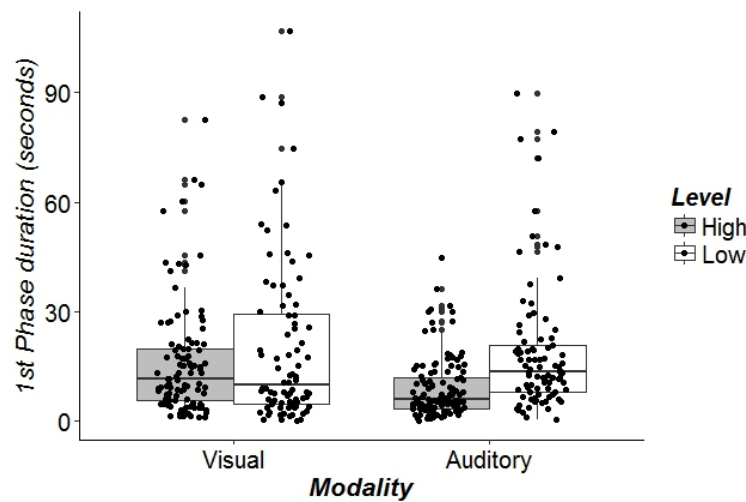


Figure 2.14. First phase duration in each perceptual task. Box represents the 25th and 75th percentile; data points over the boxes indicate data range.

2.7.5.3 Switching rates.

The overall performance for the switching rates across the four perceptual tasks is summarized in Table 2.17 and in Figure 2.15.

Table 2.17

Switching Rate Performance across the Four Perceptual Task (N = 31)

	Tasks	No. Obs	Min	Max	Mean	SD	SEM
Switching Rate	AF	104	0.01	0.36	0.07	0.08	.007
	AM	84	0.01	0.24	0.06	0.04	.004
	VT	118	0.01	0.84	0.14	0.14	.013
	AS	89	0.01	0.26	0.06	0.05	.005

Note. Maximum number of observations = 124.

The switching rate for each task (Modality x Level) is displayed in Figure 2.14. As indicated by a significant main effect of Level, $F(1, 41) = 5.67, p = .026$, the switching rates were higher in the high-level tasks than for the low-level tasks, whereas the main effect of Modality only approached significance, $F(1, 32) = 3.55, p = .069$. These main effects were qualified by an interaction between Modality and Level, $F(1, 34) = 10.89, p = .002$. The interaction was further disentangled with pairwise comparisons, adjusted for multiple testing. For the auditory task, the results showed that the switching rate in verbal transformations was higher than in auditory streaming, $t(26) = 3.47, p = .006$. There was no difference between the switching rate in ambiguous motion and ambiguous figure $t(23) = -0.25, p = .804$. Secondly, the impact of Level across modalities (high vision vs. high auditory and low vision vs. low auditory) was also investigated. Results showed that the switching rate in ambiguous figure was significantly lower than in verbal transformations, $t(28) = -3.42, p = 0.006$, while the switching rate in ambiguous motion and auditory streaming was not significantly different, $t(24) = 0.484, p = 0.804$.

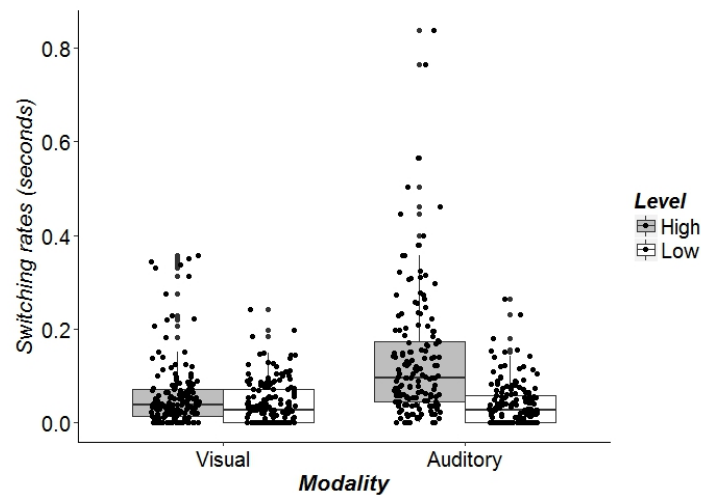


Figure 2.15. The switching rate in each perceptual task. Box represents the 25th and 75th percentile; data points over the boxes indicate data range.

2.7.6 Commonalities Between Perceptual Switching Across Tasks

The results showed that there were positive correlations between the switching rate in ambiguous figure and verbal transformations and between ambiguous figure and ambiguous motion (Table 2.20). The time it took participants to select the first interpretation correlates strongly across tasks, except for the initial reaction times in ambiguous figure and auditory streaming which do not relate to each other (Table 2.18). The time it took participants to switch to the second interpretation (i.e., the first phase duration) did not relate across tasks (Table 2.19).

2.7.7 Individual Differences

Day-Night Stroop task. Performance accuracy (percentage of correct responses) on Stroop was very high (97.24% correct). Performance accuracy on Stroop memory control was overall very high (average 99.66% correct). Because performance accuracy was so high, the analyses were focused on reaction times (RTs) on the following variables: Stroop RTs, Stroop memory control RTs and inhibition RTs (Stroop RTs minus Stroop memory control RTs). A paired t-tests revealed that the memory control RTs ($M = 0.67s$, $SD = .20$) were significantly shorter than in the Stroop RTs ($M = .82s$, $SD = .24$), $t(34) = 4.35$, $p < .001$, suggesting that it took longer time to correctly solve the inhibition task than the memory task.

Ego-Resiliency. The average score was 42.26 ($SD = 4.7$) [range: 32-51]. According to Block & Kremen (1996) this score is interpreted as a high resiliency trait.

2.7.8 Perceptual Switching and Individual Differences Measures

The switching rates correlated negatively with inhibition RTs; the faster participants responded to the verbal transformations, the faster they responded in the inhibition task. However, there was also a negative correlation with the memory control task and no correlation with the Stroop effect. The switching rates in the auditory streaming task

correlated positively with Stroop effect, the higher the switching rate, the lower the inhibition.

There was a positive correlation between first phase duration in the verbal transformations and inhibition. People that had a longer first phase durations in verbal transformations also had longer reaction times in the inhibition task. The first phase duration in auditory streaming correlated negatively with Stroop effect. Initial reaction times in all the tasks (except auditory streaming) correlated positively with the inhibition task.

Table 2.18

Correlations between the Initial Reaction Times across the Four Perceptual Tasks and the Individual Differences Measures

Task		AF	AM	VT	AS	Inhibition	Mem Control	Stroop Effect	Ego-Resiliency
AF	S rho	--	.552**	.545**	.479**	.476**	.361*	.101	-.183
	N		31	31	31	31	31	31	31
AM	S rho		---	.493**	.176	.426*	.239	.194	-.035
	N			31	31	31	31	31	31
VT	S rho			--	.582**	.365*	.303	.111	-.224
	N				31	31	31	31	31
AS	S rho				--	.284	.288	.077	-.070
	N					31	31	31	31

Note. S rho = Spearman correlation coefficient

Table 2.19

Correlations between the 1st Phase Durations across the Four Perceptual Tasks and the Individual Differences Measures

Task		AF	AM	VT	AS	Inhibition	Mem Control	Stroop Effect	Ego-Resiliency
AF	S rho	--	.401	.320	.305	.023	.032	-.160	-.076
	N		23	29	22	29	29	29	29
AM	S rho			.210	.152	.169	-.059	.314	.246
	N			24	18	24	24	24	24
VT	S rho				.045	.437*	.287	.307	-.019
	N				23	31	31	31	31
AS	S rho					-.138	.189	-.424*	.093
	N					23	23	23	23

Note. S rho = Spearman correlation coefficient.

Table 2.20

Correlations between the Switching Rates across the Four Perceptual Tasks and the Individual Differences Measures

Task		AF	AM	VT	AS	Inhibition	Mem Control	Stroop Effect	Ego-Resiliency
AF	S rho	--	.452*	.376*	.289	-.271	-.149	-.130	-.053
	N		23	29	22	29	29	29	29
AM	S rho			-.085	.082	-.189	-.055	-.260	.009
	N			24	18	24	24	24	24
VT	S rho				.139	-.499**	-.447*	-.163	-.030
	N				23	31	31	31	31
AS	S rho					.331	.063	.442*	.051
	N					23	23	23	23

Note. S rho = Spearman correlation coefficient.

2.8 Discussion Experiment 2

2.8.1 The Effects Of Level Of Processing On Perceptual Switching

The level of processing had an effect only in the auditory tasks. Specifically, initial reaction time and first phase duration were larger in auditory streaming than in verbal transformations, while the switching rate was higher in verbal transformations than auditory streaming, suggesting that perceptual switching in auditory modality is strongly influenced by the content of the stimuli as expected from research on visual modality (Strüber & Stadler, 1999). The switching rate was higher and the first phase duration shorter in the verbal transformations compared to ambiguous figure. The fact that verbal transformations had the fastest switching rate shows that stimuli with language content might be processed differently. Language processing is a highly exercised skill for humans. Extensive use can lead to strong representations for words, which allows a faster retrieval and processing for words (Rayner & Clifton, 2009).

The switching rate in the two visual tasks did not differ across levels as expected from previous research (Strüber & Stadler, 1999). There might be a couple of reasons for this. On the one hand, a difference from previous studies might be due to the stimuli used. Even if the

interpretations in the ambiguous motion do not have a specific nameable content, they are very easily accessible as people choose every day between right and left. This, together with the fact that the stimulus is more animated because it is moving, might sustain more the attention on the stimulus, which is known to increase perceptual switching (Alais, Newell, & Mamassian, 2010). On the other hand, more participants never switched in ambiguous motion than ambiguous figure, so the average of the switching rate in ambiguous motion was calculated on less number of observations, which could have influenced the difference between the switching rates between the two visual tasks. Another possibility might be that for vision the level of processing might be more influential on voluntary conditions (when participants are asked to switch as fast as possible or to hold an interpretation for as long as possible) than on neutral conditions as tested in this experiment. This hypothesis is supported by recent neuroscientific studies which show that the activity in the prefrontal cortex is functionally involved in the voluntary perceptual switching (when observers are asked to switch as fast as possible/hold as much as possible) and not in the passive perceptual switching conditions (de Graaf, de Jong, Goebel, van Ee, & Sack, 2011; Knapen, Brascamp, Pearson, van Ee, & Blake, 2011). Nevertheless, the effects of level of processing reported in the visual literature (Strüder & Stadler, 1999) were found here in auditory bistability, supporting previous studies which show that the switching rate in verbal transformations are higher than in auditory streaming (see Kondo & Kashino, 2007).

2.8.2 Commonalities Between Perceptual Switching Across Perceptual tasks

The second finding of this experiment regards the commonalities between the switching rates across modalities. The switching rates in the two high-level tasks (ambiguous figure and verbal transformations) were positively related, while there was no relationship between the switching rates in the low-level tasks (auditory streaming and ambiguous motion). Moreover, the switching rates in the two visual tasks were positively related while

there was no correlation between the switching rates in the two auditory tasks. Initial reaction time across all the tasks correlated, except for initial reaction time between ambiguous motion and ambiguous figure which did not relate. The positive correlations between perceptual switching in ambiguous figure and verbal transformations support the results of Kondo et al (2012) and are not consistent with results from Experiment 1. One reason why in this experiment there are correlations, while in Experiment 1 the correlations were not found, might be the stimulation time difference. It might be possible that for the short viewing/listening time in Experiment 1 there were not enough data points to capture the relationships. Thus, the data presented here should be complimented by more studies with longer stimuli presentation. This will be addressed in Study 3 and Study 4.

2.8.3 Common Individual Differences in Perceptual Switching?

The third finding refers to the correlations between perceptual switching and the individual differences manifested in the Stroop task and Ego-resiliency questionnaire. The reaction times in the inhibition task correlated negatively with the switching rates in auditory streaming, indicating that the better the inhibition, the larger the switching rates in auditory streaming. This result is contradictory with a recent finding of a positive correlation between inhibition and auditory streaming (Farkas et al., 2016). The initial reaction time in all the perceptual tasks, except auditory streaming, correlated positively with the reaction times in the inhibition task, showing that the better the inhibition (smaller RTs), the faster is the first perceptual decision made.

2.9 General Conclusion Study 1

The main aim of this study was to examine to what degree perceptual switching (initial reaction time, first phase duration and switching rate) in visual and auditory bistability is domain general or modality specific. For this purpose, participants' perceptions for two

auditory and two visual tasks with one low- and one high-representation-level stimulus were recorded continuously. Two experiments using similar tasks were conducted. Based on the limitations from Experiment 1, the stimulus duration was increased from 60 seconds to 120 seconds and a within participant study was designed in Experiment 2 compared to a between participant study in Experiment 1. The results showed consistent patterns as well as some dissimilarities that can be attributed to the changes performed.

The first main finding regards the commonalities between the perceptual switching in the two modalities. In Experiment 1, the switching rate across tasks did not correlate. In Experiment 2, with increased stimulus presentation, the switching rate in ambiguous figure and verbal transformations correlated positively, as well as the switching rate between verbal transformations and ambiguous motion. These results indicate that the number of perceptual switches across modalities correlate in the high-level tasks, while in the low-level tasks the perceptual switches do not relate. However, as the correlations are not only between the high-level tasks, but also between the high-auditory and low-visual tasks, it cannot be assumed that the correlations are due only to the content of the stimulus. This result is in line with the common hypothesis theory that supports that perceptual switching across modalities have common mechanisms of functioning (Kondo et al., 2012). Moreover, consistent evidence indicates that a common mechanism might be involved in the initial perceptual decision and the inertia of the first phase duration in both visual and auditory modalities (Hupé & Pressnitzer, 2012).

In both experiments, the level of processing affected perceptual switching in both modalities. For *the initial reaction time*, results were similar across experiments. Participants were much faster to report their perception in verbal transformations than in auditory streaming, whereas the initial reaction time in ambiguous figure and ambiguous motion did not differ. For the *first phase durations* in Experiment 1, the first phase duration was longer

in ambiguous motion than in ambiguous figure. In Experiment 2, the difference between levels within the visual modality was no more significant. The first phase durations and initial RT are considered to be sensitive to stimulus parameters (Denham et al., 2013). While these parameters are usually disregarded and not analysed (van Ee, 2009), some authors use these parameters to characterise the processes underlying perceptual switching (Denham et al., 2012; Denham et al., 2013). During the first phase duration, the system accumulates information about the other possible alternative interpretations (Barniv & Nelken, 2015). The results presented here suggest that this process is faster for high-level tasks than the low-level tasks, especially in the auditory tasks. At the beginning of the first perceptual phase, it is also thought that the perceptual system is essentially concerned with the formation of perceptual organisations (Denham et al., 2013). Phonological representations of words are more easily available than the representations of simple sounds. For visual tasks, the effects of the level of processing were not consistent across experiments. Therefore, it is not clear if the initial representations of concrete objects (e.g., a duck or a rabbit) are easier accessed than representations of moving objects (e.g., moving to the left or to the right).

For the *switching rate*, the level of processing difference was consistently found to affect perceptual switching in audition but not in vision. The fact that the switching rates in verbal transformations (high-level) were larger than in auditory streaming (low-level task) is consistent with previous findings of Kondo et al. (2012) who found similar results.

The differences between the two experiments might be due to the change in the verbal transformations task, which resulted in a much larger switching rate in Experiment 2 compared to Experiment 1. For example, the switching rate for verbal transformations in Experiment 2 increased from 0.14 (in Experiment 1) to 0.17 (in Experiment 2). However, there are differences for the other three perceptual tasks as well. The switching rate for ambiguous figure in Experiment 2 ($M = 0.07$) decreased significantly compared to

Experiment 1 ($M = 0.16$). A decrease is also noticed for ambiguous motion; the switching rate in Experiment 1 was higher ($M = 0.09$) than in Experiment 2 ($M = 0.06$). Another explanation for this difference might be that having all the four tasks in a single experimental session—with each block for 120 seconds—may introduce performance factors such as tiredness and loss of attention, which can decrease the switching rates. Finally, the switching rate for the auditory streaming in Experiment 2 ($M = 0.06$) increased slightly compared to Experiment 1 ($M = 0.04$). One reason for this result might be that the switching rate for auditory streaming in Experiment 1 was the lowest to start with, and it might have benefitted the most from an extended listening time. Thus, an extended listening time might have stronger impact on the switching rates for auditory streaming than the other stimuli and it can help to stabilize the switching rates in the other tasks (Suzuki & Grabowecky, 2007).

The third important finding is that inhibition, creativity, and cognitive flexibility correlate with perceptual switching in some of tasks, although not consistently. This indicates that the role of these factors for perceptual switching is task specific.

Overall, the results of Study 1 showed that there are commonalities in perceptual switching rates across modalities and that the individual differences tested were not consistently related with perceptual switching, indicating they do not play a common role across modalities/tasks.

3 STUDY 2: VISUAL AND AUDITORY MULTISTABILITY IN 6, 8, 10-YEAR OLD CHILDREN

3.1 Introduction

The present study carries on from the previous study to investigate the development of children's perceptual switching in visual and auditory multistability. As detailed in the general introduction, only a few studies have investigated the onset of perceptual switching in children (Doherty & Wimmer, 2005; Gopnik & Rosati, 2001; Mitroff et al., 2006; Rock et al., 1994; Warren & Warren, 1966; Wimmer & Doherty, 2011) and no study has investigated both visual and auditory multistability in the same sample of children. Although the ability to perceive both interpretations of an ambiguous stimulus is developed by the age of 5 (Warren & Warren, 1966; Wimmer & Doherty, 2011), it is unclear how the number of switches changes with age. It is also unclear whether the number of switches in visual and auditory tasks follows the same developmental schedule and/or draw on common cognitive mechanisms. Despite lively debate in the literature concerning the domain specificity of adults' perceptual switching abilities, there has been no attempt to observe the development of these skills to determine if they show the same developmental trajectory. This is important, as the common hypothesis would predict that visual and auditory switching abilities should emerge together and associate with the same cognitive skills. Producing the first systematic developmental picture of this phenomenon, the present study tracks the visual and auditory multistability of 6-, 8-, and 10-year-old children.

If the processing of ambiguous stimuli is domain general (Kondo et al., 2012), the number of switches in response to ambiguous visual and auditory tasks should be positively correlated. If perceptual switching is domain specific (Pressnitzer & Hupé, 2006), then no correlations should emerge. Moreover, based on the previous findings (Strüber & Stadler, 1999) higher switching rates may be found in response to ambiguous figure and verbal

transformations (the high-level tasks) than for ambiguous motion and auditory streaming (low-level tasks). In addition, there might be correlations between perceptual switching in ambiguous figure and verbal transformations but not between ambiguous motion and auditory streaming (Pressnitzer & Hupé, 2006). This would occur if the content of the stimuli were a major common top-down effect involved across modalities. Based on the early evidence from Warren and Warren (1966) the number of switches for the verbal transformation is expected to increase with age. However, it is unclear whether this would also be found in ambiguous figures, motion, and auditory streaming tasks. If ambiguous switching rate underlies general maturational processes and perceptual processing styles (Bremner et al., 2016; Doherty et al., 2010; Nayar et al., 2015) then we would expect to find the same age effects across all ambiguous perception tasks. If switching rate does not underlie overall shifts in perceptual processing then we should find task specific effects across different age groups.

Another aim is to explore whether executive functions and creativity affect perceptual switching in both vision and audition throughout childhood. In light of previous research on visual bistability (Wimmer & Doherty, 2011; Wiseman et al., 2011), the ability to switch should be predicted by executive function and creativity. However, it is unclear what role, if any, creativity and executive function play in the *number of switches* in both the visual and auditory domain in childhood. Moreover, it is currently unclear whether visual and auditory switching draws on distinct cognitive processes. If this is the case, we should see a dissociation between the developmental skills that correlate with visual and perceptual switching tasks. Alternatively, if the number of switches across ambiguous tasks is consistently related to executive functions or creativity this will indicate a domain-general role of these abilities in the development of ongoing perceptual switching as a characteristic of a flexible perceptual system.

3.2 Method

3.2.1 Participants

Sixty-six native English speaking children (6-, 8-, and 10-year-olds; 26 girls, 37 boys) from a primary school in Plymouth, England took part. There were twenty-two 6-year-olds ($M = 72$ months, $SD = 4$), twenty-three 8-year-olds ($M = 96$, $SD = 4$), and twenty-one 10-year-olds ($M = 122.86$, $SD = 2$). Children who took part had parental consent and gave their own assent on the day of testing.

3.2.2 Design

Each child performed the following tasks: Ambiguous figure, Ambiguous structure-from-motion, Verbal transformations, Auditory streaming, visual and auditory control tasks, Day-Night Stroop, Pattern meanings test, and Verbal fluency. The tasks are described in detail in Chapter 2.

3.2.3 Materials and Procedure

The procedure and the methods are identical to the one described in Experiment 2 with a few exceptions. Firstly, there were three blocks instead of five, and the viewing/listening time was 60 seconds per block. Although from Experiment 2 we could see that stimulus duration makes a difference to the results, increasing the stimulus duration in this study would have made it very hard for children to pay attention to the tasks for so long. Secondly, for practical reasons, data collection was conducted in two sessions instead of one. Each child was tested in two sessions of 25-35 minutes in a quiet room in a school. Children sat in front of the computer at approximately 60 cm distance from the screen. The sounds were presented binaurally through headphones (Sony, MDRNC7B) adjusted to a comfortable volume. The experimenter sat next to the child for the entire duration of the experiment. Whenever needed, the experimenter reminded children not to move their heads and to focus

on their task until the break. The tasks were counterbalanced within every session and between sessions with the constraint that there were never two perceptual tasks from the same modality in a row.

3.3 Results

3.3.1 Preliminary Analyses

Participants were excluded if they failed to report a switch in perception in the auditory control task and the horse/sheep tasks or if they reported a switch in the girl task. Three participants were excluded from the final analyses based on these criteria: two from 6-years-olds group (one for failing the auditory control task and one for failing the visual control task) and one participant from 8-years-olds group (for failing the auditory control task). The final sample size is $N = 63$: twenty 6-year-olds ($M = 72$ months, $SD = 4$), twenty-two 8-year-olds ($M = 96$, $SD = 4$), and twenty-one 10-year-olds ($M = 122.86$, $SD = 2$).

3.3.2 Data Analyses

Two kinds of analyses were performed. Firstly, the ability to switch *per se* in the first block and the number of switches were analysed with rANOVA. Secondly, Spearman correlations were calculated between the perceptual variables (the ability to switch and number of switches across tasks) and inhibition, set-shifting, and creativity. Subsequently, partial Spearman correlations were computed to control for the effects of age.

3.3.3 Proportion of Time Spent in Each Alternative Interpretation

Table 3.1 shows the average percentage of time spent in each of the alternative interpretations and for the “none” response (i.e., when no key was pressed), together with the percentage of participants reporting one of the interpretations (here percept 1) as their first perception of the stimulus. For ambiguous figure: percept 1 = duck, for ambiguous structure-

from-motion: percept 1 = clockwise, for verbal transformations: percept 1 = life, and for auditory streaming: percept 1 = integrated. The percentage of time during which neither response button was pressed is overall higher than for adults (where the maximum was 9%), showing that for children it takes more time to make a change in perception. However, the high proportion of none response can also be due to children finding it much harder to focus continuously on the task.

Table 3.1

The Proportion of Time Spent in Each Alternative Interpretation

Task	AGE	Percept 1 Mean (SD)	Percept 2 Mean (SD)	None Response Mean (SD)	Percept 1 as first interpretation
AF	6	.21 (.17)	.43 (.15)	.36 (.22)	42%
	8	.33 (.20)	.48 (.21)	.19 (.09)	40%
	10	.29 (.17)	.51 (.16)	.20 (.10)	43%
AM	6	.41 (.14)	.41 (.14)	.18 (.09)	74%
	8	.44 (.17)	.44 (.17)	.12 (.06)	52%
	10	.45 (.12)	.47 (.12)	.09 (.09)	65%
VT	6	.29 (.13)	.45 (.14)	.26 (.10)	78%
	8	.35 (.13)	.46 (.14)	.19 (.05)	83%
	10	.34 (.07)	.49 (.08)	.16 (.05)	73%
AS	6	.30 (.19)	.40 (.18)	.30 (.16)	52%
	8	.32 (.19)	.47 (.22)	.21 (.15)	52%
	10	.39 (.16)	.45 (.20)	.16 (.12)	.44%

Note. Percept 1 and Percept 2 in each of the 4 tasks were as follow: AF: Percept 1 = duck, Percept 2 = rabbit; AM = Percept 1 = clockwise, Percept 2 = anticlockwise; VT: Percept 1 = life, Percept 2 = fly, AS: Percept 1 = integrated; Percept 2 = segregated.

A mixed ANOVA was computed on the proportion of “none” with task (ambiguous figure, ambiguous motion, verbal transformations, auditory streaming) as a within factor and age (6, 8, and 10) as a between factor. Results showed an effect of task, $F(3, 180) = 13.31$, $\eta_p^2 = .182$, $p < .001$. Post-hoc comparisons showed that the proportion of none responses was

higher in ambiguous figure ($M = .24$, $SD = .16$) than ambiguous motion ($M = .13$, $SD = .08$), ($p < .001$). Proportion of none responses for ambiguous motion was also significantly lower than for verbal transformations ($M = .20$, $SD = .08$) and auditory streaming ($M = .22$, $SD = .14$) (all $ps < .001$). The proportion of none responses for verbal transformations and auditory streaming did not differ ($p > .05$).

There was an effect of age, $F(1, 60) = 14.61$, $\eta_p^2 = .328$, $p < .001$. Post-hoc comparisons that the proportion of none responses was higher in the group of participants aged 6 compared to the ones aged 8 ($p = .003$) and 10 ($p < .001$). The proportion of none did not differ between the groups of participants aged 8 and 10 ($p = .350$). Overall, results show that 6-year-old children take more time to make a change in perception compared to 8- and 10-year-olds.

3.3.4 Perceptual Ambiguous Tasks: The Ability To Switch In Block 1

To compare with the typical developmental literature in ambiguous perception it was first examined whether children between 6 and 10 years of age would be able to switch across the different tasks in the first 60 seconds block. See the overall performance in the ability to switch (switch/no switch) in Figure 3.1.

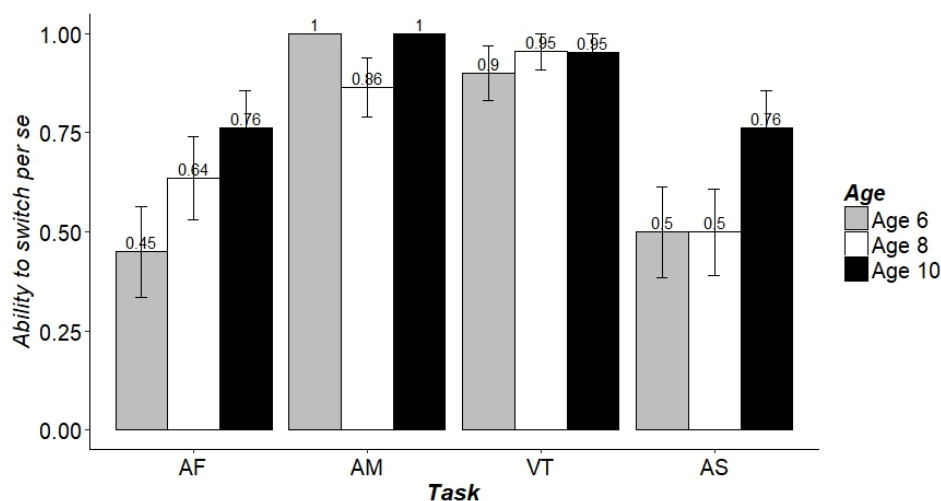


Figure 3.1. Average ability to switch (0 = no switch; 1 = switched) for each perceptual task across the three age groups. Error bars show standard errors of means.

The ability to switch was examined with a three (age group: 6- vs. 8-vs. 10-year-olds) x four (tasks: ambiguous figure vs. ambiguous motion vs. verbal transformation vs. auditory streaming) rANOVA on the ability to switch was conducted, where age was the between participants variable and task the within variable.

There was an overall effect of age, $F(2, 60) = 3.29, p = .044, \eta_p^2 = .10$, but separate age group comparison showed only marginal age differences. Six-year-olds ($M = .71$) were not more likely to reverse than the 8-year-olds ($M = .74, p = 1.00$) or 10-year-olds ($M = .87, p = .06$). 8-year-olds did not differ from either age group ($ps > .14$) (Bonferroni post-hoc). There was also a difference in the ability to switch across tasks, $F(3, 180) = 18.82, p < .001, \eta_p^2 = .24$, where children were less likely to switch during the first 60 seconds period ($ps < .001$) in both the ambiguous figure ($M = .62$) and auditory streaming tasks ($M = .59$) than in the ambiguous motion ($M = .96$) and verbal transformation tasks ($M = .94$) who did not differ ($p = .100$) (Bonferroni post-hoc). This difference between tasks was the same for all age groups as there was no interaction between age and task, $F(6, 180) = 1.39, p = .22, \eta_p^2 = .04$.

3.3.5 Perceptual ambiguous tasks: Number of Switches

The number of switches is summarised in Figure 3.2 for each of the three blocks, separately for each age group

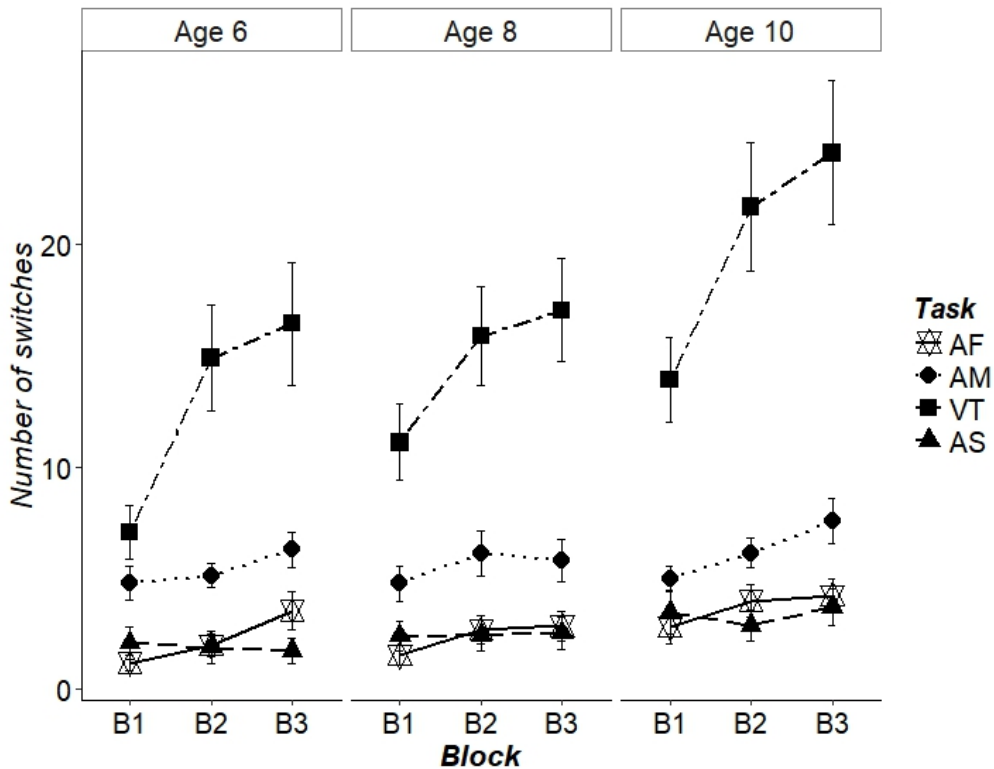


Figure 3.2. The mean number of switches, for each task across the three blocks, separately for age group. Error bars show standard errors of means. B1-B3 = Block1-3.

To examine switching rate in ambiguous tasks a three (age group: 6- vs. 8- vs. 10-year-olds) x four (tasks: ambiguous figure vs. ambiguous motion vs. verbal transformation vs. auditory streaming) x three (blocks: 1 vs. 2 vs. 3) rANOVA was computed on the number of switches with block and task as within participant variables and age group as between participant variable.

The number of switches increased with age, $F(2, 60) = 4.06, p = .022, \eta_p^2 = .12$, where 6-year-olds ($M = 5.56$) had fewer number of switches than 10-year-olds ($M = 8.27, p = .025$). Eight-year-olds ($M = 6.24$) did not differ in their number of switches from either age group ($ps > .12$) (Bonferroni post-hoc). There was an effect of block, $F(2, 120) = 50.37, p < .001, \eta_p^2 = .46$, where the number of switches increased between each adjacent, block 1 ($M = 4.98$) and block 2 ($M = 7.12, p < .001$) and in turn block 3 ($M = 7.97, p = .004$) (Bonferroni post-

hoc). The number of switches also differed across tasks, $F(3, 180) = 98.43, p < .001, \eta_p^2 = .62$, where higher number of switches occurred in the verbal transformation task ($M = 15.80$) than all other tasks ($ps < .001$). There were also higher number of switches in the ambiguous motion task ($M = 5.70$) than both ($p < .001$) the ambiguous figure ($M = 2.72$) and auditory streaming task ($M = 2.54$) who did not differ ($p = 1.00$) (Bonferroni post-hoc).

These two main effects were qualified by a block x task interaction, $F(6, 360) = 18.07, p < .001, \eta_p^2 = .23$ (Figure 3.2). This interaction occurred because the number of switches did not equally increase across all blocks for all tasks. Specifically, for the *ambiguous figure* task the number of switches increased between block 1 and both 2 ($p = .004$) and 3 ($p = .001$) but not between 2 and 3 ($p = .38$). For the *ambiguous motion* task the number of switches increased between block 1 and 3 ($p = .001$) and not between 1 and 2 ($p = .14$) nor 2 and 3 ($p = .30$). For the *verbal transformation* task the number of switches increased between block 1 and both 2 and 3 ($ps < .001$) and not between 2 and 3 ($p = .25$). In contrast, for the *auditory streaming* task the number of switches did not increase between any block ($ps > .81$) (Figure 3.2).

3.3.6 Inhibition, Set-shifting and Creativity

A summary of performance on Day-Night Stroop, Pattern meanings, and Verbal Fluency tasks is presented in Table 3.2.

Table 3.2

The Overall Performance on Stroop, Pattern Meanings Task, and Verbal Fluency

Age	Task	Measure	N	Minimum	Maximum	Mean	SD
Age 6	Stroop	Stroop RT	20	.793	2.227	1.444	.415
		Memory control RT	20	.664	1.693	1.141	.292
		Inhibition RT	20	-.414	.831	.303	.347
	Pattern meaning	cTotal	20	8	27	19.25	5.24
		cUnique	20	0	16	7.55	3.36
		cUnusual	20	0	7	3.85	1.95
	Verbal fluency	VF total	20	7	26	12.25	4.33
		VF switches	20	1	17	5.85	3.56
Age 8	Stroop	Stroop RT	22	.736	1.984	1.086	.312
		Memory control RT	22	.689	1.467	.968	.241
		Inhibition RT	22	-.526	1.265	.117	.445
	Pattern meaning	cTotal	22	10	43	24.41	8.98
		cUnique	22	1	22	8.18	6.50
		cUnusual	22	2	15	8.55	3.38
	Verbal fluency	VF total	22	9	25	16.91	4.99
		VF switches	22	4	14	9.05	2.70
Age 10	Stroop	Stroop RT	21	.509	1.541	.886	.280
		Memory control RT	21	.437	1.217	.722	.220
		Inhibition RT	21	-.153	.468	.164	.162
	Pattern meaning	cTotal	21	9	42	23.24	8.85
		cUnique	21	1	23	8.19	5.20
		cUnusual	21	1	15	6.00	3.59
	Verbal fluency	VF total	21	10	33	19.38	5.62
		VF switches	21	4	17	9.67	3.54

Note. Reaction times (RT) are displayed in seconds; Inhibition RT was calculated by subtracting Stroop memory control RT from Stroop RT.

3.3.6.1 Stroop.

Performance accuracy (percentage of correct responses) on Stroop was overall very high (Age 6: $M = .90$, Age 8: $M = .93$, Age 10, $M = .95$). A one way ANOVA on the performance accuracy showed no difference between the three age groups, $F(2, 62) = .98, p = .38$. Performance accuracy (percentage of correct responses) on memory control Stroop was also very high (Age 6: $M = .96$, Age 8: $M = .94$, Age 10, $M = .93$). A one way ANOVA on

the performance accuracy showed no difference between the three age groups, $F(2, 62) = .09$, $p = .92$.

As performance accuracy was very high, the analyses were focused on reaction times (RTs) on the following variables: Stroop RTs, memory control RTs and inhibition RTs (Stroop RTs minus memory control RTs). There was a significant difference between the three age groups on the Stroop RTs, $F(2,62) = 14.23$, $p < .001$, $\eta_p^2 = .322$. Post-hoc tests showed that Stroop RTs at age 6 were shorter than at age 8 ($p = .009$), at age 6 shorter than at age 10 ($p < .001$), while Stroop RTs at age 8 and age 10 did not differ ($p = .082$). There was also a significant difference between age groups for the Stroop memory control RTs, $F(2,62) = 14.40$, $p < .001$, $\eta_p^2 = .324$. The post-hoc test showed that memory control RTs at age 6 did not differ from age 8 ($p = .107$). However, 6-year-olds responded slower than 10-year-olds ($p < .001$) while age 8-year-olds were slower than 10-years-olds ($p = .003$). The performance for the inhibition RTs did not differ with age, $F(2, 62) = 1.66$, $p = .199$.

3.3.6.2 Verbal fluency.

Inter-rater reliability was calculated for the variable VF Switching. Approximatively 10% of the data was coded by two independent raters. A two-way mixed intraclass correlation analysis on the total number of switches showed that there was a substantial agreement between the two raters, $ICC = .973$ (95% *CI*, .887 to .994), $p < .001$.

There were two variables of interest: the total number of words generated (VF total) and the number of switches between clusters (VF switches). A one way ANOVA on the VF switches showed a significant effect of age, $F(2, 62) = 7.951$, $p = .001$, $\eta_p^2 = .210$ where 6-year-olds switched less than both 8-year-olds ($p = .007$) and 10-year-olds ($p = .004$). There was no difference between participants aged 8 and 10 ($p = .796$). There was also an effect of age on VF total, $F(2, 62) = 10.63$, $p < .001$, $\eta_p^2 = .261$, where 6-year-olds gave fewer

solutions than both 8-year-olds ($p = .007$) and 10-year-olds ($p < .001$). There was no difference between 8- and 10-year-old children ($p = .291$).

3.3.6.3 Pattern Meanings.

Data were split between two independent raters, and the inter-rater reliability score was calculated for a subset of the data that both raters coded. A two-way mixed intraclass correlation analysis on the unique number of responses showed that there was a substantial agreement between the two raters, $ICC = .919$ (95% CI , .868 to .950), $p < .001$. A substantial agreement between the two raters was also for the number of unusual responses, $ICC = .868$ (95% CI , .784 to .919), $p < .001$.

There were three variables of interest: the total number of solutions (cTotal), the number of unusual responses (cUnusual), and the number of unique responses (cUnique). Two one-way ANOVAs on the cTotal [$F(2, 54) = 1.47$, $p = .293$] and cUnique, [$F(2, 54) = .206$, $p = .815$] showed no effect of age. There was a significant effect of age on cUnusual, $F(2, 54) = 5.19$, $p = .009$, $\eta_p^2 = .166$, where 6-year-olds had a lower number of unusual solutions compared to 8-year-olds ($p = .002$). There were no further differences (all $ps > .05$).

3.3.7 The relation Between The Ability to Switch, Inhibition, Set-Shifting, And Creativity.

Spearman correlations were used to assess the relationships between the ability to switch in each of the four ambiguous tasks, inhibition, verbal fluency, and creativity (Table 3.3). Spearman partial correlations were run subsequently to control for the effect of age (Table 3.4). The ability to switch in the Ambiguous figure correlated with the ability to switch in the Auditory streaming. The ability to switch in Ambiguous figure also correlated with measures of creativity. After controlling for the effect of age, the ability to switch in ambiguous figure remained associated with the ability to switch in auditory streaming.

Table 3.3

Spearman Correlations between the Ability to Switch in the Ambiguous Tasks, Stroop, Pattern Meaning, and Verbal Fluency Tasks

	AF	AM	VT	AS	Stroop	MC	Inhibition	CTotal	CUnique	CUnusual	VFtotal	VFswitch
Age	.258*	.007	.085	.217	-.561***	-.584***	-.164	.180	-.001	.245	.525***	.441***
AF	--	-.022	.064	.338**	-.327**	-.261*	-.138	.262*	.178	.298*	.171	.242
AM		--	-.058	-.036	-.061	-.016	-.020	.160	.169	.115	.193	.095
VT			--	.046	-.075	.004	-.140	-.027	-.136	.137	-.027	.061
AS				-	-.326**	-.229	-.085	.222	.048	.085	.063	.003

Note. MC = Stroop memory control, Inhibition = Stroop – memory control, C = creativity total, CUnique = Creativity unique, CUnusual = Creativity Unusual VFtotal = verbal fluency total, VFswitch = verbal fluency switch.

Table 3.4

Spearman Partial Correlations between the Number of Switches in the Ambiguous Tasks, Stroop, Pattern Meaning, and Verbal Fluency Tasks, Controlling for the Effects of Age

	AF	AM	VT	AS	Stroop	MC	Inhibition	CTotal	CUnique	CUnusual	VFtotal	VFswitch
AF	-	-.024	.043	.299*	-.228	-.140	-.101	.227	.185	.251*	.043	.147
AM		-	-.059	-.038	-.070	-.016	-.020	.162	.169	.117	.223	.102
VT			-	.028	-.033	.066	-.128	-.043	-.137	.120	-.085	.026
AS				.	-.253*	-.129	-.051	.190	.049	.033	-.061	-.106

Note. MC = Stroop memory control, Inhibition = Stroop – memory control, C = creativity total, CUnique = Creativity unique, CUnusual = Creativity Unusual, VFTotal = verbal fluency total, VFswitch = verbal fluency switch.

3.3.8 The Relation Between The Number of Switches, Inhibition, Set-Shifting, And Creativity.

Spearman correlations were used to assess the relationships between the number of switches across the four ambiguous tasks, inhibition, memory, verbal fluency, and creativity (Table 3.5). Spearman partial correlations were run subsequently to control for the effect of age (Table 3.6) The number of switches in the ambiguous figure correlated with the number of switch in the auditory streaming. The number of switches in the verbal transformations also correlated with the number of switches in auditory streaming. The number of switches in ambiguous figure further correlated with scores of creativity and verbal fluency. The number of switches in verbal transformation correlated with measures of creativity and verbal fluency. After controlling for the effect of age, perceptual switching in the auditory streaming task remained associated with the number of switches in ambiguous figure. The number of switches in verbal transformations also remained associated with measures of creativity (pattern meaning).

Table 3.5

Spearman Correlations Between the Number of Switches in the Ambiguous Tasks, Stroop, Pattern Meaning, and Verbal Fluency Tasks

	AF	AM	VT	AS	Stroop	MC	Inhibition	CTotal	CUnique	CUnusual	VFtotal	VFswitch
Age	.201	.107	.270*	.232	-.561**	-.584**	-.164	.180	-.001	.245	.525**	.441**
AF	--	.167	.146	.468**	-.250*	-.333**	-.012	.273*	.256*	.271*	.160	.278*
AM		---	.081	.009	.000	-.038	.029	.099	.129	.027	.146	.213
VT			---	.477**	-.329**	-.289*	-.130	.314*	.254*	.155	.293*	.184
AS				---	-.370**	-.280*	-.178	.190	.020	.159	.066	.043

Note. MC = Stroop memory control, Inhibition = Stroop – memory control, C = creativity total, CUnique = Creativity unique, CUnusual = Creativity Unusual, VFtotal = verbal fluency total, VFswitch = verbal fluency switch.

Table 3.6

Spearman Partial Correlations between The Number of Switches In The Ambiguous Tasks, Stroop, Pattern Meaning, And Verbal Fluency Tasks, Controlling For The Effects Of Age.

	AF	AM	VT	AS	Stroop	MC	Inhibition	CTotal	CUnique	CUnusual	VFtotal	VFswitch
AF	--	.149	.097	.442***	-.169	-.272*	.022	.246	.262*	.234	.065	.216
AM		-	.055	-.016	.074	.030	.048	.081	.130	.001	.106	.186
VT		.	-	.442	-.222	-.168	-.090	.281*	.264*	.095	.184	.076
AS				-	-.297	-.183	-.145	.155	.021	.108	-.068	-.068

Note. MC = Stroop memory control, Inhibition = Stroop – memory control, C = creativity total, CUnique = Creativity unique, CUnusual = Creativity Unusual VFtotal = verbal fluency total, VFswitch = verbal fluency switch.

3.4 Discussion

In the first study of this kind, both visual and auditory ambiguous tasks were investigated in the same sample, in 6, 8 and 10 years old children. Results show that perceptual switching underline both general maturational processes and tasks specific processes.

Firstly, the results show that the ability to switch *per se* in the first 60 seconds is not significantly different in children between 6, 8 and 10 years of age. Therefore, the ability to switch is present from the age of six onwards, in a larger set of perceptual tasks. There were nevertheless differences between the four tasks in how likely children reported at least one perceptual switch in the first 60 seconds. Specifically, it was less likely for children to switch to the second interpretation in ambiguous figure and auditory streaming than in verbal transformations and ambiguous motion. Thus, although in average children switch in all the four tasks by the age of six, stimulus characteristics made some of the tasks more difficult than others. Concretely, the ambiguous figure and the auditory streaming may pose stimulus specific switching difficulties for the perceptual system compared to ambiguous motion and verbal transformations. One possibility is that switching to another interpretation in an ambiguous figure such as duck/rabbit requires not only the processing of contours and shapes, but the switching between two representations with semantic content, which seems computationally more difficult/complex than the processing of motion direction and 3D perception.

On the other hand, the ongoing perceptual switching developmental changes with increasing age. Specifically, 6-year olds had fewer perceptual switches than 10-year olds did, while the 8-year olds did not differ significantly from either 6- or 10-

year old children. This increase in the number of switches is consistent with previous findings where an increase in the number of switches was observed in either visual or auditory tasks (see Ehlers et al., 2015; Warren & Warren, 1966). This indicates that the mechanisms behind perceptual switching and thus, perceptual flexibility, have a protracted development. However, as there was no interaction between age and task and because the increase was present across tasks, it is suggested that general maturational processes are also present.

In the adult literature, perceptual switching has been suggested to happen because of the activity in both sensory and non-sensory brain regions such as fronto-parietal cortex (Leopold & Logothetis, 1999; Tong et al., 2006; Toppino & Long, 2004). An increase in the number of switches in children is consistent with the continuous maturation of fronto-parietal cortex (Ehlers et al., 2015; Sowell, Thompson, & Toga, 2004; Toga, Thompson, & Sowell, 2006) and because of the maturation of connections between higher-level brain regions and sensorial cortices that take place until late adolescence (Dekker, Schwarzkopf, de Haas, Nardini, & Sereno, 2017). As attention still develops until late adolescence (Plude, Enns, & Brodeur, 1994; Porporino, Iarocci, Shore, & Burack, 2004), it is also possible that these developments participate in the ongoing increase in perceptual switching, especially as it is known from adults that attention affects perceptual switching (Alais, Newell, et al., 2010; Alais, van Boxtel, Parker, & van Ee, 2010; Meng & Tong, 2004; Tong et al., 2006). An increase in the number of switches could also be explained by results showing that children's perceptual processing styles develop from feature based processing to more global processing (Bremner et al., 2016; Doherty et al., 2010; Kovács, 2000; Nayar et al., 2015). Thus, while children's

perceptual style becomes more global, the perceptual system also gets more flexible to switch between global percepts.

Although there were generic age effects, the number of switches was different across task, indicating task specific mechanisms. Specifically, the largest number of switches was observed in verbal transformations, showing that concrete, verbal content stimuli determines a larger number of perceptual switches compared to simple tones. Language processing is a highly exercised skill for humans. Extensive use can lead to strong representations for words, which allows a faster retrieval and processing for words than for simple tones (Rayner & Clifton, 2009).

In the visual modality, the number of switches was larger in ambiguous motion than in ambiguous figure. The result is contrary to what was expected based on the findings of Strüder and Stadler (1999) who found that adults had a larger number of switches for content-based stimuli (such as the duck-rabbit figure) than for abstract stimuli (such as Necker cube). One reason for this is that ambiguous motion is a more animated task and thus attracts greater attention from children, which is known to increase the number of switches (Alais, Newell, et al., 2010; Tong et al., 2006). This suggests that for children, content-based visual stimuli do not necessarily lead to a larger number of switches than non-content based visual stimuli; more important might be the sustained attention a stimulus can attract. Another explanation for this difference is that the ability to perceive global motion is present even in 2 years olds (Yu et al., 2013). Overall, the differences in the number of switches across tasks suggest that children's perceptual switching is affected either by the strength of the stimuli's representations (e.g., verbal forms), or because some tasks (i.e., ambiguous motion) can engage more attentional resources necessary for perceptual switching.

This conclusion consistent with adult findings where similar findings were found (Klink, van Ee, & van Wezel, 2008; Meng & Tong, 2004; Paffen et al., 2006).

After controlling for the age effects, the ability to switch in ambiguous figure correlated with the ability to switch in auditory streaming. The number of switches in the two tasks also correlated positively after controlling for the age effects. The reason for these relationships could be that both tasks posed more difficulties to the perceptual switching system than the other tasks in the first place. Other than these, the number of switches in the other tasks did not relate with each other, which indicates separate underlying mechanisms. This result is consistent with recent findings from our lab that also found separate underlying mechanisms for perceptual switching in a visual and an auditory task (Denham et al., 2018).

Additionally, the ability to switch did not relate consistently to measures of executive functions or creativity, indicating that in children between 6 and 10 years of age the ability to switch is no more related with executive functions as found in Wimmer and Doherty (2011). Similarly, the number of switches across the ambiguous tasks did not show either consistent relationships with executive functions and creativity. The only correlation that maintained after controlling for the effects of age was between verbal transformations and the creativity score. Due to the nature of these two tasks, possibly the correlation is because both tasks involved verbal fluency. However, considering the problem of multiple comparisons, this result could also be only a spurious result (Benjamini, 2010). The lack of consistent correlation between perceptual switching and executive functions or creativity mirrors the results found in adult literature where similar results were found (Chamberlain et al., 2017; Díaz-Santos et al., 2017). Overall, these findings suggest that the capacity to reorganise the interpretations of ambiguous stimuli is task-

specific, more than just modality-specific. This is consistent with prevailing proposals in adult literature that put forward the hypothesis that perceptual switching in response to different bistable stimuli happens in distributed networks across the brain, in a task specific manner (Hupé et al., 2008; Pressnitzer & Hupé, 2006).

In conclusion, the present study has shown that ongoing perceptual switching develops gradually with increasing age in a larger set of visual and auditory ambiguous tasks. The results also indicate that perceptual switching underlines task characteristic mechanisms, a notion in line with the distributed mechanism of perceptual switching proposed and supported in adult studies (Pressnitzer & Hupé, 2006). Overall, these results show that perceptual switching is present from early ages, supported by general maturational processes, while it continuous to develop in a task specific manner.

4 STUDY 3: IS SUPERIOR PARIETAL CORTEX COMMONLY INVOLVED IN VISUAL AND AUDITORY MULTISTABILITY?

4.1 Introduction

In the previous studies, the question of domain generality/specificity of perceptual switching was investigated at a behavioural level. A step further in understanding the relationship between visual and auditory multistability is to see whether they share common neural underpinnings (WHERE). It is currently unknown to what degree there are common brain regions responsible for perceptual switching in both visual and auditory modalities. A potential common brain region involved in both modalities is parietal cortex. It is known that different regions of parietal cortex are involved structuring the sensory information regardless of modality (Gonzalez & Flindall, 2015). Moreover, fMRI studies indicate that regions of parietal cortex are active when participants perform visual or auditory tasks with multistable stimuli. For instance, in a fMRI study, participants' brain activity was measured while they listened to an auditory streaming task and asked to report whether they hear the integrated or the segregated interpretation (Cusack (2005). There was a greater activation in the right intraparietal sulcus when participants heard segregated than when they heard integrated, suggesting that the activity of the parietal cortex is associated with participants' ability to hear segregation. Parietal cortex was also found to be involved in visual multistability (Kanai et al., 2010). Participants viewed an ambiguous structure-from-motion stimulus while their brain activity was measured with fMRI. It was examined how cortical thickness, local grey-matter density and local white matter integrity correlates with perceptual switching rates across participants. Results showed a negative correlation between cortical thickness and percept duration in the right and left superior parietal cortex

and the bilateral postcentral gyrus. That is, the thicker the cortex in these areas, the faster the switching rate. Voxel-based morphometry⁷ of the grey matter density showed a similar correlation between the grey matter density and percept duration in the right superior parietal lobe (right SPL).

The fMRI studies described above measured the brain activity associated with perceptual switching behaviour. However, since fMRI studies cannot show causal effects, it is unclear whether the brain activation manifests as a consequence or as a cause of the perceptual change. Studying patients with lesions in certain brain regions is one of the most important sources of information about the causal roles of specific brain regions in manifesting certain behaviours and cognitive functions. It has been shown, for example, that patients with bilateral lesions of the intraparietal sulcus have impairments in the ability to manipulate mentally the object's features, suggesting that this region is involved in the organisation of the sensory information (Cusack, 2005). Permanent lesions of the parietal cortex in patients with unilateral neglect, destabilises the normal brain functioning, leading to prolonged dominance durations in binocular rivalry (Bonneh, Pavlovskaya, Ring, & Soroker, 2004). Another important tool to use to assess the causal role of a specific brain region for a particular behaviour (including perceptual switching) is by using Transcranial Magnetic Stimulation (TMS). TMS is used to cause "virtual lesions" of the stimulated area, which in turn affects the way a specific task is performed. TMS leads to a performance impairment or improvement of. There are two major types of TMS protocols used to perturb the activity of a specific brain region: online

⁷ Voxel-based morphometry (VBM) is a neuroimaging analysis technique that allows the investigation of focal differences in brain anatomy, using the statistical approach of statistical parametric mapping.

stimulation (while performing a task) or offline stimulation (administered between two experimental sessions).

In the visual modality, the involvement of the parietal cortex in perceptual switching has been demonstrated in several studies using TMS (see Table 4.1). Carmel et al. (2010) employed a binocular rivalry task where participants saw Gabor gratings⁸ oriented to the left for the left eye or to the right for the right eye. An offline protocol with repetitive TMS pulses had an effect of shortening the average phase durations after stimulating the right superior parietal cortex and no effect after stimulating the left homologous site. It was argued that the activity in the parietal cortex has the role of maintaining and stabilising the new dominant percept after the switch happened. Moreover, it was suggested that TMS on right superior parietal cortex leads to a weaker top-down signal, which makes it easier for the suppressed interpretation to become dominant. In the same year, Kanai et al. (2010) applied continuous theta burst stimulation (cTMS) to the right superior parietal cortex, to the left superior parietal cortex, and to vertex (the highest point of the head; as a control site). Transient disruption of the parietal regions decreased the number of perceptual switches in response to an ambiguous structure-from-motion task (i.e., increased the phase durations) compared to vertex, suggesting a direct involvement of the right and left SPL in perceptual switching rates.

The two TMS studies described above posed a controversy. While Carmel et al. (2010) reported that the phase durations increased after TMS, Kanai et al. (2010) found that the phase durations decreased after TMS. To investigate this controversy, a collaborative study (Kanai et al., 2011) was conducted subsequently to understand

⁸ A sine wave grating seen through a Gaussian window (graphical representation of the data). Gabor patches are popular stimuli in vision laboratories because they have characteristics that match the receptive field properties of neurons in primary visual cortex.

whether the discrepancies found were due to the tasks used or due to different TMS protocols. The correlations between the brain structures and the percept durations reported in Kanai et al. (2010) were revised. They made a more detailed region of interest (ROI) analysis of grey matter density in the superior parietal cortex of the data reported in Kanai et al. (2010). Standardised grey matter density was extracted for the site initially stimulated by Kanai et al. (2010). The ROI analysis revealed a positive correlation between the grey matter density and the switching rate in ambiguous-structure-from motion. This result is the opposite of what Kanai et al. (2010) found in a posterior location of the SPL. This result was interpreted to suggest that the structure of anterior and posterior SPL (see Figure 4.1) have opposite roles in the dynamics of multistable perception. After this stage, they applied cTMS to this newfound anterior location and showed that phase durations were shorter after stimulation of the right anterior SPL (aSPL) compared to vertex (the control site). The results of this collaborative study were interpreted to reflect a fractionation of parietal cortex function; different regions within parietal cortex play opposing roles in the control of multistability: right posterior SPL increases the phase durations while right aSPL decreases phase durations.

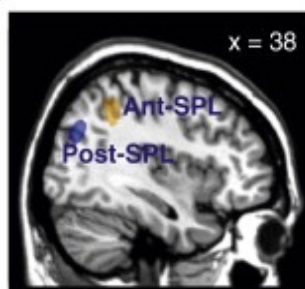


Figure 4.1. Visualisation of the right ($x = 38$) anterior and posterior superior parietal lobe sites stimulated in Kanai et al. (2011).

In another study, Zaretskaya et al. (2010) found that online TMS over right intraparietal sulcus prolonged the phase durations compared to vertex. They provided causal evidence for a destabilising influence of the right intraparietal cortex on perceptual stability in binocular rivalry, similarly to what Kanai et al. (2010) reported for an ambiguous structure-from-motion task. Right SPL did not affect the phase durations as previously found in Carmel et al. (2010). The authors argued that TMS removes attentional resources from the parietal-frontal system, making it less likely to re-select new interpretations of the ambiguous stimulus (leading to an increase in the phase durations). In contrast with Carmel et al. (2010) and Kanai et al. (2011), it was argued that the function of parietal cortex is neither to stabilise nor to select new interpretations. Instead, Zaretskaya et al. (2010) suggest that its role is to optimise the perceptual input, seeking and maintaining access to relevant information and that TMS disrupts this process.

Wood, Schauer, Bak, and Carmel (in preparation) investigated the controversy between the results of Carmel et al. (2010) and Zaretskaya et al. (2010) in two experiments. In Experiment 1 they used orthogonal gratings, while in Experiment 2 they used binocular rivalry with the picture of a face and a house presented to each eye. In both experiments, offline stimulation on anterior SPL shortened phase durations compared to vertex, similarly to the results of Carmel et al. (2010). The effect of online stimulation compared to vertex was not found in either experiment. However, in Experiment 2, a comparison between online-parietal and online-vertex TMS stimulation of the parietal cortex (similar to the protocol of Zaretskaya et al. (2010) led to a lengthening of the phase durations. The authors show that different TMS protocols can indeed lead to different results, suggesting that offline stimulation inhibits the neural activity while online stimulation increases neural

noise. Moreover, the authors state that their findings support the view that right anterior SPL mediates the maintenance of current interpretations rather than generating the switches. They also found that online TMS lengthens the dominance for the face/house but not for the grating stimulus, suggesting a stimulus-dependence effect of TMS in perceptual switching.

Table 4.1

Summary of Four Previous TMS studies that Investigated the Role of the Parietal Cortex in Perceptual Switching

Study	Carmel, Walsh, Lavie & Rees (2010)	Kanai, Bahrami & Rees (2010)	Kanai, Carmel, Bahrami & Rees (2011)	Zaretskaya et al. (2010)
Task	Binocular rivalry	Ambiguous structure-from-motion	Ambiguous structure-from-motion	Binocular rivalry
Main results	Right anterior SPL => TMS shortens dominance durations	Right posterior SPL TMS => lengthens dominance duration	Right anterior SPL TMS => decreased phase duration	Parietal sulcus TMS => Prolonged phase durations compared to vertex
Site stimulated	<u>Right anterior SPL</u> MNI : x = 36, y = -45, z = 51 <u>Left SPL:</u> MNI x = -36, y = -45, z = 51.	<u>Right posterior SPL</u> MNI: x = 34, y = -66, z = 34), <u>Left SPL</u> MNI: x = -21, y = -63, z = 61	<u>Right anterior SPL:</u> MNI: x = 36, y = -45, z = 51	<u>Intraparietal sulcus</u> (a location close to the area stimulated by Kanai et al. (2011))
TMS protocol	Online 1 Hz for 30 minutes	Theta burst stimulation (cTBS)	Theta burst stimulation (cTBS)	Online 2Hz continuous TMS
Number of participants	8	12	8	15

In sum, these studies found that anterior and posterior superior parietal cortex are causally involved in visual multistability, and that these effects are dependent on the TMS protocols and stimuli used.

For auditory multistability no study has applied TMS to assess whether parietal cortex is causally involved. The aim of the present study was, therefore, to investigate whether superior parietal cortex regions—previously found to be involved in the visual modality—are causally involved in the perceptual switching in auditory multistability as well. Specifically, this study examined whether the right posterior (MNI: $x = 34$, $y = -66$, $z = 34$) and anterior parietal cortex (MNI: $x=36$, $y=-45$, $z=51$) have a similar role in visual and auditory multistability. If anterior and posterior superior parietal cortices play similar roles in visual and auditory multistability, we should find the same influence in both visual and auditory tasks. On the other hand, if these regions are not commonly involved in both modalities, the results would speak for modality-specific role of parietal cortex.

As previously the role of parietal cortex on perceptual switching has been investigated with ambiguous-structure-from motion (Kanai et al., 2010; Kanai et al., 2011), this task was more suitable in the current study than the duck/rabbit figure for which no TMS studies have been conducted before. Similarly, the indication that parietal regions are involved in perceptual switching in auditory modality come from studies using auditory streaming task (Cusack, 2005) which makes it the most suitable task to represent the auditory modality in this study. Although in Study 1 of this thesis the highest correlation was found between the two high-level tasks, there was also a correlation between verbal transformations and ambiguous motion, thus there was no strong support for a correlation only for the high-level tasks. For these reasons, in this TMS study it was important to use tasks for which there was

evidence that parietal cortex activity plays a role. Moreover, as there are indications that the neural underpinning of perceptual switching is task specific (Kleinschmidt et al., 2012; Sterzer & Kleinschmidt, 2007; Sterzer et al., 2009), using ambiguous structure-from-motion and auditory streaming was the best option available.

4.2 Method

All the paradigms used previously have tested perceptual switching behaviour using multistable paradigms where participants were instructed to report only the two main interpretations, while the non-frequent interpretations were not actively investigated. However, it has been shown that in both auditory streaming (Denham et al., 2014; Farkas et al., 2016) and ambiguous structure-from-motion (Hol, Koene, & van Ee, 2003) participants report hearing/seeing more than two interpretations. For instance, in auditory streaming, participants can hear a sound that is a combination of the Integrated and Segregated such as “-LH-/---L”, “-HL-/L---”. Similarly, the ambiguous motion sphere can be perceived as moving simultaneously right and left. In the current study, participants were instructed to report a third interpretation (see Section 4.2.4.1.- 4.2.4.2)

4.2.1 Participants

Thirty adults ($M_{\text{age}} = 21.6$, $SD = 6.82$, age range 19–50 years, 14 females) participated in this study. One of the participants was one of the authors of this study. All participants had normal or corrected-to-normal vision and hearing. Seven participants participated previously in studies involving visual bistability in the same lab. All subjects gave written informed consent before participation. The study was approved by the local ethics committee of the University of Edinburgh. Participants received money in compensation for their participation.

4.2.2 Design

A within-group design was adopted in which each participant took part in a training session and three experimental sessions.

In each of the sessions, participants had to report their perceptions in one visual task (ambiguous structure-from-motion) and one auditory task (auditory streaming). Each experimental session consisted of a pre-TMS session, Transcranial Magnetic Stimulation, then a post-TMS run (see Figure 4.2). In each of the three sessions, the stimulation was made on one of the three brain regions: anterior superior parietal cortex (aSPL), posterior superior parietal cortex (pSPL), and vertex (control site) (see Procedure in Section 4.2.3). Overall, each participant was tested in 12 conditions: 3 Sites (aSPL, pSPL, Vertex) x 2 Tasks (Visual and Auditory) x 2 Runs (Pre and Post TMS).

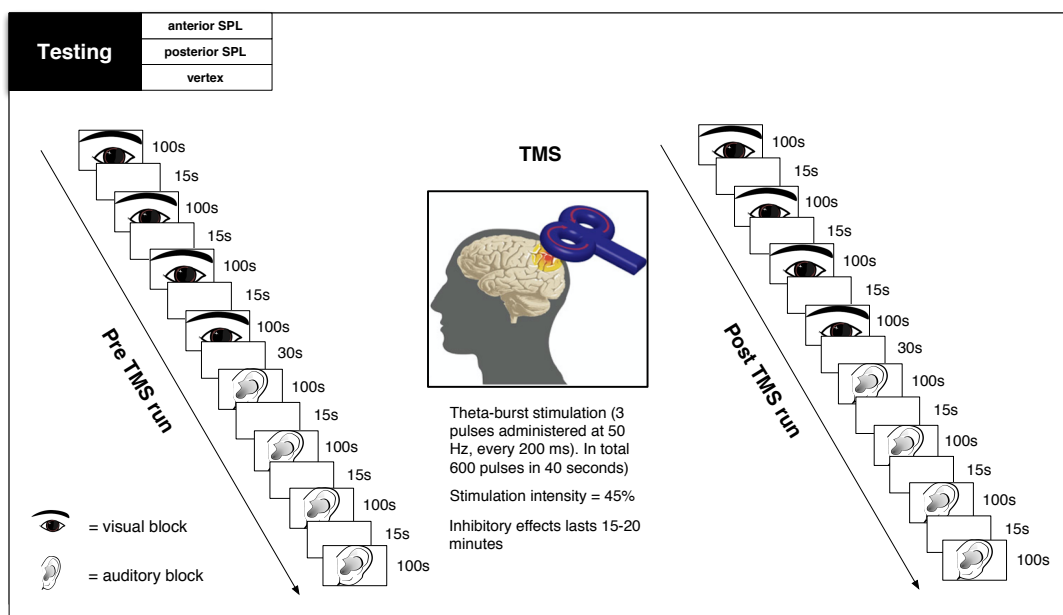


Figure 4.2. Timeline for one experimental session. The TMS stimulation was administered between a pre- and a post-run.

4.2.3 Procedure

The study started with a screening procedure to check the suitability of participants to be exposed to TMS according to the *Guidelines for Non-invasive brain stimulation (NIBS) studies from The University of Edinburgh*. After this stage, participants had MRI scans at the hospital. In the experimental sessions participants were seated in a comfortable chair with their head on a chin rest, 55 cm away from the monitor. The tasks were presented on a Latitude E6440 Dell, 14-inch display (refresh rate 60 Hz). The auditory task was delivered through Sennheiser HD600 headphones (Sennheiser electronic GmbH & Co. KG) using Psych toolbox stimulus presentation software under MATLAB (The MathWorks Inc.).

Two multistable paradigms were investigated: auditory streaming and ambiguous structure-from-motion. Participants came to the lab on four separate occasions, on consecutive days, at least 24 hours apart to prevent TMS carry over effects. Each session lasted approximately 60 minutes. The first session involved training with the visual and auditory tasks. The next three sessions were experimental sessions where TMS stimulation was applied. Each session consisted of sixteen stimulus blocks, where four visual (V) and four auditory (A) were administered in the pre-TMS run (VVVVAAAA or AAAAVVVV) and another four visual and four auditory were administered in the post-TMS run. Each block was 100s long. Between the pre- and post-runs, TMS stimulation was applied on one of the three brain regions: vertex, aSPL, and pSPL. Stimulation was made at the same hour on separate days (with some exceptions where it was difficult to schedule the participants according to the rule). Between blocks there was a 15 seconds break, while between tasks there was a break of 30 seconds. The order of presentation of the visual and auditory blocks was counterbalanced across participants. The order of the

TMS stimulation was counterbalanced for each participant. During the experimental sessions, participants were instructed to listen to the sounds/look at the ambiguous structure-from-motion stimulus and asked to report their perception using the mouse buttons. To check whether the participants maintained focus during the experiment, catch trials (unambiguous interpretations of the stimuli) were appended at the end of each block for brief periods (8 seconds). The timeline of the experimental sessions is summarised in Figure 4.2.

4.2.3.1 Site localisation.

The specific regions in the superior right parietal cortex previously associated with perceptual switches in ambiguous structure-from-motion were anterior superior parietal cortex ($x = 36, y = -45, z = 51$) and posterior superior parietal cortex ($x = 34, y = -66, z = 34$). After participants had MRI scans taken, the two parietal locations were localised with standard MNI brain coordinates on each participant's anatomical MRI scan using the neuro-navigation Brainsight 2 system (Rogue Research Inc., Montreal, Canada). The stimulation site on each participant's scalp was identified with a navigation system using a Polaris infrared camera (Northern Digital Inc., Waterloo, Canada) that was operated to co-register participant's structural scan with their head. The angle of the stimulation was measured on each scan so that the stimulation was made from the smallest distance from the skull.

The vertex was localised using externally visible anatomical landmarks. The midpoint between the nasion (top of the nose) and inion (end of the skull) and halfway between the tragus of the two ears was identified. This midpoint was then marked on a polyester swim cap that participants wore during the stimulation. For the vertex stimulation, the coil was held against the participant's scalp, with the handle pointing straight behind the participant, the coil was held parallel to the floor.

4.2.3.2 TMS protocol.

The parameters for the TMS stimulation were identical to those used in Kanai et al. (2011). The theta-burst TMS stimulation protocol consisted of 3 pulses at 50Hz repeated at 200ms intervals for 40 seconds at 45% stimulator output. One 40s rTMS train was applied in each experimental session, as explained above, in an offline procedure (i.e., separately from the behavioural tasks which were performed before and after TMS). Participants were wearing earplugs during the TMS stimulation. The coil was held manually by a trained investigator, positioned according to the pre-planned stimulation sites trajectories on the individual MRI scans.

4.2.3.3 Perceptual tasks.

Visual stimulus. The stimulus consisted of a revolving sphere with 300 randomly positioned white dots on a black background moving in opposite directions with a sinusoidal speed profile (see Klink, van Ee, Nijs, et al., 2008). The viewing angle of the cylinder subtended 7.1 degrees. The sphere was presented in the centre of a computer monitor at a distance of 55 cm, with a white fixation cross in its centre. It was 3 degrees in diameter, while the dots were 3 arcmin. The position of the dots changed from one frame to the next with a sinusoidal speed profile (a peak angular speed of 57.1 degrees per second.), with roughly half moving leftwards and half moving rightwards (see Klink, van Ee, Nijs, et al., 2008). Each visual (V) stimulus block consisted of a 100 seconds presentation of the moving dot stimulus, with an additional 8-second disambiguated segment appended without a break at the end of each block (see Section 4.2.5 for details of the disambiguated segment).

Auditory stimulus. Sinusoidal tones of 75 milliseconds duration (including 10ms rise and fall times) were arranged according to the auditory streaming paradigm. The hearing threshold was set up to a comfortable level for each participant, which was maintained constant during the experiment. The frequency difference between the two tones was 4 semitones with the ‘L’ tone’s frequency set at 400 Hz and the ‘H’ tone frequency at 504 Hz. The stimulus onset asynchrony (SOA, onset-to-onset time interval) was 150ms. Participants were presented with 100 seconds long LHL- tone sequences. Each auditory (A) block consisted of a 100-second tone sequence of the repeating LHL_ pattern with an additional 8-second disambiguated segment appended without break at the end of each block (see Section 4.2.4.2 for details). The disambiguated segment served as a control condition to check for false positives and false negatives.

4.2.4 Training

4.2.4.1 Ambiguous structure-from-motion.

Due to structure-from-motion effects, the moving dots create the impression of a three-dimensional rotating sphere (H. Wallach & O'Connell, 1953). Because there are depth cues to indicate which dots belong to the front or the back of the sphere, the direction in which the sphere rotates is ambiguous. Viewers can see the sphere moving clockwise or anticlockwise. The stimulus can also be perceived as two half-spheres, moving in opposite directions (Chen & He, 2004; Hol et al., 2003).

Participants were trained to report *Left* if they perceived the front face of the rotating sphere moving leftwards, *Right* if they perceived the front face of the rotating sphere moving rightwards, and *Combined* if they perceived two half spheres moving in opposite directions. These interpretations were demonstrated using disambiguated examples. For the *Left* example, rightward moving dots were

coloured grey; the reduced luminance of rightward-moving dots biases perception in favour of leftward rotation. Similarly, for the *Right* example, leftward moving dots were coloured in grey. *Combined* was demonstrated by colouring leftward moving dots in yellow and rightward moving dots green, to encourage participants to see opposite moving surfaces of similar luminance.

Visual depictions for each interpretation are shown in Figure 4.3. These were used to facilitate the training and later to remind participants about the interpretations.



Figure 4.3. The three interpretations for the ambiguous structure-from-motion (Left, Right and Combined).

4.2.4.2 Auditory streaming.

The tone sequence consisting of a repeating LHL_ pattern can be perceived in different ways. Participants were instructed to report continuously four possible categories of interpretations: a) *integrated* (“LHL-”; pressing the left button of a mouse), b) *segregated* (“L-L-/H---” or “H---/L-L-”; pressing the right button of a mouse), c) *combined* (“-LH-/---L”, “-HL-/L---”; pressing both buttons simultaneously), and d) *none* (releasing both mouse keys). See Figure 4.4 for a visual depiction of the interpretations.

The *integrated percept* emphasised hearing all tones as part of a single repeating pattern. The *segregated percept* emphasised hearing two parallel sound streams; one heard in the foreground, the other in the background, one at slower rate, one at a faster rate. The *combined percept* emphasised the perception of two parallel

streams of sound, at least one of which included a repeating pattern composed of both high and low tones. Finally, the *none response* allowed participants to indicate that they did not hear any repeating pattern or could not decide between the patterns previously described to them (or they were confused).

These interpretations were demonstrated to participants by using disambiguated examples.

The *integrated* example was introduced by reducing the difference between the “L” and “H” tones to 1 semitone (L = 400Hz, H=426 Hz).

The *segregated* example was demonstrated by increasing the difference between the two tones to 10 semitones (L =400 Hz, H= 713 Hz). To demonstrate that either the L_L_L stream, or the H__H__ stream may be perceived in the foreground (with the other in the background) they were also presented with examples in which the intensity and timbre of the background stream were modified (with frequency difference between the H and L tones of 4 semitones); intensity was reduced by 18dB, timbre was changed by adding 2-8 harmonics (with equal weight) to the fundamental.

Combined was demonstrated using the same intensity and timbre manipulation described above for *Segregated* to emphasise either the LH_ or the HL_ pattern. For each LHL_ pattern in the sequence, either the first or second L was manipulated either like LHL_ or like LHL_, which promotes the perception of a foreground HL or LH stream. The visual depictions in Figure 4.4 were used to help explaining the emphasised patterns and later as reminders before each test block.

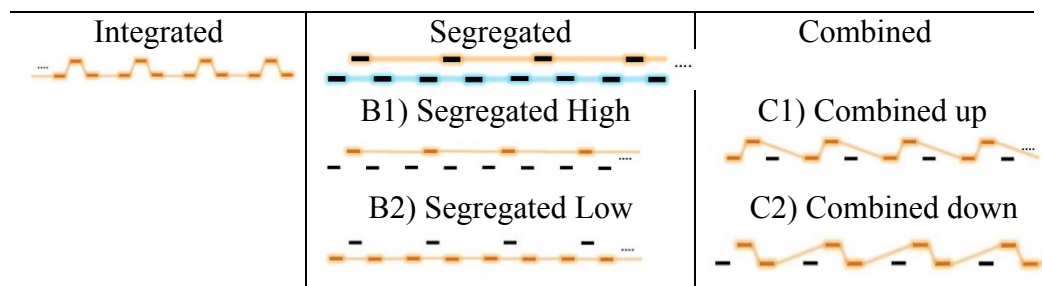


Figure 4.4. A visual illustration of the auditory streaming paradigm (from Farkas et al, 2016). *Note.* The Segregated and Combined interpretations could have been heard in different ways, but participants were instructed to focus only on the main category when they reported their perception.

4.2.5 Training Procedure

Both tasks had similar training procedures. Participants were presented with an ambiguous sound/image for 10 seconds and informed about its ambiguity. Afterwards, they were introduced to each disambiguated version of the stimuli described above and with the mouse keys assignments. This step was repeated until participants understood the perceptual categories and how to report them. The training continued with two types of practice.

1) The ambiguous sound was presented for 60 seconds and participants were asked to categorise their perceptions as if they were doing the main experiment. The timing and the content of the instruction screen sequence were the same as in the main experiment; 1s blank screen, 10s instructions showing visual depictions of the interpretations and key assignments, 2s blank screen, 2s central fixation cross. This was followed by 60 seconds of the ambiguous stimulus, plus a 8-second disambiguated segment which was randomly chosen from *Left, Right (Segregated, Integrated)*. At the end of the practice trial, participants were given feedback on the number of perceptual switches they had made, the proportion of time they reported

for each category and the proportion of time they correctly categorised the disambiguated segment.

2) Participants had to identify correctly the disambiguated versions presented above. There were five 6-9 seconds segments of disambiguated visual and auditory examples presented randomly from *Left*, *Right*, *Combined (Integrated, Segregated low, Segregated high, Combined up, Combined down)*. Participants were given feedback on the proportion of time they had correctly categorised each disambiguated segment. This allowed the experimenter to check whether participants were able to categorise and report their perceptions quickly and accurately. The training continued with these practice trials and further explanations from the experimenter, until participants could report the intended categories with at least 70% accuracy.

4.2.6 Testing

A key press initiated the start of a test block. Participants then saw an instructions screen that reminded participants of the perceptual categories and the key assignment. The 108-second (100 + 8) visual or auditory stimulus was then presented and participants were required to press continuously the key corresponding to their current percept. These perceptual reports were recorded by polling the key status every 10ms for the duration of the stimulus. Participants were instructed to keep either the left, the right or both mouse keys pressed for as long as they continued to hear/see one of the perceptual interpretations, and to switch to another interpretation as soon as their perception changed. They were asked to refrain from trying to voluntarily change their perception to one of the interpretations and to indicate naturally their perception.

4.2.7 Extracting Perceptual Data

The key press data was processed to extract continuous periods during which the same perceptual category was reported. The reports from the disambiguated segments were separately extracted and the proportion of time the participant reported the category corresponding to the intended category (defined by the disambiguation) was recorded. These proportions provide a measure of how well a participant understood the perceptual categories and the key assignment. Any participant who scored less than 60% in both categories was excluded from further analysis. Please note that in this experiment it was possible for participants to report their failure to recognise their current percept by not pressing any key.

For the analysis, median phase durations of each block were computed instead of the switching rates as done in the previous two studies, to allow a direct comparison with the previous published TMS work that has exclusively used phase durations and not the switching rates. Nevertheless, the choice to use in this study the median phase duration instead the switching rates does not change the overall results; phase durations are inversely proportional with the number of switches; the shorter the phase durations, the larger the number of switches and the larger the phase durations, the fewer the number of switches.

4.3 Results

4.3.1 Preliminary Analyses

4.3.1.1 Catch trials analysis.

To check whether participants were paying attention to the tasks, their catch trials performance was preliminarily analysed. The performance for all participants is displayed in Figure 4.5.

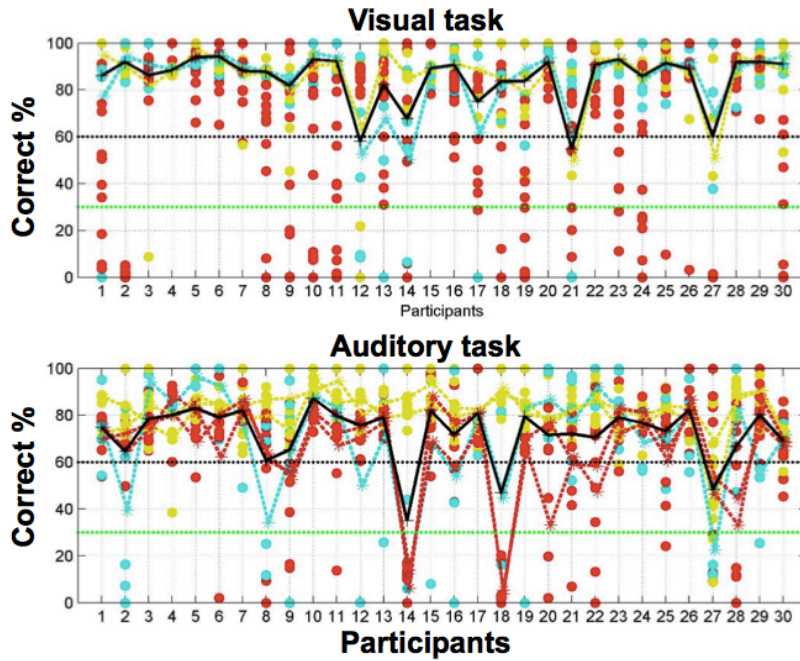


Figure 4.5. Catch trial accuracy performance for the visual task (upper figure) and for the auditory task (lower figure). The dotted lines represent the average percentage of correct responses for each catch-trial-answer category. The blue and yellow colours depict the accuracy for the clockwise/anticlockwise interpretations (for the visual task) and the integrated/segregated interpretations (for the auditory task). The red line summarizes the accuracy for the combined interpretations. The black line is the average across the main interpretations (except the combined).

Participants were excluded if the average of the correct answers was less than 60%. Applying this criteria, five participants were excluded from the subsequent analyses. Another participant was excluded because he/she was an extreme outlier (who predominantly reported *Combined* and in some blocks never switched). Yet another participant was excluded for technical problems. In total, there were 23 participants in the final sample.

4.3.2 Proportion of Time Spent into Each Interpretation

There were 12 conditions altogether: Site (Anterior, Posterior, Vertex) x Modality (Visual, Auditory) x Run (pre-TMS, post-TMS). The proportion of time spent in one interpretation for each TMS condition is displayed below (Figure 4.6). The three interpretations seem relatively well balanced, with the proportion of *Combined* relatively higher compared to previous reports.

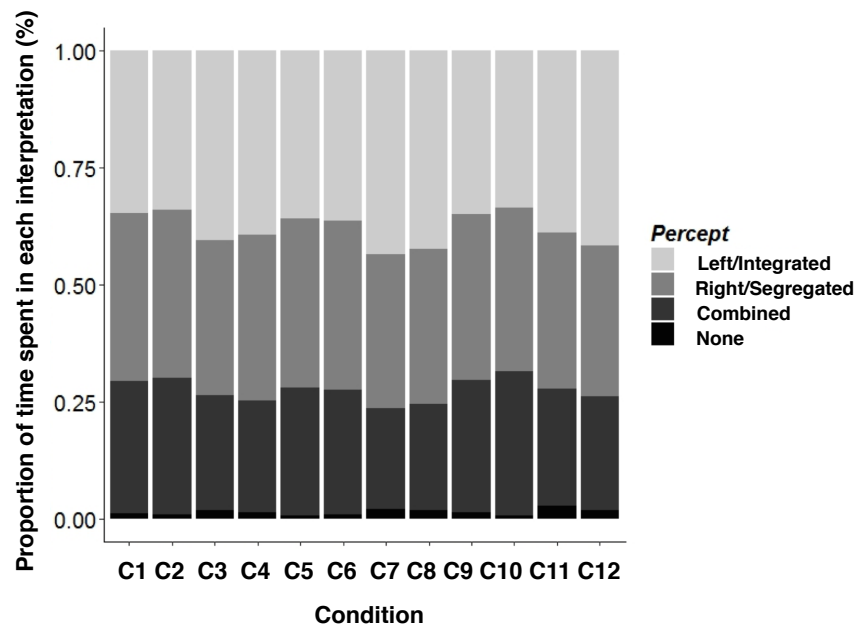


Figure 4.6. The proportion of time spent in each interpretation. The 12 conditions are displayed on the x-axis. The average proportion of time for each interpretation is displayed on the y-axis (from 0 to 1). C1-AntVisPre, C2-AntVisPost, C3-AntAudPre, C4-AntAudPost, C5-PosteriorVisPre, C6-PosteriorVisPost, C7-PosteriorAudPre, C8-PosteriorAudPost, C9-VertexVisPre, C10-VertexVisPost, C11-VertexAudPre, C12-VertexAudPost.

4.3.3 Change in Phase Durations after TMS

The performance on the two tasks (separately for two and three percepts) is displayed in Table 4.2 and Table 4.3.

Median change in the phase durations after TMS was calculated by subtracting the median phase duration of the post-TMS run from the pre-TMS run (post-pre). The analyses were run separately for the two and three percepts to investigate whether they lead to different results. Two mixed model level analyses on the median change were conducted with TMS (Anterior, Posterior, and Vertex) and Modality (Visual, Auditory) as factors and participant ID as a random effect.

For the two percepts the analyses showed no significant effects. The effect of TMS was not significant, $F(2, 110) = .36, p = .70$. The effect of Modality was not significant either, $F(1, 110) = .001, p = .9$. There was no interaction effect, $F(2, 119) = 2.07, p = .13$. See Figure 4.7 for a visualisation of these results.

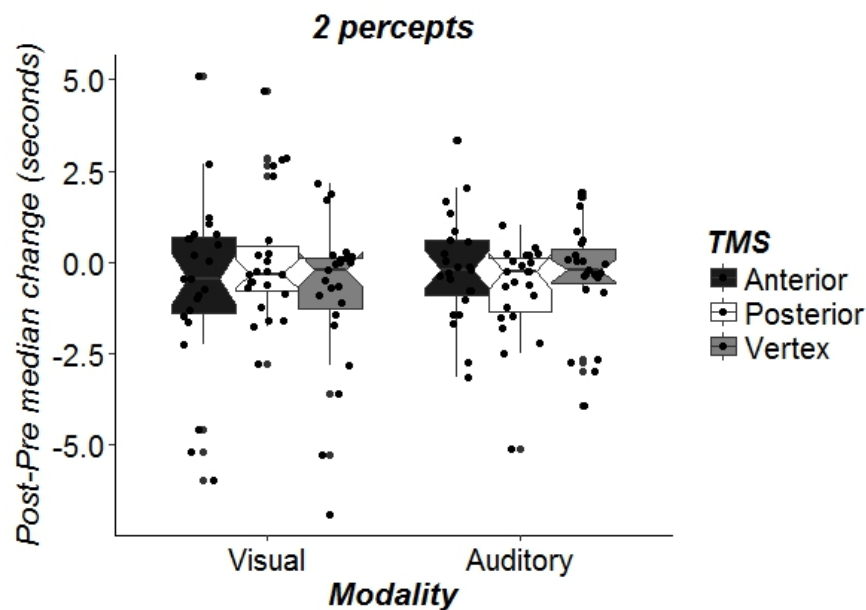


Figure 4.7. Median change duration after TMS on each of the three brain regions stimulated (see agenda), only for the two main interpretations.

Table 4.2

Summary Statistics for Each of the 12 conditions (2 Percepts)

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
Median	4.58	4.38	4.17	3.81	4.20	4.54	4.27	3.79	5.43	4.32	4.60	4.55
Mean	5.08	4.54	4.45	4.28	4.89	5.03	5.18	4.44	5.12	4.29	5.27	4.91
SE.mean	0.53	0.42	0.36	0.42	0.43	0.65	0.59	0.49	0.45	0.40	0.62	0.49
CI.mean.	1.09	0.87	0.75	0.86	0.90	1.35	1.23	1.02	0.93	0.82	1.28	1.02
Std.dev	2.52	2.00	1.73	2.00	2.08	3.13	2.84	2.36	2.15	1.90	2.97	2.37
Skewness	1.17	0.80	0.31	0.98	1.24	1.33	0.65	0.69	0.58	0.21	0.96	0.22
Kurtosis	0.87	0.38	-1.13	0.18	0.93	1.14	-0.61	-0.65	0.27	-0.69	0.02	-1.46

Note. C1-AntVisPre, C2-AntVisPost, C3-AntAudPre, C4-AntAudPost, C5-PosteriorVisPre, C6-PosteriorVisPost, C7-PosteriorAudPre, C8-PosteriorAudPost, C9-VertexVisPre, C10-VertexVisPost, C11-VertexAudPre, C12-VertexAudPost.

Table 4.3

Summary Statistics for Each of the 12 conditions (3 Percepts)

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
Median	4.15	3.92	3.97	3.63	3.62	3.78	4.15	3.49	4.37	3.66	4.55	4.25
Mean	4.07	4.10	4.21	4.00	4.06	4.11	4.71	4.02	4.62	3.94	5.12	4.51
SE.mean	0.29	0.25	0.36	0.41	0.30	0.43	0.52	0.44	0.32	0.26	0.60	0.47
CI.mean	0.60	0.51	0.74	0.86	0.62	0.89	1.08	0.92	0.66	0.53	1.24	0.97
Std.dev	1.38	1.19	1.71	1.98	1.43	2.05	2.51	2.12	1.52	1.24	2.86	2.25
Skewness	0.61	0.89	0.32	0.78	1.11	1.73	0.74	0.87	1.37	0.00	0.64	0.20
Kurtosis	-0.63	0.85	-1.09	-0.42	0.72	2.96	-0.52	-0.29	2.64	-1.29	-0.83	-1.51

Note. C1- AntVisPre, C2- AntVisPost, C3-AntAudPre, C4-AntAudPost, C5-PosteriorVisPre, C6-PosteriorVisPost, C7-PosteriorAudPre, C8-PosteriorAudPost, C9-VertexVisPre, C10-VertexVisPost, C11-VertexAudPre, C12-VertexAudPost.

For the three percepts the analyses showed no significant effects. The effect of TMS was not significant, $F(2, 110) = 1.81, p = .17$. The effect of Modality was not significant either, $F(1, 110) = 1.52, p = .22$. There was no interaction effect, $F(2, 119) = 0.98, p = .38$. See Figure 4.8 for a visualisation of these results.

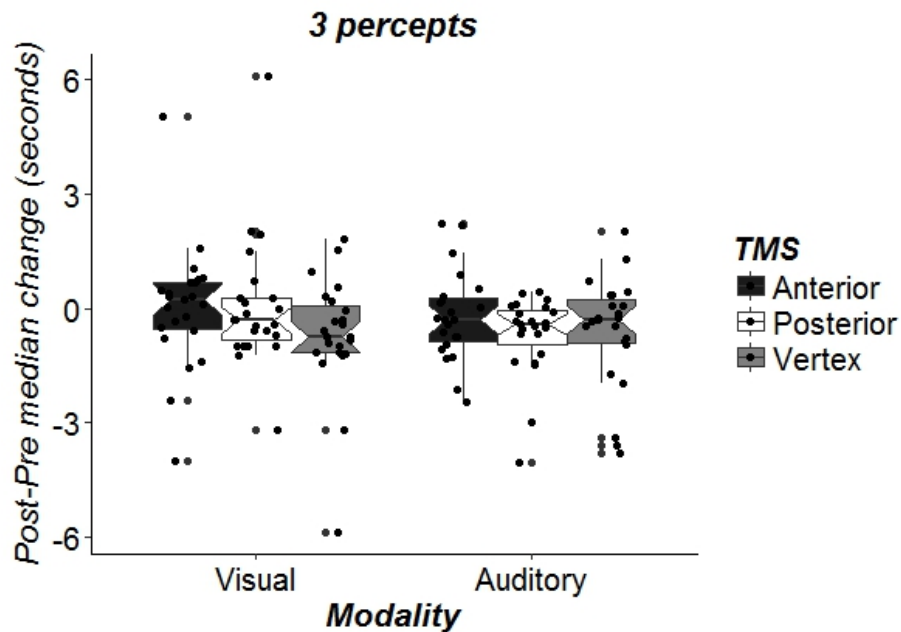


Figure 4.8. Median change duration after TMS on each of the three brain regions stimulated (see agenda), for all three interpretations.

4.3.4 Individual Differences

One reason why there was no significant TMS effects could be the large individual differences across conditions (i.e., for each participant, the pre- and post-TMS runs were different for the three site locations). To investigate this possibility, normalised median phase durations were calculated for each participant and for each condition, separately for two and three percepts. Normalisation was performed in order to control for the possible differences existing in the baseline, in the pre TMS runs. The normalisation was computed with the following formula: $(\text{post-pre})/\text{pre}$, where *post* and *pre* represent the conditions before (pre) and after (post) TMS. The

data are displayed in Figure 4.9 and Figure 4.10 and show large individual differences. Specifically, TMS stimulation did not have similar effects on all participants. The individual differences are high regardless of whether the median phase duration is calculated for the two main interpretations or for all three interpretations.

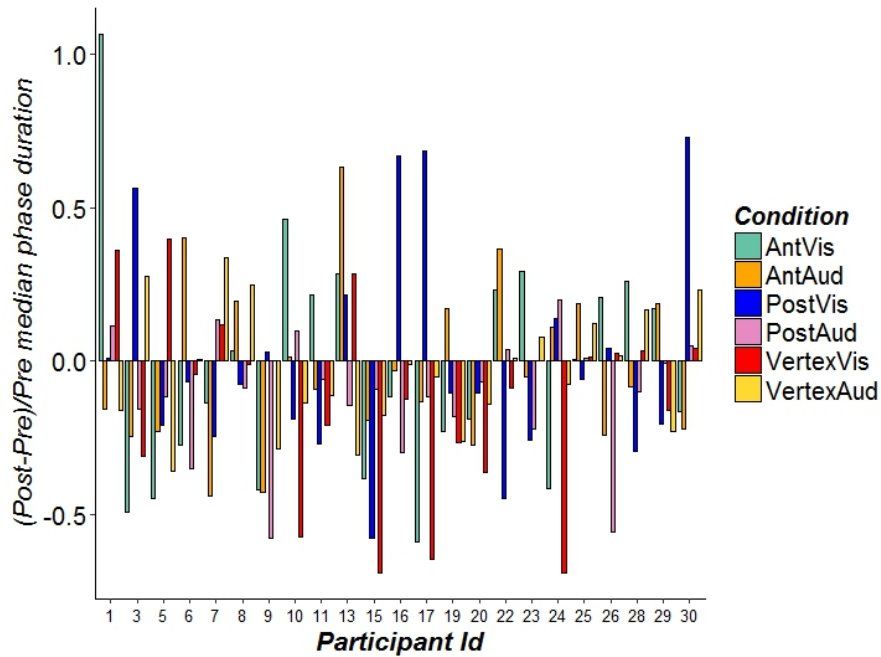


Figure 4.9. Normalised median phase durations for each participant for the main two interpretations. The accompanying agenda represents the six conditions.

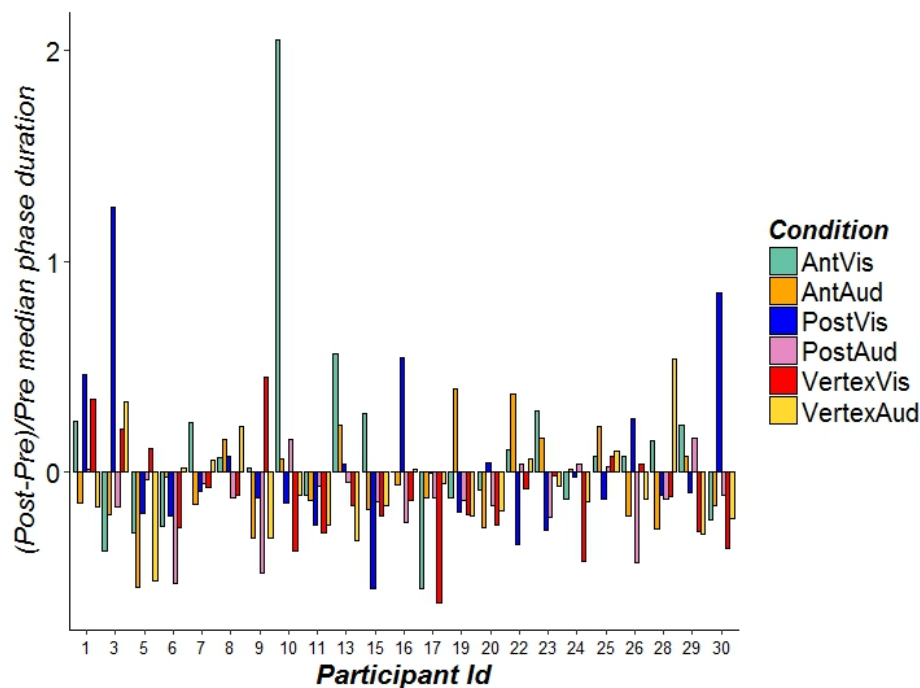


Figure 4.10. Normalised median phase durations for each participant, for all three interpretations. The accompanying agenda represents the six conditions analysed.

4.3.5 Domain General Or Domain Specific?

The question of domain generality/specificity was further investigated in the current study. Spearman correlations between median phase durations for each of the 12 conditions were computed and displayed in Table 4.4 and Table 4.5. Three main aspects of interest were:

- a) Do the phase durations across modalities, in the same TMS condition, in the pre-TMS run correlate (e.g., Does AnteriorVisualPre correlate with AnteriorAuditoryPre)?
- b) Do phase durations across modalities in the same TMS condition, in the post-TMS run correlate (e.g. Does PosteriorVisualPost correlate with PosteriorAuditoryPost)?

- c) Does TMS affect the correlations pre- and post-TMS stimulation in the same condition (e.g., Does AnteriorVisualPre correlate with AnteriorVisualPost)?

The results indicate significant correlations between phase durations across modalities especially in the Anterior SPL TMS condition. In the Vertex and Posterior SPL TMS the correlations did not reach significance. This can be due to a decreased power because of the large number of conditions tested. Nevertheless, these results show that perceptual switching rates in the ambiguous structure-from-motion and auditory streaming are also positively related.

Another result of interest here is the strong correlation between the median phase durations in the pre- and post- in the same TMS condition. Results suggest that TMS stimulation did not disturb the correlation cross modalities between median phase durations in the pre- and post-TMS runs of the same TMS condition.

Table 4.4

Spearman Correlations Matrix between the Median Phase Durations across the 12 conditions (2 interpretations)

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
C1	—	0.693 ***	0.489 *	0.560 **	0.268	0.547 **	0.459 *	0.423 *	0.532 **	0.524 *	0.455 *	0.438 *
C2		—	0.176	0.368	0.262	0.386	0.136	0.149	0.575 **	0.749 ***	0.109	0.148
C3			—	0.785 ***	0.244	0.396	0.923 ***	0.856 ***	0.109	0.355	0.886 ***	0.852 ***
C4				—	0.236	0.360	0.814 ***	0.772 ***	0.236	0.465 *	0.820 ***	0.790 ***
C5					—	0.702 ***	0.367	0.228	0.305	0.492 *	0.292	0.181
C6						—	0.402	0.256	0.587 **	0.590 **	0.515 *	0.450 *
C7							—	0.897 ***	0.009	0.330	0.891 ***	0.876 ***
C8								—	0.064	0.243	0.850 ***	0.865 ***
C9									—	0.495 *	0.131	0.103
C10										—	0.321	0.336
C11											—	0.934 ***

Note. 1- AntVisPre, 2- AntVisPost, 3-AntAudPre, 4-AntAudPost, 5-PosteriorVisPre, 6-PosteriorVisPost, 7-PosteriorAudPre, 8-PosteriorAudPost, 9-VertexVisPre, 10-VertexVisPost, 11-VertexAudPre, 12-VertexAudPost

Table 4.5

Spearman Correlations Matrix between the Median Phase Durations across the 12 conditions (3 interpretations)

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
C1	—	0.391	0.424 *	0.342	0.197	0.384	0.448 *	0.375	0.567 **	0.664 ***	0.473 *	0.531 *
C2		—	0.229	0.339	0.239	0.241	0.149	0.182	0.267	0.553 **	0.087	0.172
C3			—	0.876 ***	0.059	0.336	0.911 ***	0.884 ***	0.203	0.371	0.851 ***	0.907 ***
C4				—	-0.056	0.149	0.811 ***	0.814 ***	0.272	0.248	0.746 ***	0.866 ***
C5					—	0.478 *	0.060	0.064	0.175	0.365	0.066	0.042
C6						—	0.180	0.195	0.304	0.573 **	0.267	0.244
C7							—	0.916 ***	0.184	0.262	0.846 ***	0.913 ***
C8								—	0.201	0.210	0.841 ***	0.875 ***
C9									—	0.637 **	0.279	0.245
C10										—	0.230	0.229
C11											—	0.894 ***

Note. 1- AntVisPre, 2- AntVisPost, 3-AntAudPre, 4-AntAudPost, 5-PosteriorVisPre, 6-PosteriorVisPost, 7-PosteriorAudPre, 8-PosteriorAudPost, 9-VertexVisPre, 10-VertexVisPost, 11-VertexAudPre, 12-VertexAudPost

4.4 Discussion

The main aim of this study was to examine whether the effects of TMS on the anterior and posterior superior parietal cortex (SPL) in visual and auditory multistability are similar or not. Based on Kanai et al. (2011) it was expected that the median phase durations after TMS on posterior SPL would increase. On the other hand, phase durations were expected to decrease after TMS on anterior SPL. The results showed no effect of TMS on the phase durations in either of the two tasks. Surprisingly, the findings for the visual modality by Kanai et al. (2011) were not replicated. Results also showed that the phase durations across the two tasks did correlate in some of the conditions, suggesting that perceptual switching in the low-level tasks manifest domain general patterns (Kashino & Kondo, 2012). However, as the correlations were not consistent across the three site locations, this will be further explored in Study 4.

The non-significant effect of TMS on the phase durations could be explained by the large individual differences found in the study. The phase durations across conditions and individuals manifested large variability, which could have undermined the possibility of finding the real effects of TMS. There could be a couple of reasons for this variability. Firstly, a large rate of individual differences in response to TMS may be due to brain's lateralisation. While the focus here was just on the right SPL, it cannot be ruled out that for some participants left SPL and not right SPL might be the dominant region involved in perceptual switching. Zaretskaya et al. (2010) identified two anatomically distinct parietal regions that appeared in most individual subjects and the group analysis: the right superior parietal lobule (right SPL) and right anterior intraparietal sulcus (right IPS). However, it was reported that although in the group analysis the right hemisphere

was statistically significant, not every subject manifested higher activation in the right hemisphere. In their study, nine subjects tended toward a right-lateralised fMRI response and six subjects tended toward a left-lateralised response, suggesting that the activation in parietal cortex associated with perceptual switching is not exclusively found in the right hemisphere. In the study of Kanai et al. (2011), and in the current experiment, the regions of interest were not determined by observing the actual activation in the brain in response to the tasks. Possibly this is the most ecologically valid approach to take. If brain lateralisation is an important factor for the individual differences manifested, this could have introduced an uncontrolled variability in the current study.

Another reason for the large variability in the TMS effects on each condition could be that participants used different strategies to perform the tasks. When participants are instructed to report neutrally their interpretations, as done in this study, it cannot be ruled out that participants use different cognitive control strategies. For instance, Braver (2012) talks about two kinds of cognitive control that participants can assert on a given task: *reactive and proactive*. Although the distinction is made in terms of cognitive control involved in tasks such as Stroop or Go-No-Go, his model could be relevant here to explain the large individual differences obtained in this experiment. In the reactive control, participants rely upon detection of an interference, which drives the reactivation of the task's goals (i.e., "I see a switch then I press this button"). In this control mode, the aim of the task is not actively maintained; participants only react passively to the given task. The neutral instructions bias tested in this experiment was designed to tap more into a reactive control mode, because participants were told to pay attention to the tasks and report their perceptions after switching happened. On the other hand, when participants are

actively involved in sustaining the goals during the task, a proactive control is engaged. In multistable literature, a proactive control is engaged in conditions where participants are asked to switch as fast as possible or to hold on to a specific interpretation for as long as possible (Farkas et al., 2016; van Ee et al., 2005; Windmann et al., 2006). Although in the neutral condition participants will predominantly exert reactive control, participants could actually use both types of control. This might induce significant individual differences. These two types of control have different neural underpinnings as well. While both involve the activity in the lateral prefrontal cortex, the proactive one involves a much wider brain network, such as medial temporal regions or anterior cingulate cortex. Future research should investigate whether the individual differences are higher in the neutral conditions than in the switch and hold conditions, and to what degree the performance in the neutral conditions could be explained by different strategies used to report the percepts.

Thirdly, multistability is commonly described as involving alternations between two mutually exclusive perceptual states. However, in practice, participants often report different perceptual mixtures, both in visual (Anstis & Saida, 1985; Hol et al., 2003; Knapen et al., 2011) and auditory multistability (Denham et al., 2012; Sterzer et al., 2009). For example, while exposed to the ambiguous structure-from-motion there might be a perception of flashing or of two halves of a sphere that move in the opposite direction toward the front face of the image. In the auditory streaming stimulus, the listeners can perceive more interpretations than the most frequent investigated, the integrated and segregated interpretations (Denham et al., 2012). According to Overgaard (2015), when using subjective reports it is preferable and more accurate to ask participants to report at least three response options

(interpretation A, interpretation B and combined) than only the two most frequent. Similarly, other authors suggest that combined percepts are valid interpretations and should not be disregarded from the analysis (Denham et al., 2012; Denham et al., 2014; Hol et al., 2003; Pastukhov et al., 2013). On the other side, some researchers advise taking measures to minimise the occurrence of combined percepts as they are considered to add noise to the data (Knapen et al., 2011). It is unknown to what degree having to report three alternative interpretations leads to significant differences in perceptual switching compared to having to report two alternative interpretations. If there is a difference based on the number of interpretations that participants are instructed to report, then this could have impacted the TMS results. This possibility will be explored further in Study 4.

Other factors that can explain the variability of responses under TMS could be genetic, age, brain connectivity or ongoing brain activity (Vernet, Brem, Farzan, & Pascual-Leone, 2015). In fact, it has been proposed that the effects of TMS are state-dependent which means that all these factors are mediating the role of TMS because the state of the baseline cortical activation varies from moment to moment, for each individual (Silvanto & Pascual-Leone, 2008).

5 STUDY 4: DOMAIN GENERALITY/SPECIFICITY AND EFFECTS OF VOLUNTARY CONTROL ON VISUAL AND AUDITORY MULTISTABILITY

5.1 Introduction

The previous studies did not show consistent evidence for cross-modal correlations. In Study 1, the number of switches in the two high-level tasks (ambiguous figure and verbal transformations) correlated positively while perceptual switching in the two low-level tasks (ambiguous motion and auditory streaming) did not correlate. In Study 4, the median phase durations in the auditory streaming and ambiguous structure-from motion were related in some of the TMS conditions. The correlations between the two low-level tasks were further examined in this study.

One systematic and informative way to summarise the effect sizes of the previous studies that looked at the correlations between modalities is by running a meta-analysis correlation. Such a method takes into account the effect sizes and the number of participants in each study and calculates an average effect size. Thus, before the current study was conducted, a meta-analysis was conducted on the correlation coefficients reported by Pressnitzer and Hupé (2006), Kondo et al. (2012), and from Study 1 and Study 3 in this thesis. The results are summarised in Figure 5.1 and indicate an overall medium correlation (a correlation coefficient around 0.3). Based on these results in present study it expected to find a positive correlation between perceptual switching across modalities.

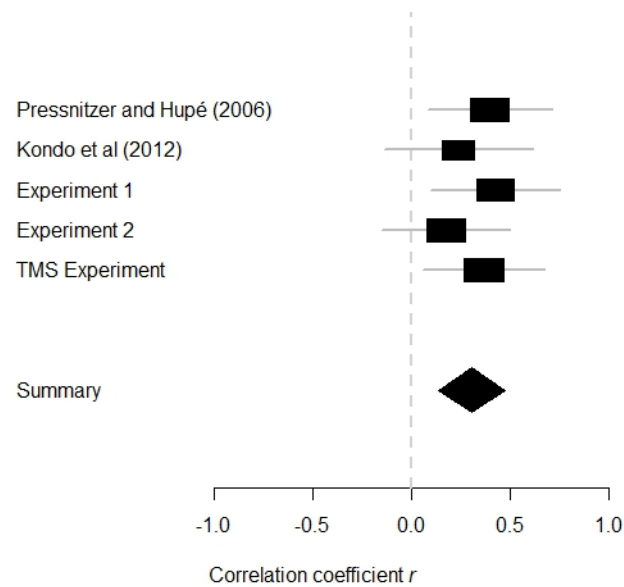


Figure 5.1. Meta-analysis of the correlation coefficients from the previous studies that investigated perceptual switching across modalities. Note that the size of the rectangle symbolises the number of participants—the bigger the rectangle, the larger the N —while the lines on the sides are the confidence intervals. The diamond shape is the overall effect size of the correlation between perceptual switching in visual and auditory modalities based on the previous results.

A further question raised in the TMS study was whether having to report two or three alternative interpretations affects the cross-modal correlations. This was addressed in this study by investigating two groups of participants; in one group participants were instructed to report only the two main interpretations, in the second group participants were instructed to report three interpretations (the two main ones plus the combined interpretations). Finally, in the TMS study, perceptual switching across participants and conditions was highly variable. One possible reason why participants manifested large individual differences could be that they used different attentional strategies when reporting perceptual switching. As such, the current study

investigates whether attentional manipulations have an impact on the correlations across modalities. This was achieved by comparing perceptual switching in a visual and an auditory task in three attentional conditions: *Neutral* (participants had to respond naturally to the tasks), *Hold* (try to maintain each interpretation for as long as possible), and *Switch* (try to switch as fast as possible). If the correlations between perceptual switching cross modalities happen in the neutral condition, as well as in the Hold and Switch conditions, this will bring evidence for a more distributed stimulus driven hypothesis. On the other hand, if perceptual switching cross modality correlates only in the Hold and Switch conditions and not in the Neutral condition, this will support a central system that is sensitive to top-down effects of attention.

5.2 Methods

5.2.1 Participants

Ninety-five adults participated in this study. The study was run at two separate locations: in Hungary at Research Centre for Natural Sciences of the Hungarian Academy of Sciences (RCNS) ($N = 55$), and in the UK at University of Plymouth (UoP) ($N = 40$). Testing at two locations was done for practical reasons (the amount of time and resources needed for data collections was very high) and partly to assure that results are replicable. At each location participants were assigned to one of two groups according to the number of response categories they were instructed with (see Section 5.2.2); RCNS: two-response alternative group (24 adults: 18 females, $M_{\text{age}} = 21.5$, $SD_{\text{age}} = 1.98$), three-response alternative group (31 adults: 21 females, $M_{\text{age}} = 21.48$, $SD_{\text{age}} = 2.08$); UoP: two-response alternative group (20 adults: 16 females, $M_{\text{age}} = 22.9$, $SD_{\text{age}} = 9.52$), three-response alternative

group (20 adults: 13 females, $M_{\text{age}} = 21.5$, $SD_{\text{age}} = 3.05$). All participants had normal or corrected-to-normal vision and normal hearing. Hearing was tested before the experiment started. Participants also completed The Edinburgh Handedness Questionnaire (Oldfield, 1971) (see Appendix G. All subjects gave written informed consent before participation. At each testing location, the study was ethically approved by the local ethics committee: at RCNS by the Unified Committee for Psychological Research Ethics (EPKEB); at UoP by the Faculty of Health and Human Sciences Research Ethics Subcommittee from University of Plymouth. Participants received money or credit points in compensation for their participation.

5.2.2 Design

A between-group design was adopted, participants being assigned to two groups. In the *two-response alternative* group participants were asked to report their perceptions using two perceptual categories, while in the *three-response alternative group* participants were asked to report their perceptions using three categories. Every participant took part in one session with three conditions. For each participant, one session consisted of preliminary assessments, a Stroop task, training with two perceptual tasks, and three experimental conditions interleaved with two supplementary tasks. For reference, the experimental design is summarised in Table 5.1. Each participant received a visual and an auditory task, each with three conditions. Experimental conditions were distinguished by the task instructions given to participants: Neutral (should not try to influence their perceptions), Hold (should try to hold onto each percept they experienced for as long as possible), and Switch (should try to switch to a new percept as quickly as possible). Each condition consisted of eight stimulus blocks, four visual (V) and four auditory (A). In the visual task, participants were asked to report the direction of motion of the front face

of a rotating sphere (ambiguous structure-from-motion), while in the auditory task, they were asked to report on the perceptual grouping of tones in a sequence (auditory streaming). Testing happened in a quiet, dimly lit room, one testing session lasting between 2.5 to 3 hours.

Table 5.1

Experimental Design Summary for Study 4

Stage	Activity	Description
1	Preliminary activities	consent form handedness questionnaire hearing test Stroop test
2	Training (Different for the two response groups)	Response Categories:
	A) Visual task	Two-alternative group LEFT RIGHT
	B) Auditory task	Three-alternative Group LEFT RIGHT COMBINED INTEGRATED SEGREGATED COMBINED
3	Condition 1: Neutral	8 stimulus blocks: VVVVAAAA or AAAAVVVV; order counterbalanced across participants
4	Supplementary activity	Ego-resiliency questionnaire, creativity questionnaire, order counterbalanced across participants (stages 4, 6, 8)
5	Condition 2: Hold/Switch;	8 stimulus blocksOrder counterbalanced with stage 7
6	Supplementary activity	See stage 4
7	Condition 3: Switch/Hold	8 stimulus blocks, Order counterbalanced with stage 5
8	Supplementary activity	See stage 4

5.2.3 Materials and Procedure

At RCNS, the visual stimuli were presented on a Samsung 17" TFT 740B screen with a resolution of 1280 x 1024 pixels, and at UoP on a Dell screen with a resolution of 1920 x 1080 pixels.

Auditory stimulus. The stimulus was similar with the one administered in the TMS experiment. In this experiment the stimulus onset asynchrony (SOA, onset to onset time interval) was 125 ms. Each auditory (A) block consisted of a 180-second tone sequence of the repeating LHL_ pattern with an additional 8-second disambiguated segment appended without break at the end of each block (see description). The sounds were delivered through Sennheiser HD600 headphones by an IBM PC computer using the MATLAB Psychtoolbox at both locations.

Visual stimulus. 500 white dots (each subtending a viewing angle of 4.7 arcmin) were plotted on a black background, with initial position determined by projecting the dots with uniform random distribution onto a virtual sphere, which subtended a viewing angle of 3.3 degrees (a chin rest was used to fix the distance of the head relative to the screen). The virtual sphere rotated around a central vertical axis with an angular velocity of 75 degrees/second at a frame rate of 1000 Hz. The position of the dots therefore changed from one frame to the next with a sinusoidal speed profile with roughly half moving leftwards and half rightwards (see Klink, van Ee, & van Wezel, 2008). Each visual (V) stimulus block consisted of a 180-second presentation of the moving dot stimulus, with an additional 8-second disambiguated segment appended without break at the end of each block (see section 4.2.4.1 for details of the disambiguated segments).

As discussed in Section 2.4.4, the length of each experimental block might be important for the correlations to emerge. In both auditory research (see Denham et al., 2012; Denham et al., 2014) and visual research (see van Ee et al., 2005) block durations are generally larger than three minutes. Considering that longer duration time is required for perceptual switching to stabilise (Suzuki & Grabowecky, 2007), in this experiment each block was 180 seconds long.

5.2.4 Training

The training for both tasks was the same as in the TMS experiment, with the exception that participants in the two-response alternative group were instructed only with two main interpretations: *Left/Right (Segregated/Integrated)*. The third interpretation (*Combined*) in each of the two tasks was never mentioned in this group.

5.2.5 Testing

The main experiment consisted of three conditions defined by three different types of instructions. In each condition, there were four visual blocks and four auditory blocks in the following order: VVVVAAAA or AAAAVVVV, counterbalanced across participants. A key press initiated the start of each block, so it was possible for participants to take short breaks between blocks. The disambiguated segment concatenated to each block was randomly chosen from *Left, Right (Integrated, Segregated)* for both the two- and three-alternative response groups. Note that no disambiguated *Combined* segments were appended to the sequences for either participant group.

5.2.6 Experimental conditions

Conditions were defined by differences in the task instructions; i.e., the presence and nature of the voluntary control participants were instructed to exert over their perception. In the *Neutral* instruction condition, participants were asked to report their perceptions as they occurred without trying to influence them in any way. In the *Hold* condition, participants were asked to faithfully report their perceptions, while at the same time trying to hold onto each percept for as long as possible. In the *Switch* condition, participants were asked to faithfully report their perceptions while at the same time trying to switch to a new percept as quickly as possible. The order of the Hold and Switch conditions were counterbalanced across participants. Instruction screens at the start of each condition explained to participants what they were required to do. At the start of each *Hold* (*Switch*) block, the participants were reminded of their current task.

5.2.7 Procedure

A key press initiated the start of a test block. Participants then saw an instruction screen sequence reminding them of the perceptual categories and the key assignment (and the attentional task in the Hold and Switch conditions). Participants were asked to fixate the cross and move their eyes as little as possible. The 188 (180 + 8)-second visual or auditory stimulus was then presented and participants were required to continually press the key corresponding to the category of their current percept. To keep fatigue effects to minimum, participants were able to take short breaks between blocks and even longer breaks between sessions. Between each of the three conditions, participants stood up from the chair and walked in another room to complete the individual measures tasks (see Table 5.1). Another way in which the

effects of tiredness and fatigue were controlled was by having the catch trials at the end of each block (see Section 4.2.4.1). Participants that did not perform well at these catch trials were excluded from the final analyses. Potential influences of tiredness and fatigue are indirectly analysed in Appendix H. It was hypothesized that in case fatigue has a significant impact on perceptual switching then we should see a difference between the number of switches in Block 1 (when people's vigilance is the highest) and Block 4 (when people get tired). The results did not support this hypothesis.

5.2.8 Individual Differences Tasks

5.2.8.1 Stroop task.

Participants' inhibitory control was measured using a computerised version of the Stroop task (Lansbergen et al., 2007). Words coloured in red, green or blue were presented on the same screen used for the experiment, which subtended a vertical angle of 0.9 degrees and a horizontal angle of 3.3-4.7 degrees (the horizontal angle changing with word length). Participants were instructed to respond as quickly and accurately as possible using the arrow keys on a standard computer keyboard mapped as follows: ↑ for red, ← for blue, and → for green. Stimuli were shown on the screen until one of the response keys was pressed. Each response was followed by a white blank screen for 250 ms.

The task consisted of four conditions. In each condition, there were 60 trials. Condition 1 was the neutral-word condition; the names of the three colours (red, blue and green) appeared on the screen written in black. Participants were required to press the arrow key corresponding to the colour name. Condition 2 was the neutral-colour condition; four X's appeared on the screen in one of the three colours and participants had to press the arrow key corresponding to the colour of the X's.

Condition 3 was the congruent-incongruent condition; colour names appeared on the screen either in the corresponding-colour Congruent trial (e.g., “red” coloured red), or in the two-colours Incongruent trial (e.g., “red” coloured blue). Participants were required to press the arrow key corresponding to the colour of the letters (not the word). There were equal numbers of Congruent and Incongruent trials with an order separately randomised for each participant. The Congruent/Incongruent condition had two blocks of 60 trials.

The Stroop interference effect was measured in two steps. Firstly, the median reaction times of the correct responses in the colour-neutral condition and the word-neutral conditions were averaged to obtain a neutral reaction time. Secondly, the reaction times in the neutral condition (computed above) was subtracted from the median reaction time of the correct responses in the incongruent conditions. Thus, a smaller reaction time difference indicates stronger inhibitory control of a prepotent response. The variable computed in step two was the measure of interest for the analyses with the number of switches.

5.2.8.2 Ego-resiliency questionnaire.

The same questionnaire described in Study 1 (Section 2.2.7) was used here as well. The main measure of interest was the sum of all responses. Cronbach’s α was 0.671 across all participants.

5.2.8.3 Creative behaviour questionnaire.

The paper-based Biographical Inventory of Creative Behaviours (BICB) (Batey, 2007) was used as a measure of individual creativity. BICB is a 34-item questionnaire, in which participants are instructed to indicate whether they have participated in various creative activities (e.g., “*Invented a game or other form of*

entertainment”, “*Composed a poem*”, “*Started a club, association, or group*”).

Participants were asked to tick all the items that applied to them. The BICB score for each participant is the sum of the items ticked. See the questionnaire in Appendix J.

5.3 Results

5.3.1 Preliminary Analyses

5.3.1.1 Catch trials analysis.

To check whether participants understood the instructions correctly their catch trials performance was analysed separately. Any participant who scored an average of less than 30% in one category, or less than 60% over both categories, was excluded from further analyses. Applying this criteria, 33 out of 95 participants were excluded from the analyses based on poor catch-trial matching in the auditory streaming or ambiguous structure-from-motion task. In group-two alternative response, there were 11 participants excluded (four for failing the visual catch trials). In group-three alternative group, 22 participants were excluded (three for failing in the visual catch trials and one for technical reasons).

The sample analysed was based on 33 participants in the two-alternative response group (25 females; 19–25 years; $M_{\text{age}} = 21.45$, $SD_{\text{age}} = 1.82$) and 29 participants in the three-alternative response group (19 females; 19–26 years; $M_{\text{age}} = 21.79$, $SD_{\text{age}} = 2.53$).

The data from for each of the two groups in the two locations was pooled together after it was verified that the number of switches from both locations followed the same trend. Specifically, it was verified that the number of switches in the Switch condition was the highest, followed by the Neutral condition and the

lowest number of switches in the Hold condition. The overall performance for each of the two response groups is summarised in Table 5.2 and Table 5.3 below.

5.3.2 Effects Of Voluntary Control On Multistability

The overall performance across the six conditions in the two-alternative response group is summarised in Table 5.2 and for three-alternative group in Table 5.3.

Table 5.2

Descriptive Statistics for the Number of Switches in Each of the Six Conditions in the Two-Alternative Response Group (N = 33)

	An	Ah	As	Vn	Vh	Vs	Average Auditory	Average Visual
Median	20.50	11.50	26.25	22.25	13.75	32.50	19.58	22.75
Mean	23.14	14.27	29.50	23.74	16.58	34.13	22.30	24.82
SEM	1.94	1.62	2.69	1.53	1.63	3.79	1.79	1.78
SD	11.16	9.32	15.45	8.76	9.38	21.76	10.27	10.20
Skewness	0.94	1.03	0.60	0.57	0.80	1.90	0.79	0.83
Kurtosis	0.30	0.33	-0.51	-0.64	0.01	5.20	-0.11	0.29

Note. An: Auditory neutral; Ah: Auditory hold; As: Auditory switch; Vn: Visual neutral; Vh: Visual hold; Vs: Visual switch.

Table 5.3

Descriptive Statistics for the Number of Switches in Each of the Six Conditions in the Three-Alternative Response Group (N = 29)

	An	Ah	As	Vn	Vh	Vs	Average Auditory	Average Visual
Median	24.50	14.00	30.75	25.25	16.00	36.00	21.92	24.50
Mean	25.14	17.58	32.10	27.38	16.53	35.25	24.94	26.39
SEM	1.97	2.05	3.16	1.85	1.83	2.65	2.08	1.72
SD	10.63	11.06	17.01	9.99	9.88	14.29	11.19	9.25
Skewness	0.59	1.52	2.13	0.50	0.96	0.90	1.52	0.66
Kurtosis	0.45	2.03	6.35	-0.59	1.52	0.58	3.12	-0.69

Note. An: Auditory neutral; Ah: Auditory hold; As: Auditory switch; Vn: Visual neutral; Vh: Visual hold; Vs: Visual switch.

A repeated-measures analysis of variance (rANOVA) was conducted on the number of switches with voluntary control (neutral, hold, switch), modality (visual, auditory) as within factors, and group as between factors (two-alternative vs. three-alternative) (see Figure 5.2). Greenhouse-Geisser correction was applied to control for violation of sphericity. As the distribution of the number of switches was lognormal, a \log_{10} -correction was applied before the analysis was conducted. However, as the analysis without the correction led to the same results, the non-transformed data are presented here. Voluntary control was the only significant effect, $F(2,120) = 58.38, p < .001, \eta^2_{\text{partial}} = .493$, Figure 5.2). Pairwise comparisons with Bonferroni correction showed that the number of switches in the Hold condition ($M = 16.24, SD = 9.85$) was significantly lower than in both the Neutral ($M = 24.85, SD = 10.25$) and the Switch condition ($M = 32.75, SD = 17.34$). The number of switches in the Switch condition was also higher than in the Neutral condition, all $ps < .001$. No other main effects, no between group effects and no interactions were significant, all $Fs < 2.72$, all $ps > .105$.

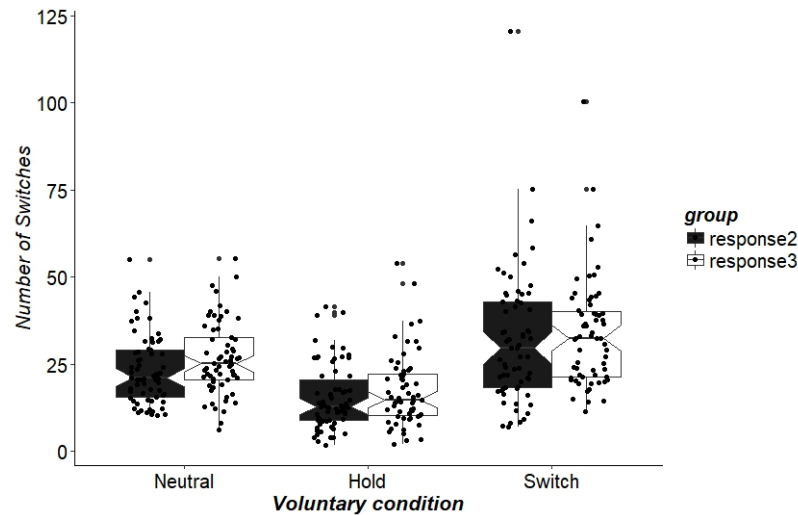


Figure 5.2. The number of switches (y-axis) in the three voluntary conditions (on the x-axis). The accompanying agenda represents the two groups tested. Box = 25th and 75th percentile; the dots over the box plots indicate the data range.

5.3.3 Domain Generality Or Modality Specificity?

To test whether the number of switches across the two perceptual tasks are related, Spearman correlation tests were run between each of the six conditions (Auditory Neutral, Auditory Hold, Auditory Switch, Visual Neutral, Visual Hold, Visual Switch), separately for the two groups tested. The results are summarised in Table 5.4 and Table 5.5 below.

5.3.3.1 Two-alternative response group

Table 5.4

Cross-modal Spearman Correlations between the Number of Switches across the Six Conditions ($N = 33$) in Two-Alternative Response Group

Condition	Spearman	An	Ah	As	Vn	Vh	Vs
An	Rho		0.569***	0.733***	0.443**	0.209	0.518**
Ah	Rho			0.329	0.449**	0.481**	0.144
As	Rho				0.415*	0.108	0.691***
Vn	Rho					0.513**	0.666***
Vh	Rho						0.029

Note. An: Auditory neutral; Ah: Auditory hold; As: Auditory switch; Vn: Visual neutral; Vh: Visual hold; Vs: Visual switch

Two Fisher r -to- z transformation tests were computed to investigate whether the coefficient correlations between the neutral and each of the two attentional conditions differed. The correlation coefficient for the neutral condition ($r = .443$) did not differ significantly from the hold condition ($r = .481$), $z = -.19$, $p = .849$. Neither did the correlation coefficient for the neutral condition ($r = .443$) differ significantly from the switch condition ($r = .691$), $z = -1.45$, $p = .147$.

An average of the three visual conditions was computed to create an overall number of switches for the visual conditions, and a similar computation was done to obtain an overall number of switches for the auditory conditions. A Spearman correlation test between the two overall number of switches showed a strong cross-modal correlation, $r(33) = .587$, $p < .001$.

5.3.3.2 Three-alternative response group

Table 5.5

Cross-modal Spearman Correlations between the Number of Switches across the Six Conditions ($N = 29$), in Three-Alternative Response Group

Condition	Spearman	An	Ah	As	Vn	Vh	Vs
An	Rho		0.528**	0.497**	0.314	0.326	0.310
Ah	Rho			0.565**	0.362	0.586***	0.291
As	Rho				0.305	0.277	0.484**
Vn	Rho					0.513**	0.547 **
Vh	Rho						0.358

Note. An: Auditory neutral; Ah: Auditory hold; As: Auditory switch; Vn: Visual neutral; Vh: Visual hold; Vs: Visual switch

Two Fisher r -to- z transformation tests were computed to investigate whether the coefficient correlations between the neutral and each of the two attentional conditions differed. The correlation coefficient for the neutral condition ($r = .314$) did not differ significantly from the hold condition ($r = .586$), $z = -1.25$, $p = .211$.

Neither did the correlation coefficient for the neutral condition ($r = .314$) differ significantly from the switch condition ($r = .484$), $z = -0.73$, $p = .465$.

An average of the three visual conditions was computed to create an overall number of switches for the visual conditions, and a similar computation was done to obtain an overall number of switches for the auditory conditions. A Spearman correlation test between the two overall number of switches showed a strong cross-modal correlation, $r(30) = .489$, $p = .006$.

A Fisher r -to- z transformation test was also computed to investigate whether the correlation coefficients for the neutral conditions across the two groups are different or not. The results revealed that the two coefficient correlations were not significantly different $z = .59$, $p = .55$. This suggests that the relationship across the two modalities in the neutral conditions in both groups is not different, indicating that the insignificant result for the neutral condition in the three-alternative group might be due to sample size and sampling population.

Three partial correlations were further conducted between the number of switches in the two modalities in each voluntary condition while controlling for the group effects. Even after controlling the effect of group, the correlations between the number of switches in the visual neutral condition and the auditory neutral condition remained significant in all three attentional control conditions: neutral [$r(59) = .415$, $p = .001$], switch [$r(59) = .586$, $p < .001$] and hold [$r(59) = .60$, $p < .001$].

5.3.4 Relation Between The Number Of Switches, Ego-Resiliency, Inhibition, And Creativity.

Descriptive statistics for the individual differences are summarised in Table 5.6.

Table 5.6

Summary Statistics for Each of the Three Additional Measures, separately for the two groups tested

	Two alternative group (N =33)			Three alternative group (N = 29)		
	Ego Resiliency	Creativity	Stroop	Ego resiliency	Creativity	Stroop
Median	41.00	7.00	143.00	41.00	7.00	62.00
Mean	41.36	7.58	141.48	40.03	7.97	85.12
SEM	0.72	0.56	13.68	0.95	0.73	13.99
SD	4.15	3.21	78.59	5.12	3.91	75.36
Skewness	-0.32	0.65	0.67	-0.61	0.63	1.97
Kurtosis	-0.52	-0.67	0.15	-0.22	-0.21	4.35

Note. For each of the three measures, the variable of interest is defined in Section 5.2.8.

Spearman correlations between the number of switches and inhibition, ego-resiliency and creativity were run separately for each group response. For the two-alternative group no relationships were found (Table 5.7) while the for three-alternative group (Table 5.8) there was a positive correlation between creativity and Visual Switch.

Table 5.7

Spearman Correlations between the Number of Switches across the Six Conditions and Individual Differences Measures for the Two-Alternative Response Group

	Spearman	Stroop	Ego Resiliency	Creativity	An	Ah	As	Vn	Vh	Vs	auditory	visual
Stroop	Rho	—	-0.275	-0.206	0.036	0.128	-0.174	0.035	-0.087	-0.091	-0.015	-0.039
Ego Resiliency	Rho		—	0.368	-0.151	-0.135	-0.060	-0.260	0.057	-0.287	-0.132	-0.193
Creativity	Rho			—	0.038	-0.153	0.074	-0.063	-0.061	-0.049	-0.012	-0.036

Table 5.8

Spearman Correlations between the Number of Switches across the Six Conditions and Individual Differences Measures for the Three-Alternative Response Group

	Spearman	Stroop	Ego Resiliency	Creativity	An	Ah	As	Vn	Vh	Vs	auditory	visual
Stroop	rho	—	-0.011	-0.037	-0.104	-0.013	-0.232	-0.142	-0.073	-0.056	-0.139	-0.087
Ego Resiliency	rho		—	0.398 *	0.046	-0.158	0.069	0.123	-0.029	0.253	0.011	0.180
Creativity	rho			—	-0.232	-0.326	0.195	0.259	-0.151	0.421 *	-0.040	0.268

5.4 Discussion

The aim of this study was to bring stronger evidence on whether the number of switches in visual and auditory multistability correlates at the individual level and to what extent attentional control or the number of response categories instructions influences this relationship. It was found that perceptual switching across modalities correlates strongly, regardless of the attentional control condition and the number of response categories participants were asked to report. This supports that there are strong similarities between perceptual switching regardless of modality.

Results showed strong correlations between the number of switches in the visual and auditory tasks in each of the three conditions (Neutral, Switch, Hold). Compared to the insignificant correlation results of Pressnitzer and Hupé (2006), in this experiment the number of switches between modalities correlated strongly. Probably, the lack of significance in Pressnitzer and Hupé (2006) was due to a small sample size. Small sample sizes and short stimulus presentation could be responsible also for the non-significant results between the number of switches in the low-level tasks from the Study 1 in this thesis. The strong positive correlations found here regardless of the voluntary condition suggest that similar functional mechanisms control multistability in vision and audition.

The number of switches in each modality was higher in the Switch condition, followed by the Neutral and then by the Hold condition, where the number of switches was the lowest. This result is in line with previous findings showing that perceptual switching can be controlled voluntarily, in either the visual (Farkas et al., 2016; van Ee et al., 2005) or auditory modality (Farkas et al., 2016; Billing et al., 2018). It was hypothesised that in case the correlations are higher in the Switch and Hold rather than Neutral condition, this will support a common influence of top-down attentional influences on perceptual switches. Although the correlation coefficient for the number of switches between modalities was the highest in the Switch

condition, followed by Hold and then Neutral, these were not statistically significant, indicating that the commonalities observed are unlikely due to arise principally from some attention-related top-down effect. In line with previous findings, it is possible that there are common principles of functioning at a more low-level sensorial level. For instance, Paffen et al. (2006) showed that attention can speed the number of switches in a binocular rivalry task, but has no inherent control over the switching processes. In recent computational work of Cao et al. (2016) perceptual switching was investigated across a variety of multistable paradigms in vision and audition. They showed convincing evidence that the timing of perceptual decisions shows invariant dominance distributions (i.e., the characteristics of the distribution are similar). This also supports perceptual switching in all multistable paradigms share common principles of functioning. The authors found also that the distributions of the phase durations in multistable paradigms are less variable and more skewed than the distributions of reaction times in a choice task. Although some authors (see Gallagher & Arnold, 2014) are sceptical to the fact that distributions of phase durations in multistability reflect intrinsic characteristics to the phenomenon, it seems that phase durations distributions are indeed a reliable statistical summary of the multistable phenomena as it differs from other response time choices. These results have implications for understanding the neural underpinning of multistable phenomena, considering that they follow closely the “intrinsic stochastic dynamics of neocortical activity, which is dominated by connected local assemblies, such as cortical columns or clusters of columns” (Cao et al., 2016). Recently, Denham et al. (2018) analysed phase durations at a much finer detail and found that there are in fact differences across perceptual switching across modalities. Firstly, it was found that although the distributions of the phase durations were lognormal, they were significantly different (as examined with a two-sample Kolmogorov Smirnov test). To understand further the underlying differences, two parameters of each lognormal distribution were compared: μ (which determines the mean of the distribution) and σ (which determines the variance of the data or standard deviation). While the μ parameter was no different between modalities, the σ

parameter differed between the two distributions, suggesting that the distributions of the phase durations have modality-specific characteristics despite being lognormal. Evidence for distinct characteristics of the phase durations across domains were also shown in the analyses of the successive phase durations. Specifically, the correlations between successive phase durations were higher in the auditory task than in the visual task, showing that in the auditory modality there was a higher dependence between successive phases. This indicates that there are modality specific characteristics at more detailed levels of analyses.

Additionally, the results indicated no difference in the number of switches between the two-alternative and three-alternative response group, suggesting that the mechanisms involved in the frequency of perceptual switches are similar when participants are asked to report two versus three alternative interpretations. Therefore, the non-significant TMS results cannot be explained only by the fact that participants were asked to report three interpretations instead of two. Nevertheless, the number of participants incorrectly categorising the disambiguated segments appended to the end of the test blocks was much higher in the auditory three-response task than in the visual three-response task. This difference may have affected the correlations in the cross-modal comparisons.

Finally, the correlation results with the additional measures further showed no common influence of inhibition, creativity, or ego-resiliency on the number of switches. This suggests that although there is a positive correlation between modalities, this is not due to a central influence of inhibition, creativity, or ego-resiliency. These results are consistent with recent studies which show that perceptual switching in ambiguous motion or Necker cube does not relate to inhibition or cognitive flexibility (Chamberlain et al., 2017; Díaz-Santos et al., 2017). A lack of consistent correlation with creativity is also consistent with the lack of consistency in previous findings. Thus, while creativity was previously related with perceptual switching or the ability to switch in various visual tasks (Doherty & Mair, 2012; Wiseman et al., 2011; Wu et al., 2016), in the

auditory domain, perceptual switching in auditory streaming was not related with creativity (Farkas et al., 2016). Overall, although it is possible that the ability to switch shares common grounds with creativity (measured with divergent thinking tasks or creative potential questionnaires), the current findings suggest that creativity (measured with a creative potential questionnaire) is not related to the number of perceptual switches in either of the two modalities. It is also possible that in general, the relationships between perceptual switching and executive functions, creativity or personality are task and stimulus specific.

6 GENERAL DISCUSSION

6.1 Cross Modal Commonalities

As outlined in the introduction, it is an on-going debate whether there are commonalities between perceptual switching behaviours in visual and auditory multistability. While some studies report that the switching rate in visual and auditory tasks correlates (Kondo et al., 2012), other studies identified no correlations across modalities (Hupé et al., 2008; Pressnitzer & Hupé, 2006). To clarify this dispute, the aim of this thesis was to investigate to what extent perceptual switching behaviour in the visual and auditory modality is domain general or specific. The aim was approached through three angles of questioning: the HOW, WHAT and WHERE of multistability, in three studies on adults and one study on children.

In Study 1, perceptual switching behaviour in two visual and two auditory tasks was investigated in adults. Based on studies that show differences across tasks (Kashino & Kondo, 2012; Struber & Stadler, 1999; van Ee, 2005; Wolf & Hochstein, 2011), it was reasoned that one aspect that could clarify the dispute over the correlations from previous studies is to look at stimuli that differ in their complexity and which require different cognitive demand levels (WHAT). For this reason, perceptual switching in one visual high-level task (ambiguous figure) and one auditory high-level task (verbal transformations) was compared to perceptual switching in one visual low-level task (ambiguous motion) and one auditory low-level task (auditory streaming). The results of Experiment 2 in Study 1 showed positive correlations between the perceptual switching rates across modalities, especially in the high-level tasks (ambiguous figure and verbal transformations), but also between ambiguous motion and verbal transformations, indicating the correlations found are not only between tasks with semantic content. Moreover, these correlations emerged when the stimuli were presented for 120 seconds, but not when they were presented for 60 seconds (in Experiment 1).

The correlation between the number of switches in low-level processing tasks (ambiguous motion and auditory streaming) was further investigated in Study 3 and 4. In Study 3, in some of the conditions, the median phase durations across modalities correlated positively when stimuli were presented for 100 seconds. Moreover, Study 4 showed that, with longer stimulus durations (180 seconds) and more participants, strong correlations were consistently found, indicating common functional mechanisms in perceptual switching across modalities. It is possible that in Study 4 the commonalities were captured also between the two low-level tasks because perceptual switching was assessed in longer viewing/listening stimulus presentation compared to Study 1 and Study 3. For instance, Suzuki and Grabowecky (2007) showed that the dynamics of perceptual switching changes with increasing stimulus presentation. When participants are presented with a binocular rivalry task, perceptual switching is slow in the first 20-seconds. When the stimuli are presented for 1-2 minutes the switching rate increases, reaching a stable rate when the stimulus is presented at least for 3 minutes. Thus, possibly in Study 1 (Experiment 2), the switching rate increased in a task specific way, while in Study 4, with enough stimulus presentation, the switching rate was stable across both tasks, allowing the commonalities to be captured. This difference in the stimulus duration should be considered by future studies that look into finding commonalities cross modalities.

Another possible reason why in Study 4 the correlations were more consistent could be the number of participants (the highest from all the previous studies). Finally, another possible reason why the correlations were the strongest when the stimulus duration was higher could be the effects of fatigue or tiredness. Perceptual switching can be affected by fatigue and tiredness as any other psychological factors investigated in experiments. Because there is no research that looked specifically into the effects of fatigue and tiredness on perceptual switching, this concern is rather difficult to address. In the current study, fatigue factors were controlled by allowing participants to have many breaks between blocks/trials. Adding the catch trials at the end of each block was another way to control that participants paid attention to the tasks and to eliminate

those that did not report their perception accurately. Finally, the possible effects of fatigue were investigated in Appendix H by looking at how the number of switches changes from Block 1 (when participants start the task and their vigilance is higher) to Block 4 (at the end of the task, when participants should be more tired). If tiredness had a significant influence on the switching rates, we would expect a significant difference between the first and the last block. As discussed in Appendix H the results did not support this hypothesis, the switching rate in the first and last block were not statistically different.

Thus, whereas the correlational analyses between the switching rates across modalities were not fully convincing in Study 1 and Study 3—possibly due to the short stimulus presentation and number of participants—the results of Study 4 showed strong positive correlations. This suggests that perceptual switching across modalities share common functionalities, regardless of the level of processing, which supports the results found by Kondo et al. (2012). The strong correlations challenge the conclusions from Pressnitzer and Hupé (2006) who—based on the lack of correlations—stated that “perceptual switching is implemented independently across sensory modalities” (p. 1351). I suggest that evidence for distributed mechanisms does not come from the lack of correlations because, as it was found in this thesis and reported in Denham et al. (2018), the correlations do exist. Evidence for distributed mechanisms were found with more detailed analyses as described in Denham et al. (2018) which showed there are subtle differences between perceptual switching across modalities in the analyses of the phase durations distributions.

In Study 4, it was further tested to what extent these commonalities are dependent on whether participants are asked to report two- or three-alternative interpretations and if voluntary attentional control gives rise to stronger correlations. The results indicated that the number of switches in the two-alternative group was not different from the number of switches in the three-alternative group. It has been previously found that multi-stable perception—in contrast to bi-stable perception—offers the possibility to make better predictions for computational models that look into understanding the mechanisms behind perceptual switching dynamics (Wallis &

Ringelhan, 2013). However, the results in Study 4 showed that cross modality correlations were not different if participants were asked to report two or three interpretations. This indicates that the results found in Study 3 are unlikely to be due to the instructions to report three interpretations instead of only two.

Another aim of this thesis was to see if the common neural underpinnings of visual and auditory multistability lie in parietal cortex. In a recent literature review on the role of the fronto-parietal cortex on perceptual bistability, it was suggested that parietal cortex is the strongest candidate brain region to be associated with perceptual switching in vision (Brascamp et al., 2018). In the auditory modality, the involvement of parietal cortex is still highly disputed. Cusack (2005) and Hill et al. (2011) found that intraparietal sulcus is involved in the perception of bistable auditory stimuli. However, recent findings did not find strong support for the involvement of parietal cortex in auditory bistability (Sanders, Winston, Barnes, & Rees, 2018). On the contrary, they found that activity in the temporal and frontal areas showed the strongest correlation with perceptual switches in auditory streaming task. Results from Study 3 did not support this hypothesis. Surprisingly, the previous findings on the effects of TMS on parietal cortex in visual multistability were not replicated either. One reason why previous results were not replicated could be that the sample size was too small to test all the conditions and that participants manifested large individual differences, reducing the statistical power to find an effect of TMS. Therefore, it is possible that the results in Study 3 are false negatives. Whether or not parietal cortex is commonly involved in perceptual switching across modalities remains an outstanding question and needs to be addressed in high-powered studies.

The analyses on the differences across the four tasks tested show further that beside the commonalities found there are task specific characteristics. For instance, perceptual switching rates were the highest for verbal transformations, indicating that language-based multistability is quantitatively different from multistability determined by figure/ground, motion, or stream sound segregation. These differences can be due to the fact verbal forms are much easier to extract

because of the extensive use of language that allows faster retrieval and processing for words than for example, simple tones (Rayner & Clifton, 2009).

6.2 Theoretical Implications Of The Findings

What does the strong positive correlations between the number of switches in ambiguous structure-from-motion and auditory streaming mean for the dispute between the central versus the distributed hypothesis? On the basis of previous reports (see Carter & Pettigrew, 2003; Kondo et al., 2012; Pressnitzer & Hupé, 2006; Shannon et al., 2011) a strong positive correlation would be evidence for a common mechanism that controls multistability regardless of paradigm and modality. What form this common mechanism has is still unknown.

One view is that these commonalities are due to top-down modulations of voluntary attentional control. Attentional effects on multistability have been shown in both visual (Strüber & Stadler, 1999; van Ee et al., 2005) and auditory multistability (Billig, Davis, & Carlyon, 2018; Snyder, Gregg, Weintraub, & Alain, 2012). Study 4 confirms that voluntary control affects perceptual switching in the same way, regardless of modality. Moreover, although the correlation coefficients across perceptual switching in the two modalities were higher in the switch and hold conditions, they were not significantly different from the cross-modality correlations found in the neutral condition (where voluntary control was not present or as strong). Therefore, the strong correlations exist even in the absence of the attentional bias. This suggests that the top-down effects in the form of attentional voluntary control cannot fully explain the strong correlations found. What could then explain the relationships between multistability across modalities?

One explanation comes from computational models that seek to explain the mechanisms by which perceptual switching happens. Computational models can in fact reproduce the alternation patterns of perceptual switching behaviour without the involvement of attentional control mechanisms (Moreno-Bote et al., 2010; Noest, Van Ee, Nijs, & Van Wezel, 2007; Pastukhov & Braun, 2007). These models propose that multistability is due to activity within distributed neural

hierarchies that are in charge of resolving conflicting sensorial input. In these models, attention does not determine the competition but can only bias the competition by affecting which of the interpretations are processed faster and for how long (Pressnitzer & Hupé, 2006; N. Rubin & Hupé, 2004).

Most supported models suggest that the time course of multistability is due to three neuronal ingredients: inhibition, adaptation, and neural noise (Hock, Schöner, & Giese, 2003; Huguet, Rinzel, & Hupé, 2014; Kalarickal & Marshall, 2000; Klink, van Ee, Nijs, et al., 2008; Rankin, Sussman, & Rinzel, 2015). According to these models, the representations of the different interpretations come from separate pools of neurons exerting mutual inhibition on one another. The activity in these pools of neurons allows for one of the representations to predominate over the other. While one of the interpretations is dominating, adaptation takes place until the dominant neural pool diminishes its suppressive grip. The flip occurs and the other representation dominates. Between successive transitions, various sources of noise (e.g., neural noise) introduce variability and instability (Brascamp, Van Ee, Noest, Jacobs, & van den Berg, 2006). The involvement of these basic components is supported by findings in psychophysics (Alais, Newell, et al., 2010; Blake & Logothetis, 2002), brain imaging (Weilnhammer, Stuke, Hesselmann, Sterzer, & Schmack, 2017), and computational models (Noest et al., 2007; Pastukhov et al., 2013; Wilson, 2007).

It is yet unknown to what degree the three factors act in a similar fashion across modalities. In a recent study, Cao et al. (2016) proposed that perceptual switching across a range of multistable paradigms follows common principles of functioning at low levels of sensory processes. In their computational work, they investigated the statistical proprieties among six multistable paradigms (five visual and one auditory). Although perceptual switching rates differed considerably across paradigms, the relationships between the mean phase durations and the variance of these durations (i.e., coefficient of variation) remained consistent. Thus, it was showed that the phase durations are affected by “proportional noise” which means that slower

switching is affected by less noise and faster switching by more noise. These proprieties were further compared with the phase durations of other reaction time tasks but they differed, supporting that multistability has specific characteristics. The authors proposed that perceptual switching is similar across paradigms because of up and down activity of populations of neurons (such as cortical columns or clusters of columns). These microcircuits are organised in a highly distributed system, where specific brain regions are involved depending on the task.

Another explanation for the common mechanism is offered by predictive coding theory (Friston & Stephan, 2007). Generally, this theory has an epistemological approach to multistability, questioning why human brains manifest the behaviour of switching back and forth between different alternatives of an unchanging stimulus (Hohwy, Roepstorff, & Friston, 2008). The theory proposes that human brains are engaged in probabilistic unconscious perceptual inference about the causes of the sensory input. There is increasing support for both visual (Hohwy et al., 2008; Weilhhammer et al., 2017) and auditory modalities (Denham & Winkler, 2006) that predictive coding theories can explain multistable phenomena using the same underlying mechanisms of perceptual inferences. According to this theory, the commonalities in perceptual switching across modalities are explained by the fact the brain is engaged in common inferential processes regardless of modality. It is suggested that sensory input is initially processed in the early brain regions (such as visual cortex or auditory cortex) and that higher-level brain regions (such as frontal or parietal cortex) generate hypotheses about the most likely cause of the sensorial input. The difference between the top-down prediction and the bottom-up input gives rise to a prediction error (i.e., what is unexplained from the top-down predictions by the bottom-up signal). Top-down predictions are constantly updated to minimise the prediction error. In case the stimuli are ambiguous, the process of minimising the prediction error is continuous. The difference between the present interpretation and the alternative one exerts constant pressure to alter the current hypothesis. Using predictive coding theory assumptions, Kanai et al. (2011) proposed that anterior SPL is involved in generating the predictions (i.e., the

current interpretation). Perturbing the activity of anterior SPL should lead to an impairment in the ability to generate strong predictions, which increases the number of switches. On the other hand, posterior SPL is considered to be involved in generating the prediction error that increases the probability of a perceptual switch. TMS on this region should slow the switching rates. Further replication studies are needed to confirm if parietal cortex does indeed play the role suggested previously. This is important especially because the effects on posterior SPL were not found consistently across studies (see Sandberg et al., 2016). The results in Study 3 in this thesis did not support either the proposal of Kanai et al. (2011).

Recent neuroscientific studies challenge the association between the fronto-parietal activity and perceptual switching by asking whether the activation is a cause or an effect of the perceptual switch awareness. Non-report studies use modified versions of the binocular rivalry paradigm and tracked perceptual alternations using signals from the eyes. Results showed that that the frontal regions play a role in the act of reporting perceptual events rather than in the perceptual transitions (Frassle, 2014; Brascamp, 2015, Zou et al, 2016), suggesting that multistability happens mainly in the low-level brain regions, outside the brain's executive system. These results support the view that high-level brain regions are not involved in the competition between the alternative representations but in the act of reporting perceptual switching. Nevertheless, although the use of binocular rivalry is very widespread—and much of what is known about multistability comes from its investigation—it is arguable to what extent the findings can be generalised to other paradigms such as the ones used in the current work (Meng & Tong, 2004; Tong, 2001).

Finally, it is also possible that the correlations found are due to contextual effects such as individual motor strategies or subjective decisional criteria (Gallagher & Arnold, 2014; Hupé et al., 2008). For instance, one participant might follow the instructions closely and report all perceptual changes in real time, but he can also report transitions that did not occur because he anticipated the perceptual switches. Another participant might wait for the new percept to persist

for a longer period before reporting its existence. Participants' responses can be similar across modalities because of these confounding variables that could be used consistently. For instance, Gallagher and Arnold (2014) suggested that the correlations are most likely due to the participants' tendency to over-report or under-report perceptual switches. However, their study used mainly visual tasks. It would be important to explore this further by investigating whether these tendencies are the solely cause of the commonalities found and if they are modality and/or tasks specific.

6.3 Perceptual Switching In Development

Investigating how perceptual switching across modalities develops between 6 and 10 years of age was another aim of the thesis. The results show that overall, the ability to switch is present in all tasks by the age of 6 and it does not develop significantly between the age of 6 and 10. On the other hand, the number of switches increased with age in each of the tasks. Thus, as the increase in the number of switches is not task or modality specific, it seems that more general developmental processes are responsible for the maturation of perceptual switching. There are a few speculative ideas about these changes. Firstly, the increase in the number of switches with age might be due to the continuous development of the pathways of information flow between the fronto-parietal cortices and sensory modalities. This interpretation is in line with recent findings on the developmental trajectory of visuospatial perception. Dekker et al. (2017) found that the protracted development of visuospatial perception from 6 to 12 years of age is not limited to the tuning of the neuronal populations in visual cortex. Instead, this development reflects a more efficient use of the spatial information available in the visual system when making decisions by the higher-level mechanisms. This suggests that developments in higher order brain mechanisms influence how efficient the information in the visual system is processed. An increase in the number of switches could also be explained by the gradual changes

from local processing styles to more global processing styles (Bremner et al., 2016; Doherty et al., 2010; Nayar et al., 2015).

The increase in the number of switches with age is also consistent with the predictive coding theories on multistability, which emphasises that different regions of the fronto-parietal cortex has the role of mediating the predictions and prediction error signals. For instance, in children, the prediction error signals might undergo maturation, which leads to stronger prediction errors and to the prediction signals, being interpreted more efficient/faster, which then leads to an increase in the switching rates. However, these are only speculations, as a predictive coding approach to multistability has never been tested in children.

The findings of Study 2 show task specific effects of perceptual switching in children. Children reported fewer perceptual switches in ambiguous figure and auditory streaming than in ambiguous motion and verbal transformation, most probably due to particularities of the perceptual tasks that make some of the tasks easier. Moreover, there were no correlations found across tasks and the individual differences did not relate consistently with perceptual switching across modalities/tasks. Overall, these findings suggest that the capacity to reorganise the interpretations of ambiguous stimuli is task specific, more than just modality specific. This result is consistent with findings from adult literature which suggest that perceptual switching is generated in distributed networks across the brain in a task specific manner (Denhman et al, 2018; Hupé, Joffo, & Pressnitzer, 2008; Pressnitzer & Hupé, 2006). All these are potential explanations for the mechanisms behind the increase in the number of switches with age and with task specificity findings. The mechanisms behind these findings need to be address specifically so we could go from the description of the phenomena to the understanding of the underlying principles of functioning.

6.4 Individual Differences In Executive Functions, Creativity And Personality And Perceptual Switching.

A further aim of this thesis was to investigate whether perceptual switching across modalities is consistently related with individual differences in executive function (inhibition and set-shifting), creativity, and personality. These measures were selected because they have been investigated previously in relationship to perceptual switching, but never in a more systematic way across a variety of tasks/modalities.

6.4.1 Executive Function

In the search for the causal mechanisms of perceptual switching, it has been proposed that executive brain regions such as frontal cortex play an important role (especially in the visual domain). Support for this hypothesis comes from evidence showing that patients with lesions in the prefrontal cortex (causing impairments of executive functions) are less able to recognise and intentionally switch between two possible interpretations of ambiguous figures, compared to healthy subjects (Windmann et al., 2006). Further support comes from studies where fronto-parietal activation was elicited only by genuine perceptual switching and not by the presentation of an unambiguous stimulus sequence that matched the reported perceptual experience (Brascamp, Blake, & Knapen, 2015; de Graaf et al., 2011; Frässle, Sommer, Jansen, Naber, & Einhäuser, 2014; Knapen et al., 2011). Consequently, research with children proposed that inhibitory control is necessary for the ability to switch in visual tasks (Wimmer & Doherty, 2011). However, the relationship between executive functions and the switching rate has not been found consistently in vision and audition. Inhibitory control does not appear to affect the number of switches experienced by adults in ambiguous structure-from-motion (Chamberlain et al., 2017) or Necker cube (Díaz-Santos et al., 2017). On the other hand, inhibitory control was negatively associated with perceptual switching in the auditory streaming (Farkas et al., 2016).

The results on adults from this thesis reflect similar inconclusive results as the previous findings discussed. Inhibitory control related only with the switching rates in response to verbal transformations (Study 1, Experiment 2), suggesting that a better inhibitory control is associated with a faster switching rate in verbal transformations. In Study 4 the number of switches in the two low-level tasks were not associated with inhibitory control.

Inhibitory control did not relate with the number of switches in children either. Moreover, the ability to switch as measured by Wimmer & Doherty was not related either with inhibitory control in children between 6 and 10 years of age. The ability to switch did not increase with age and by the age of 6 the majority of children could switch. It seems therefore that once the ability to switch emerged, inhibitory control is not contributing anymore to the ongoing perceptual switching. Toppino and Long (2004) made a distinction between the experience of *ambiguity* (that the same input can produce more than a single interpretation) and of *reversibility* (changing perceptions of the same input over time). It seems that inhibitory control plays a role on in the experience of ambiguity and not in reversibility. Wimmer and Doherty (2011) suggested that children develop the ability to understand the concept of ambiguity between 3 and 4, and that they are able to switch to the second interpretation only from 4/5 years of age. Once the competing internal representations are established (i.e., ambiguity is solved and switching happened for the first time), switching back and forth between alternative interpretations might function in a relatively automatic manner, which means that the number of switches over an extended viewing/listening may be due to very different, non-cognitive factors (Toppino & Long, 2004).

The current findings on the role of inhibitory control are also supported by neuroscientific findings on the role of frontal cortex on perceptual switching. Patients with frontal lobe damage with impairments in executive functions also show impairments in the ability to switch in ambiguous visual tasks (Lumer et al., 1998; Windmann et al., 2006). Thus, overall, it seems that

inhibitory control—mediated by activity in the frontal cortex—is more involved in the ability to switch than in the number of switches.

6.4.2 Set-shifting

Set-shifting, which is the ability to switch between tasks or mental sets, is another executive function which relates to the activity in the fronto-parietal cortex (Miyake et al., 2000). As set-shifting involves switching between mental sets and perceptual switching involves the alternation between interpretations, their relationship was tested empirically but the findings are not consistent. One study on children suggested that set-shifting measure with DCCS is a key process for the ability to switch in ambiguous figures tasks (see Bialystok & Shapero, 2005). However, in another study with children, set-shifting measured with DCCS (Frye et al., 1995) did not predict the ability to switch (Wimmer & Doherty, 2011). Set-shifting was also unrelated to the number of switches in adults in ambiguous motion cylinder (Chamberlain et al., 2017) or Necker cube (Díaz-Santos et al., 2017). On the other side, set-shifting measured with a verbal fluency task (Troyer et al., 1997) correlated positively with the proportion of integrated percepts (the non-dominant interpretation) in auditory streaming (Farkas et al., 2016), suggesting that better set-shifting abilities leads to a larger number of switches in auditory streaming. All these studies used a variety of ambiguous tasks and measures of set-shifting, making it hard to draw a conclusion of whether or not set-shifting has a domain general role in perceptual switching.

In this thesis, the role of set-shifting was tested in Study 1 on adults and Study 2 on children. In either of the two studies set-shifting was not related with the number of switches. Moreover, in the child study set-shifting did not relate with the ability to switch either.

6.4.3 Creativity and Personality

Switching between different representations in divergent thinking tasks and different interpretations of an ambiguous image have been described as similar human experiences since the early 20th century by Gestalt theories. Moreover, both phenomena involve perceptual and

mental restructuring and are thought to rely on similar processes (Schooler & Melcher, 1995). In the introduction, it was described that the relationship between the two phenomena has been predominantly found in visual tasks (Doherty & Mair, 2012; Wiseman et al., 2011), while in the auditory modality creativity was not related with perceptual switching in auditory streaming. In Study 1 and 2 of this thesis, creativity measured with a divergent thinking task was investigated in relation to perceptual switching in ambiguous figure, ambiguous motion, verbal transformations, and auditory streaming. In Study 4, perceptual switching in ambiguous motion and auditory streaming was investigated in relation to creativity measured with a creative achievement self-assessed questionnaire. No relationship was found between creativity and perceptual switching across modalities, neither in Study 1 nor in Study 4. Similarly, in children, after controlling for the effects of age, the number of switches and the ability to switch did not relate either with creativity scores. It seems therefore that creativity and perceptual switching do not relate consistently across tasks/modalities and that this relationship is most likely highly task dependent.

In previous research, the relationship between personality factors and perceptual switching across various multistable paradigms was not consistent either. In the auditory modality, Farkas et al. (2016) found that personality factors such as ego-resiliency are related to switching rates in auditory streaming. In the visual modality, the previous findings were inconsistent. While some authors found a relationship between personality factors such as introversion and perceptual switching (i.e., the number of switches) in ambiguous figures (Lindauer & Reukauf, 1971), others found that perceptual switching (i.e., median phase durations) in binocular rivalry is not related to personality (Bosten et al., 2015). For a more systematic exploration, this thesis investigated ego-resiliency in relationship to the switching rates across modalities and tasks. This is a more systematic approach to measuring its relationship with ego-resiliency. The results from both Study 1 (Experiment 2) and Study 4 showed no relationship between the number of

switches and ego-resiliency. Possibly the results are dependent on the stimuli characteristics and personality measured used here.

6.5 Conclusion

To summarise, perceptual switching in visual and auditory multistability correlated positively, supporting the common mechanism hypothesis. As TMS did not affect perceptual switching across modalities, these commonalities do not seem to be due to the same neural underpinning in parietal cortex. Moreover, attentional control does not explain either the commonalities found, indicating a more low-level common mechanism or functionality. Furthermore, no central influence of inhibitory control and creativity was consistently associated with perceptual switching regardless of task/modality. It is possible that once the ability to switch develops, non-cognitive factors determine the rate of switching. Overall, the lack of consistent association between ambiguous perception tasks and measures of executive function and creativity suggests that the capacity to reorganise the interpretations of ambiguous stimuli is task specific, more than just modality specific. Although ongoing perceptual switching increases with age, task specific characteristics were found in both children and adults. This is consistent with proposals in the adult literature which suggest that perceptual switching is generated in distributed networks across the brain, in a task specific manner (Denham et al, 2018; Hupé, Joffo, & Pressnitzer, 2008; Pressnitzer & Hupé, 2006). Moreover, it suggests that these distributed networks are involved in perceptual switching in a similar way from early ages.

7 SUMMARY CONTRIBUTIONS OF THIS THESIS

The main contribution of this work is the systematic investigation of the relationship between perceptual switching in visual and auditory multistability. The question of domain generality/specificity in visual and auditory multistability was approached through three perspectives: the *How*, *What* and *Where* of bi-/multi-stability.

Perceptual switching relationships between visual and auditory tasks was analysed to see if they support a common or a distributed hypothesis (HOW). This question is perhaps one of the most fundamental and long-standing question regarding perceptual bi/multi-stability: are the switches between percepts controlled by central, high-level mechanisms, or by distributed competition within sensory representations. To see if perceptual switching is different across stimuli with different cognitive demands levels, two tasks were investigated in each modality with stimuli that have or do not have semantic content (WHAT). A neuroscientific approach was further used to explore whether the same neural underpinnings are involved in perceptual switching across modalities (WHERE). The role of voluntary attentional control on perceptual switching in both modalities was studied to investigate whether it plays a central top-down role in perceptual switching across modalities. Finally, a developmental approach was employed to understand how perceptual switching in the four tasks manifests in children and whether executive functions and creativity—that continue to develop in children—affect perceptual switching.

Firstly, in Study 1 (Experiment 2) there was a positive relationship between perceptual switching in ambiguous figure and auditory streaming (the high-content tasks). In Study 3 and 4 positive correlations were found between perceptual switching in the low-content tasks (ambiguous motion and auditory streaming). This contributes to the initial dispute between the results of Pressnitzer and Hupé (2006) and Kondo et al. (2012) and brings stronger evidence that the number of switches in vision and audition are determined by common factors. As the commonalities were found even between the two low-level stimuli, it seems that semantic content of the stimuli is not the only factor that explains the commonality found. Moreover, the relationships across modalities were not dependent on voluntary control influences or how many interpretations participants were asked to report. The results have implications for the understanding of basic perceptual mechanisms of disambiguating the environment around us. Results suggest that switching between interpretations in visual and auditory ambiguities share common mechanisms, which might further indicate that our perception (regardless of modality) is

intrinsically flexible to build and re-build representations from the visual and auditory input in our environment.

Secondly, the analyses on the impact of different cognitive demand levels from Study 1 and Study 2 indicated that especially in the auditory tasks, the content of the stimuli has a significant influence on the number of switches. These findings extend the understanding of the influence of the stimuli's semantic content on perceptual switching (Strüber & Stadler, 1999) and show that in audition, the stimuli with semantic content have a larger number of switches than non-semantic stimuli. In vision, on the other hand, a stimulus with semantic content does not always produce the largest number of perceptual switches. Experiment 2 with adults did not find any difference in the number of switches across the two visual tasks. With children, on the other hand, the results were the opposite of what was expected; ambiguous motion (low content-based) had a larger number of switches than ambiguous figure (high content-based). Thus, in visual bistability, the attention that the stimuli can attract might be more important than semantic content. Overall, results from both children and adults show differences across perceptual tasks, indicating that disambiguating the ambiguities in our environment is a highly task-specific skill, supporting findings which show that perceptual switching is generated in distributed networks across the brain in a task specific manner (Denhman et al, 2018; Hupé, Joffo, & Pressnitzer, 2008; Pressnitzer & Hupé, 2006).

Thirdly, the thesis produced the first systematic study of perceptual switching across a range of visual and auditory tasks in 6-, 8-, and 10-year-old children. This contributes to the general understanding of perceptual switching phenomena by showing that although there is an increase in the number of switches across ages, these developments are modality and task independent. It was found that executive functions and creativity do not relate consistently with the number of switches across all four paradigms investigated. This is an important finding because it suggests that cognitive factors do not play a role in the ongoing perceptual switching, while it might be

more important for the ability to switch. The distinction between the ability to switch (measured once) and the switching rates (measured continuously) is something that future studies should address more closely as it could confirm whether fronto-parietal cortex is involved only in the ability to switch (Ricci & Blundo, 1990) and not in determining the number of switches (Brascamp et al., 2015; Frässle et al., 2014; Knapen et al., 2011).

Fourthly, for the first time, the casual role of superior parietal cortex in both visual and auditory multistability was investigated using TMS. Parietal cortex was not involved in perceptual switching across modalities as expected. This finding contributes as a failed replication of previous findings in visual modality, which supported that anterior and posterior parietal cortex increases or decreases the median phase durations. The current findings raise an important question about which results are valid and require further investigations.

Finally, methodologically, the thesis indicates that the stimulus duration is important for finding the commonalities across modalities and discusses the effects of the number of blocks when studying multistability. Furthermore, it successfully adapted adult methodologies on a child population, being the first within-subjects study that investigated both visual and auditory tasks.

APPENDICES

Appendix A

Violation of data assumptions before and after log-transformation

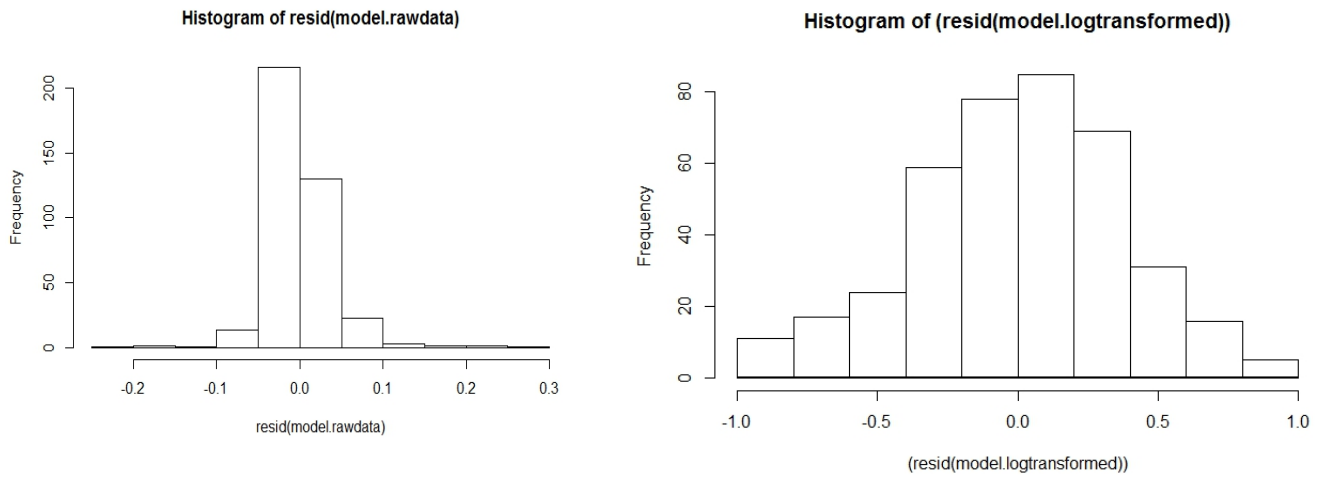


Figure 1.A. The distributions of residuals before and after log-transformation of the switching rates.

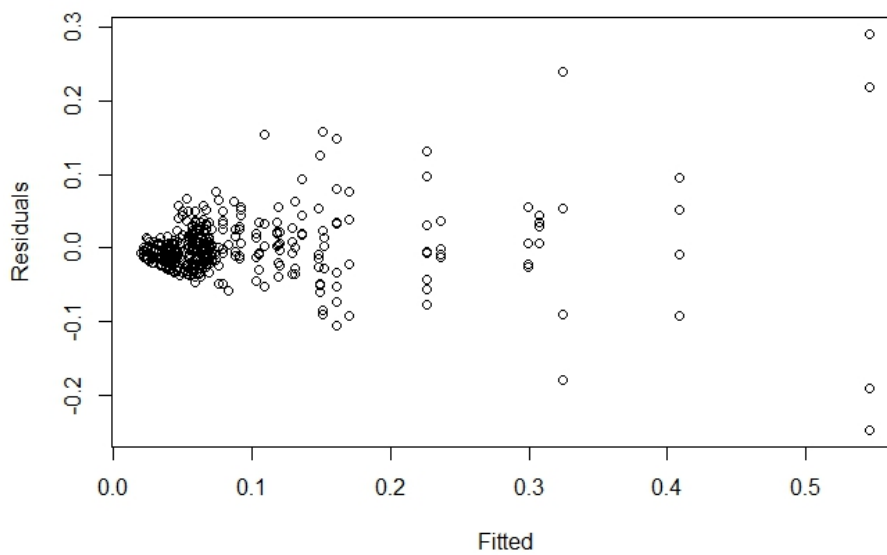


Figure 2.A. Heterogeneity of the switching rates before log-transformation.

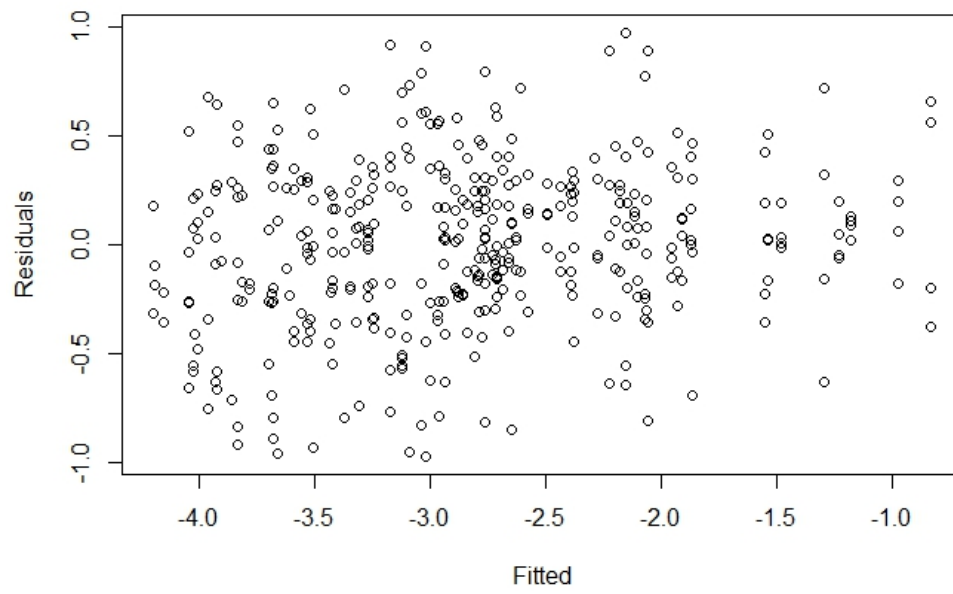


Figure 3.A. Homoscedasticity of the switching rates after log-transformation.

Appendix B

Spectrograms for the sounds used in Study 1 (Experiment 1) and Study 2

Figure 1.B below depicts the wave form and the spectrogram for the word “Fly” used in Experiment 1.

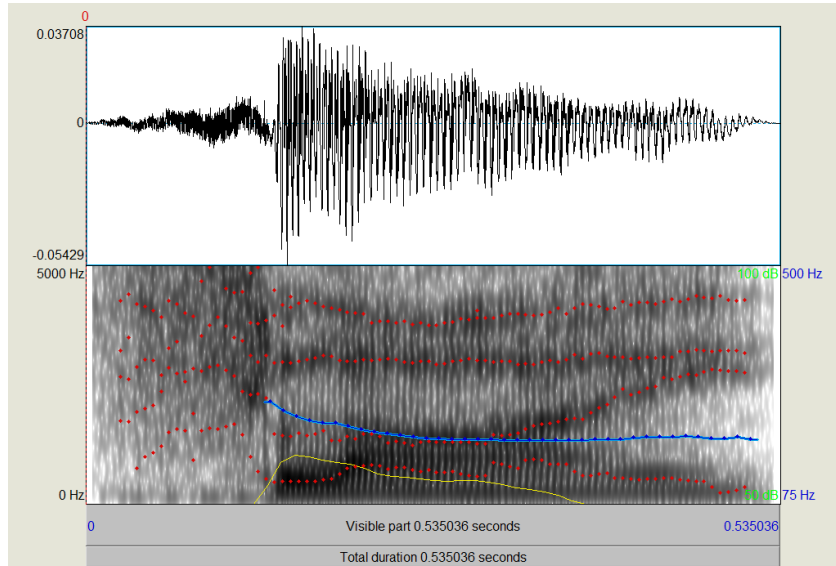


Figure 1.B. “Fly” word wave form (top) and spectrogram (bottom). The yellow line represents the intensity contour, the red dots represent the formant contour (the concentrations of acoustic energy), and the blue line represents the pitch contour.

Figure 2.B depicts the wave form and the spectrogram for the sounds used to create the auditory streaming task in Study 1 and Study 2.

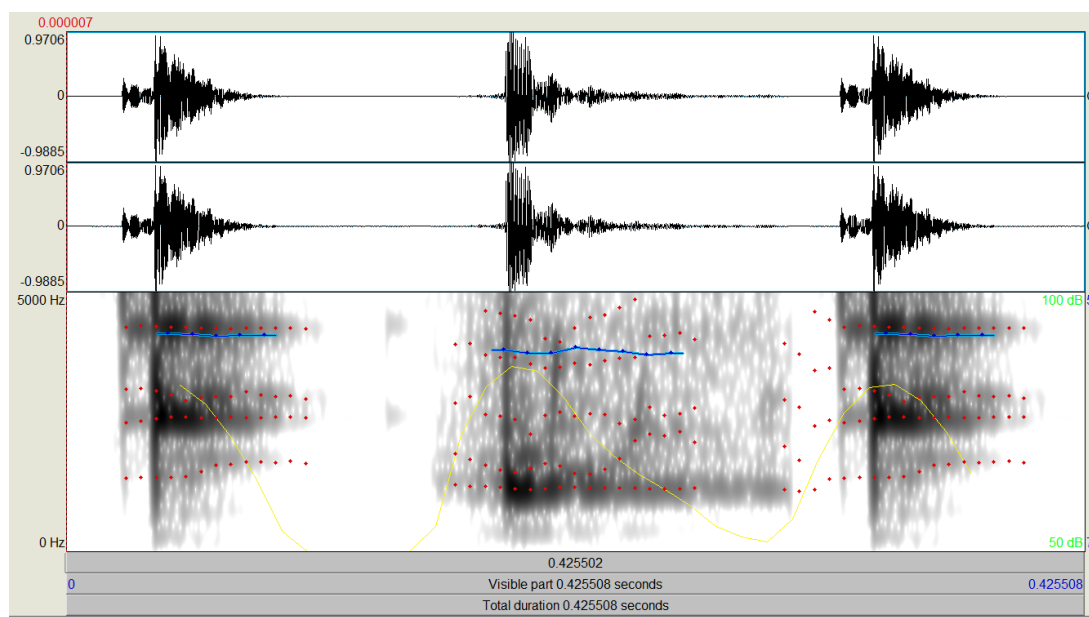


Figure 2.B. Auditory streaming sound wave form (top) and spectrogram (bottom). The yellow line represents the intensity contour, the red dots represent the formant contour (the concentrations of acoustic energy), and the blue line represents the pitch contour.

Appendix C

STUDY 1: Experiment 1

Considering no-reports as zero instead of missing values

In this appendix are found the exact same analyses performed in the main text, with the exception that in the blocks where participants did not switch, instead of missing values these were assigned zero values.

Please note that the headings and table numbers match those in the main thesis for an easier read (with the exception of the letter C that stands for Appendix C)

2.3.3.1. 1st vs. subsequent phase durations.

It was explored whether the first phase duration and the subsequent durations were significantly different (Figure 2.6.C). A Linear mixed effects model (LMEM) on the phase durations showed that there is a significant difference between 1st phase durations and the median subsequent phase durations. The first phase duration ($M = 19.79$, $SD = 17.55$) was significantly higher than the subsequent phase durations ($M = 5.20$, $SD = 6.70$), $F(1, 316) = 184.94$, $p < .001$.

2.3.3.2. Switching rates across blocks.

Before the main analyses were conducted, a LMEM was performed to assess whether switching rates differ across blocks (see descriptive statistics in Table 2.1.C). The results showed that there was an effect of Block on the switching rates, $F(4, 809) = 6.91$, $p < 0.001$. Planned comparisons using Dunnett test between block 1 and each of the subsequent blocks revealed that there was a significant difference between block 1 and 3, block 1 and block 4 and between block 1 and block 5 (all $ps < .001$).

Table 2.1.C

Means and Standard Errors for the Switching Rates in each Block

	Block1	Block2	Block3	Block4	Block5
Mean	.06	.08	.09	.09	.09
SD	.08	.10	.10	.11	.11

2.3.6.2. First phase duration.

The first phase duration performance across the four perceptual tasks is summarised in

Table 2.4.C.

Table 2.4.C

The Overall Performance for Individual Differences Measures

	Tasks	MaxObs	FinalObs	Min	Max	Mean	SD	SEM
1 st Phase	AF	184	183	0.00	43.10	5.94	7.08	0.52
Duration	AM	180	178	0.00	59.10	11.53	13.08	1.13
	VT	172	171	0.00	59.87	15.29	17.22	1.37
	AS	192	191	0.02	41.61	8.84	7.11	0.51

Note. MaxObs: the maximum number of observations per task; FinalObs: the final number of observations after omitting missing values.

Participants' first phase duration was longer in the auditory tasks than the visual tasks, as indicated by the effect of Modality, $F(1, 623) = 18.13, p < .001$. There was no effect of Level, $F(1, 717) = 0.39, p = .530$. However, there was an interaction, $F(1, 712) = 15.50, p < .001$. Firstly, there was a significant difference between the two levels within each modality (i.e., high-level vision vs. low-level vision and high-level auditory vs. low-level auditory). In the auditory modality, the first phase duration was longer for verbal transformations (high-level task) than for auditory streaming (low-level task), $t(719) = 2.394, p = .017$. In the visual modality, the first phase duration was longer in ambiguous motion (low-level task) than in ambiguous figure (high-level task), $t(719) = -3.284, p = .001$. Secondly, the impact of Level was investigated between modalities (high-level vision vs. high-level auditory and low-level vision vs. low-level auditory). Results showed that the first phase duration in the ambiguous figure is shorter than in the verbal transformations, $t(719) = -5.674, p$

$< .001$, while the first phase duration in the ambiguous motion was not significantly different than for the auditory streaming, $t(719) = 0.590$, $p = .555$.

2.3.6.3. Switching rates.

The switching rates across the four perceptual tasks is summarised in Table 2.5.C.

Table 2.5.C

Switching Rates Performance across the Four Perceptual Tasks

	Tasks	MaxObs	FinalObs	Min	Max	Mean	SD	SEM
Switching Rate	AF	184	184	0.00	0.72	0.13	0.14	.01
	AM	180	180	0.00	0.47	0.06	0.07	.01
	VT	172	172	0.00	0.55	0.13	0.11	.01
	AS	192	192	0.00	0.11	0.02	0.03	>.01

Note. MaxObs: the maximum number of observations per task; FinalObs: the final number of observations after omitting missing values.

Participants had a higher switching rate in the auditory tasks than in the visual tasks as indicated by the effect of Modality, $F(1,634) = 31.13$, $p < .001$. There was a significant effect of Level, $F(1,723) = 223.80$, $p < .001$. The switching rate was higher in the high-level tasks than in the low-level tasks. However, there was an interaction between Modality and Level, $F(1, 724) = 19.37$, $p < .001$. Pairwise tests indicated a difference between the two levels within each modality (i.e., high vision vs. low vision and high auditory vs. low auditory). In the auditory modality, the switching rate was higher in verbal transformations (high-level task) than in the auditory streaming (low-level task), $t(727) = 13.67$, $p < .001$. In the visual modality, the switching rate was higher in ambiguous figure (high-level task) than in ambiguous motion (low-level task), $t(727) = 7.33$, $p < .001$. The results showed a Level effect across modalities (high vision vs. high auditory and low vision vs. low auditory). The switching rates between the visual and auditory high-level tasks were not significantly different $t(714) = -0.22$, $p = 0.823$. On the other hand, the switching rate in the ambiguous motion (low-level visual task) was significantly higher than in auditory streaming (the low-level auditory task), $t(713) = 7.13$, $p < .001$.

2.3.8. Relationships between Perceptual Switching Variables across Perceptual Tasks

Table 2.7. C

Commonalities between Perceptual Switching across the Four Perceptual Tasks

Group	AF & VT	AM & AS	AM & VT	AF & AS
Initial RT				
S rho	.439*	.051	.426*	.391
N	21	23	22	25
1stPhase Durations				
S rho	.631**	-.120	.337	.366
N	21	23	22	25
Switching rate				
S rho	.590**	.183	.491*	.119
N	21	23	22	25

Note. S rho = Spearman test rho.

2.3.9. Relationships between Perceptual Switching Variables within the Same Perceptual Task

Table 2.8.C

Correlations for Perceptual Switching Variables (Initial Reaction Time, First Phase Duration, and Switching Rate) within Each Perceptual Task

		AF		AM		VT		AS	
		1PhD	SR	1PhD	SR	1PhD	SR	1PhD	SR
Initial	S rho	.150	-.263	.059	-.192	.090	-.255	.334*	-.166
RT	N	46	46	45	45	43	43	48	48
1stPhD	S rho	--	-.872***	--	-.576***	--	-.677***	--	-.732***
	N	--	46	--	45	--	43	--	48

Note. Initial RT = Initial reaction time, 1PhD = first phase duration, SR = Switching rates.

Table 2.9.C

Perceptual Switching and Individual Differences in Ambiguous Figure (N = 46)

	Stroop			FIST			Creativity			Verbal Fluency	
	Stroop RTs	Memory Control RTs	Inhibition RTs	Accuracy	RT	Errors	cTotal	cUnique	cUnusual	VFswitching	VFtotal
Initial Reaction time	.413**	.313*	.015	-.236	.236	.237	-.075	.013	-.071	-.168	-.131
1 st Phase Duration	.140	.038	.054	-.088	-.081	.011	.031	-.202	.037	-.058	-.217
Switching Rate	-.192	-.157	-.052	-.009	-.000	.068	.035	.239	-.046	.047	.090

Table 2.10.C

Perceptual Switching and Individual Differences in Ambiguous Structure-from-motion (N = 45)

	Stroop			FIST			Creativity			Verbal Fluency	
	Stroop RTs	Memory Control RTs	Inhibition RTs	Accuracy	RT	Errors	cTotal	cUnique	cUnusual	VFswitching	VFtotal
Initial Reaction time	.271	.115	.364*	-.056	.244	.048	-.177	-.188	-.107	-.013	-.142
1 st Phase Duration	-.140	-.192	.034	-.353*	.013	.417**	-.162	-.034	-.134	-.036	-.044
Switching Rate	-.036	.037	-.106	.197	.018	-.203	-.179	-.180	-.302*	-.178	-.116

Table 2.11.C

Perceptual Switching and Individual Differences in Verbal Transformations (N = 43)

	Stroop			FIST			Creativity			Verbal Fluency	
	Stroop RTs	Memory Control RTs	Inhibition RTs	Accuracy	RT	Errors	cTotal	cUnique	cUnusual	VFswitching	VFtotal
Initial Reaction time	.371*	.338*	.214	.113	.146	-.068	-.019	.117	-.158	-.039	.095
1 st Phase Duration	.009	-.200	-.076	.015	-.090	-.039	.424***	.314*	.359*	.047	.045
Switching Rate	-.201	-.005	-.273	.131	.169	-.185	-.306*	-.054	-.291	.042	-.051

Table 2.12.C

Perceptual Switching and Individual Differences in Auditory Streaming (N = 48)

	Stroop			FIST			Creativity			Verbal Fluency	
	Stroop RTs	Memory Control RTs	Inhibition RTs	Accuracy	RT	Errors	cTotal	cUnique	cUnusual	VFswitching	VFtotal
Initial Reaction time	.178	.260	.030	.111	.180	-.160	-.130	-.101	-.047	.066	-.117
1 st Phase Duration	.172	.002	.260	.228	.107	.005	-.216	-.124	-.149	.094	.064
Switching Rate	-.242	-.100	-.327*	-.209	-.188	.010	.264	.169	.191	-.078	.051

Appendix D
Spectrogram for the sounds used in the Verbal transformations task in Study 1
(Experiment 2) and Study 2

Figure 1.D depicts the wave form and the spectrogram for the word “Life” used in the verbal transformations task in Study 1(Experiment 2) and Study 2.

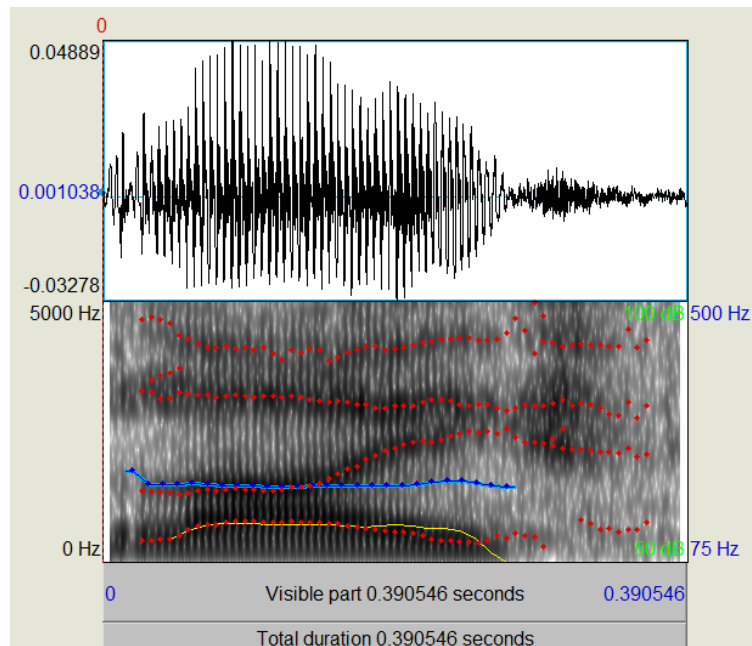


Figure 1.D. “Life” word wave form (top) and spectrogram (bottom). The yellow line represents the intensity contour, the red dots represent the formant contour (the concentrations of acoustic energy), and the blue line represents the pitch contour.

Appendix E

Ego Resiliency Questionnaire

ER89

Below are a number of statements that describe ways in which people act and think. For each statement, please indicate how much these items apply to you. If it **Does not apply at all** circle 1, if it **Applies slightly**, if at all circle 2, if it **Applies somewhat** circle 3, and if it **Applies very strongly** circle 4. Be sure to indicate your answer for every statement below.

1	2	3	4	
Does not apply at all	Applies slightly, if at all	Applies somewhat	Applies very strongly	
1. I am generous with my friends	1	2	3	4
2. I quickly get over and recover from being startled.	1	2	3	4
3. I enjoy dealing with new and unusual situations.	1	2	3	4
4. I usually succeed in making a favorable impression on people.	1	2	3	4
5. I enjoy trying new foods I have never tasted before.	1	2	3	4
6. I am regarded as a very energetic person.	1	2	3	4
7. I like to take different paths to familiar places.	1	2	3	4
8. I am more curious than most people.	1	2	3	4
9. Most of the people I meet are likeable.	1	2	3	4
10. I usually think carefully about something before acting.	1	2	3	4
11. I like to do new and different things.	1	2	3	4
12. My daily life is full of things that keep me interested.	1	2	3	4
13. I would be willing to describe myself as a pretty “strong” personality.	1	2	3	4
14. I get over my anger at someone reasonably quickly.	1	2	3	4

Appendix F

STUDY 1: Experiment 2

Considering no-reports as zero instead of missing values

2.7.5. Effects of Modality and Level of Processing on Perceptual Switching

2.7.5.1. Initial reaction time.

The overall performance for the initial reaction times across the four perceptual tasks is summarised in Table 2.15.F.

Table 2.15.F

Initial Reaction Times Performance across the Four Perceptual Tasks (N = 31)

	Tasks	No. Obs	Min	Max	Mean	SD	SEM
Initial RT	AF	124	0.29	13.74	2.69	1.84	.17
	AM	124	0.16	8.92	2.31	1.47	.13
	VT	124	0.67	7.59	2.13	0.98	.09
	AS	124	0.83	34.33	4.67	5.53	.05

Note. Maximum number of observations was 124.

The initial reaction times were longer for auditory than visual tasks as indicated by the effect of Modality, $F(1, 34) = 6.92, p = 0.02$. There was an effect of Level, $F(1, 41) = 12.20, p = .001$, the initial reaction time for the high-level tasks was shorter than for the low-level tasks. The main effects were qualified with an interaction between Modality and Level, $F(1, 44) = 23.21, p < .001$. Pairwise contrasts showed that the difference between the initial reaction time in ambiguous figure (high-level visual) and ambiguous motion (low-level visual) did not differ $t(30) = 1.69, p = 0.122$. In the auditory modality, participants responded faster to verbal transformations (high-level) than to auditory streaming (low-level), $t(30) = -5.166, p < .001$. Further, results showed that the initial reaction time in ambiguous figure (high-level visual) was not significantly different from verbal transformations (high-level auditory), $t(30) = 1.715, p = 0.122$, while the initial reaction time in the ambiguous motion (visual low-level) was significantly shorter than in auditory streaming (auditory low-level), $t(30) = -4.139, p < 0.001$.

2.7.5.2. First phase duration.

The overall performance for first phase duration across the four perceptual tasks is summarised in Table 2.16.F

Table 2.16.F

First Phase Duration Performance across the Four Perceptual Tasks (N = 31)

	Tasks	No. Obs	Min	Max	Mean	SD	SEM
1 st Phase Duration	AF	124	1.09	118.9	26.61	32.53	2.93
	AM	124	0.04	118.7	30.07	22.67	3.58
	VT	124	0.02	118.2	10.86	14.37	1.30
	AS	124	0.41	118.8	36.64	42.00	3.79

Note. Maximum number of observations was 124.

There was an effect of Modality, $F(1, 40) = 4.50, p = .04$, the first phase duration was longer in the visual than in the auditory tasks. There was a main effect of Level, $F(1, 33) = 10.37, p < .001$, the first phase duration in the high-level tasks were larger than in the low-level tasks. However, there was an interaction, $F(1, 33) = 10.04, p = .003$. Pairwise comparisons showed that only in the auditory modality there was a significant difference between the first phase duration in verbal transformations (high-level) and auditory streaming (low-level), $t(30) = -6.264, p < .001$. In the visual tasks, the first phase durations in ambiguous figure (high-level) and ambiguous motion (low-level) did not differ significantly, $t(30) = -0.891, p = 0.3802$. The results showed further that there was an impact of Level across modalities (high vision vs. high auditory and low vision vs. low auditory). In ambiguous figure (high-level visual task) there was a significantly larger first phase duration than in verbal transformations (auditory high-level), $t(30) = 4.51, p < 0.001$, while the first phase duration in auditory streaming and ambiguous motion (AM) did not differ significantly, $t(30) = -0.789, p = 0.436$.

2.7.5.3. Switching rates.

The overall performance for the switching rates across the four perceptual tasks is summarized in Table 2.17.F

Table 2.17.F

Switching Rate Performance across the Four Perceptual Task (N = 31)

	Tasks	No. Obs	Min	Max	Mean	SD	SEM
Switching Rate	AF	124	0.00	0.36	0.06	0.08	.007
	AM	124	0.00	0.24	0.04	0.05	.004
	VT	124	0.00	0.84	0.14	0.14	.013
	AS	124	0.00	0.26	0.04	0.05	.004

As indicated by a significant main effect of Level, $F(1, 42) = 20.20, p < .001$, the switching rates were higher in the high-level tasks than for the low-level tasks. There was also an effect of Modality, the switching rates in the ambiguous figure was higher than in the ambiguous motion, $F(1, 35) = 8.25, p = .007$. These main effects were qualified by an interaction between Modality and Level, $F(1, 38) = 9.92, p = .003$. The interaction was further disentangled with pairwise comparisons, adjusted for multiple testing. For the auditory task, the results showed that the switching rate in verbal transformations was higher than in auditory streaming, $t(30) = 4.676, p < .001$. There was no difference between the switching rate in ambiguous motion and ambiguous figure $t(30) = 1.73, p = .141$. Secondly, the impact of Level across modalities (high vision vs. high auditory and low vision vs. low auditory) was also investigated. Results showed that the switching rate in ambiguous figure was significantly lower than in verbal transformations, $t(30) = -3.65, p = 0.002$, while the switching rate in ambiguous motion and auditory streaming was not significantly different, $t(30) = -0.191, p = 0.895$.

2.7.8. Perceptual Switching and Individual Differences Measures

Table 2.18. F

Correlations between the 1st Phase Durations across the Four Perceptual Tasks and the Individual Differences Measures

Task	AF	AM	VT	AS	Inhibition	Mem Control	Stroop Effect	Ego-Resiliency
AF	--	.231	.292	.314	.037	.090	0.123	-.231
AM			.271	.232	.123	.029	.244	.041
VT				.114	.437*	.287	.430*	-.019
AS					.072	.088	.149	.166

Note. S rho = Spearman correlation coefficient.

Table 2.19.F

Correlations between the Switching Rates across the Four Perceptual Tasks and the Individual Differences Measures (N = 31)

Task	AF	AM	VT	AS	Inhibition	Mem Control	Stroop Effect	Ego-Resiliency
AF	--	.326*	.348	.409*	-.255	-.199	-.403*	.129
AM			.121	.262	-.128	-.122	-.152	.115
VT				.236	-.499**	-.447*	-.279	-.030
AS					.015	.028	-.069	-.104

Appendix G

Edinburgh Handedness Questionnaire (Study 4)

112

R. C. OLDFIELD

APPENDIX II

Medical Research Council Speech & Communication Unit

EDINBURGH HANDEDNESS INVENTORY

Surname..... Given Names.....

Date of Birth..... Sex.....

Please indicate your preferences in the use of hands in the following activities *by putting + in the appropriate column*. Where the preference is so strong that you would never try to use the other hand unless absolutely forced to, *put ++*. If in any case you are really indifferent *put + in both columns*.

Some of the activities require both hands. In these cases the part of the task, or object, for which hand preference is wanted is indicated in brackets.

Please try to answer all the questions, and only leave a blank if you have no experience at all of the object or task.

		LEFT	RIGHT
1	Writing		
2	Drawing		
3	Throwing		
4	Scissors		
5	Toothbrush		
6	Knife (without fork)		
7	Spoon		
8	Broom (upper hand)		
9	Striking Match (match)		
10	Opening box (lid)		
i	Which foot do you prefer to kick with?		
ii	Which eye do you use when using only one?		

L.Q.

Leave these spaces blank

DECILE

MARCH 1970

Appendix H

Study 4

Examining effects of fatigue on perceptual switching.

Effects of fatigue and tiredness can influence the cognitive performance in experiments (Langner & Eickhoff, 2013). It is hard to control for these kinds of effects in perceptual switching experiments as the experimental blocks need to be long enough for the switching rates to stabilize (Suzuki & Grabowecky, 2007). The steps taken in Study 4 to prevent fatigue are described in section 5.2.7 at pages 157-158.

One indirect way to analyse if the results in Study 4 were affected by fatigue and tiredness is by looking at the difference between the number of switches in the first and last blocks of each condition. Presumably, if fatigue influenced the number of switches, we would find a significant difference across blocks.

The number of switches in each of the four blocks, for each of the three conditions, are separately displayed for two-category response group (Figure 1.H) and three-category response group (Figure 2.H). T-tests were conducted to see if the number of switches in Block 1 was significantly different from Block 4. Except in the auditory neutral condition in the two-response group (Table 1.H and Table 2.H), the number of switches in Block 1 was not statistically different from Block 4. These results do not support the hypothesis that the findings presented in Study 4 were considerably influenced by fatigue.

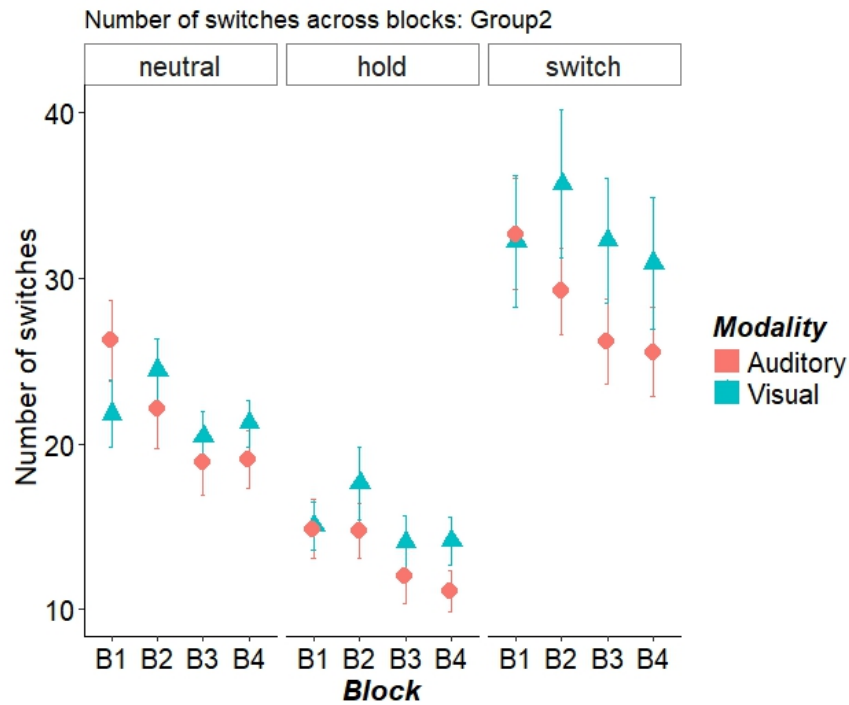


Figure 1.H. The number of switches reported by participants in two-alternative group, for each of the four blocks (B1-B4), separately for each condition and modality.

Table 1.H

T-tests between the number of switches in Block 1 and Block 4 for participants in two-response group

Task	Condition	Block 1 Average (SD)	Block 4 Average (SD)	t-test	p-value	SE	Cohen d
Auditory	Neutral	26.24 (13.97)	19.09 (9.95)	3.872	<.001	1.847	.674
Auditory	Hold	14.88 (10.46)	11.12 (7.21)	2.924	.006	1.285	.509
Auditory	Switch	32.67 (19.39)	25.55 (15.38)	2.335	.002	2.135	.581
Visual	Neutral	21.82 (11.76)	21.24 (8.05)	.365	.718	1.578	.064
Visual	Hold	15.06 (8.42)	14.12 (8.49)	.656	.517	1.433	.114
Visual	Switch	32.21 (22.74)	30.91 (22.79)	.477	.636	2.731	.083

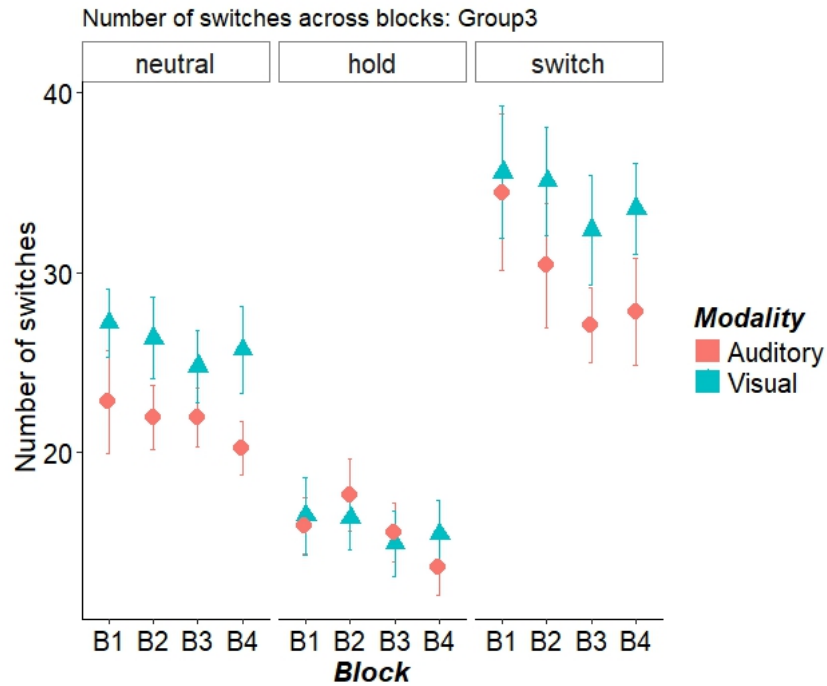


Figure 2.H. The number of switches reported by participants in three-alternative group, for each of the four blocks (B1-B4), separately for each condition and modality.

Table 2.H

T-tests between the number of switches in Block 1 and Block 4 for participants in three-response group

Task	Condition	Block 1 Average (SD)	Block 4 Average (SD)	t-test	p-value	SE	Cohen d
Auditory	Neutral	22.79 (15.39)	20.24 (7.95)	.881	.386	2.898	.164
Auditory	Hold	15.93 (8.46)	13.62 (8.60)	1.931	.064	1.196	.359
Auditory	Switch	34.48 (23.60)	29.07 (15.16)	2.021	.053	2.678	.375
Visual	Neutral	27.17 (10.13)	25.69 (12.96)	.769	.448	1.927	.143
Visual	Hold	16.45 (11.61)	15.41 (10.09)	.838	.409	1.234	.156
Visual	Switch	35.59 (19.74)	33.55 (13.48)	.680	.502	2.034	.126

Appendix J

Creativity scale (Study 4)

Please answer as truthfully as you can. Place a cross (X) in the box next to the activities you have been actively involved in.

In the past 12 months have you...

1	Written a short story	
2	Written a novel	
3	Organized an event, show, performance or activity	
4	Produced a TV/Play script	
5	Designed and produced a textile product (e.g. made an item of clothing or household object)	
6	Redesigned and redecorated a bedroom, kitchen, personal space, etc.,	
7	Invented and made a product that can be used	
8	Drawn a cartoon	
9	Started a club, association or group	
10	Produced a picture, i.e. NOT a doodle (using paint, pencils, charcoal, acrylic, etc.,)	
11	Had an article published	
12	Formed a sculpture using any suitable materials	
13	Recognised where an accepted scientific theory/approach does not explain what it purports to	
14	Produced your own food recipes	
15	Produced a short film	
16	Produced your own website	
17	Produced a theory to explain a phenomenon	
18	Invented a game or other form of entertainment	
19	Selected to lead/manage others	
20	Made someone a present	
21	Composed a poem	
22	Adapted an item and used it in a way that it was not designed to be, in what you consider to be an ingenious way	
23	Published research	
24	Choreographed a dance	
25	Designed and planted a garden	
26	Produced a portfolio of photographs (NOT photographs of a holiday, party, etc.,)	
27	Acted in a dramatic production	
28	Delivered a speech	
29	Mentored/Coached someone else to improve their performance	
30	Devised an experiment to help understand something	
31	Made up a joke	
32	Been made a leader/captain of a team/group (e.g. Debating society chairperson, Captain of the Hockey team, etc.,)	
33	Composed a piece of music	
34	Made a collage	

REFERENCES

- Alais, D., & Blake, R. (2015). *The Oxford handbook of perceptual organization. In Chapter 38: Binocular rivalry and perceptual ambiguity.*
- Alais, D., Newell, F. N., & Mamassian, P. (2010). Multisensory Processing in Review: from Physiology to Behaviour. *Seeing and Perceiving*, 23(1), 3-38.
doi:10.1163/187847510x488603
- Alais, D., van Boxtel, J. J., Parker, A., & van Ee, R. (2010). Attending to auditory signals slows visual alternations in binocular rivalry. *Vision Research*, 50(10), 929-935.
- Andrews, T. J., & Purves, D. (1997). Similarities in normal and binocularly rivalrous viewing. *Proceedings of the National Academy of Sciences of the United States of America*, 94(18), 9905-9908.
- Anstis, S., & Saida, S. (1985). Adaptation to Auditory Streaming of Frequency-Modulated Tones. *Journal of Experimental Psychology-Human Perception and Performance*, 11(3), 257-271. doi:10.1037/0096-1523.11.3.257
- Barniv, D., & Nelken, I. (2015). Auditory Streaming as an Online Classification Process with Evidence Accumulation. *Plos One*, 10(12). doi:10.1371/journal.pone.0144788
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3), 255-278.
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., & Singmann, H. (2015). R package version, 1(4).
- Benjamini, Y. (2010). Simultaneous and selective inference: Current successes and future challenges. *Biometrical Journal*, 52(6), 708-721. doi:10.1002/bimj.200900299
- Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science*, 7(325-339).
- Bialystok, E., & Shapero, D. (2005). Ambiguous benefits: The effect of bilingualism on reversing ambiguous figures. *Developmental Science*, 8(6), 595-604.
- Billig, A. J., Davis, M. H., & Carlyon, R. P. (2018). Neural decoding of bistable sounds reveals an effect of intention on perceptual organization. *Journal of Neuroscience*, 3022-3017. doi:10.1523/JNEUROSCI.3022-17.2018
- Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, 3(1), 13-23. doi:DOI 10.1038/nrn701
- Block, J., & Kremen, A. M. (1996). IQ and ego-resiliency: Conceptual and empirical connections and separateness. *Journal of Personality and Social Psychology*, 70(2), 349-361. doi:10.1037/0022-3514.70.2.349
- Bonneh, Y. S., Pavlovskaya, M., Ring, H., & Soroker, N. (2004). Abnormal binocular rivalry in unilateral neglect: evidence for a non-spatial mechanism of extinction. *Neuroreport*, 15(3), 473-477. doi:10.1097/01.wnr.0000109988.85243.c6
- Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Bargary, G., Hogg, R. E., & Mollon, J. D. (2015). A population study of binocular function. *Vision Research*, 110(Pt A), 34-50. doi:10.1016/j.visres.2015.02.017
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological review*, 108(3), 624-652.
- Brascamp, J., Blake, R., & Knapen, T. (2015). Negligible fronto-parietal BOLD activity accompanying unreportable switches in bistable perception. *Nature Neuroscience*, 18(11), 1672-1678. doi:10.1038/nn.4130
- Brascamp, J., Sterzer, P., Blake, R., & Knapen, T. (2018). Multistable perception and the role of frontoparietal cortex in perceptual inference. *Annual Review of Psychology*, 69(1). doi:10.1146/annurev-psych-010417-085944

- Brascamp, J., Van Ee, R., Noest, A. J., Jacobs, R. H., & van den Berg, A. V. (2006). The time course of binocular rivalry reveals a fundamental role of noise. *Journal of Vision*, 6(11), 1244-1256. doi:10.1167/6.11.8
- Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences*, 16(2), 106-113. doi:10.1016/j.tics.2011.12.010
- Bremner, A. J., Doherty, M. J., Caparos, S., Fockert, J., Linnell, K. J., & Davidoff, J. (2016). Effects of culture and the urban environment on the development of the Ebbinghaus illusion. *Child Development*, 87(3), 962-981. doi:10.1111/cdev.12511
- Britz, J., Landis, T., & Michel, C. M. (2008). Right parietal brain activity precedes perceptual alternation of bistable stimuli. *Cerebral Cortex*, 19(1), 55-65. doi:10.1002/hbm.21117
- Cao, R., Pastukhov, A., Mattia, M., & Braun, J. (2016). Collective Activity of Many Bistable Assemblies Reproduces Characteristic Dynamics of Multistable Perception. *Journal of Neuroscience*, 36(26), 6957-6972. doi:10.1523/JNEUROSCI.4626-15.2016
- Carlson, S. M., & Meltzoff, A. M. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, 11(282-298).
- Carmel, D., Walsh, V., Lavie, N., & Rees, G. (2010). Right parietal TMS shortens dominance durations in binocular rivalry. *Current Biology*, 20(18), R799-R800. doi:10.1016/j.cub.2010.07.036
- Carter, O., Konkle, T., Wang, Q., Hayward, V., & Moore, C. (2008). Tactile rivalry demonstrated with an ambiguous apparent-motion quartet. *Current Biology*, 18(14), 1050-1054.
- Carter, O., & Pettigrew, J. D. (2003). A common oscillator for perceptual rivalries? *Perception*, 32(3), 295-305. doi:10.1068/p3472
- Chamberlain, R., Swinnen, L., Heeren, S., & Wagemans, J. (2017). Perceptual flexibility is coupled with reduced executive inhibition in students of the visual arts. *British Journal of Psychology*. doi:10.1111/bjop.12253
- Chen, X. C., & He, S. (2004). Local factors determine the stabilization of monocular ambiguous and binocular rivalry stimuli. *Current Biology*, 14(11), 1013-1017. doi:10.1016/j.cub.2004.05.042
- Cosmelli, D., David, O., Lachaux, J. P., Martinerie, J., Garnero, L., Renault, B., & Varela, F. (2004). Waves of consciousness: ongoing cortical patterns during binocular rivalry. *Neuroimage*, 23(1), 128-140.
- Crick, F., & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex*, 8(2), 97-107. doi:DOI 10.1093/cercor/8.2.97
- Cusack, R. (2005). The intraparietal sulcus and perceptual organization. *Journal of Cognitive Neuroscience*, 17(4), 641-651. doi:Doi 10.1162/0898929053467541
- de Graaf, T. A., de Jong, M. C., Goebel, R., van Ee, R., & Sack, A. T. (2011). On the Functional Relevance of Frontal Cortex for Passive and Voluntarily Controlled Bistable Vision. *Cerebral Cortex*, 21(10), 2322-2331. doi:10.1093/cercor/bhr015
- Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, 70(2), 200-227.
- Dekker, T. M., Schwarzkopf, D. S., de Haas, B., Nardini, M., & Sereno, M. I. (2017). Population receptive field tuning properties of visual cortex during childhood. *bioRxiv*, 213108. doi:10.1101/213108
- Denham, S. L., Bendixen, A., Mill, R., Toth, D., Wennekers, T., Coath, M., . . . Winkler, I. (2012). Characterising switching behaviour in perceptual multi-stability. *Journal of Neuroscience Methods*, 210(1), 79-92. doi:10.1016/j.jneumeth.2012.04.004
- Denham, S. L., Bohm, T. M., Bendixen, A., Szalardy, O., Kocsis, Z., Mill, R., & Winkler, I. (2014). Stable individual characteristics in the perception of multiple embedded

- patterns in multistable auditory stimuli. *Frontiers in Neuroscience*, 8. doi:10.3389/fnins.2014.00025
- Denham, S. L., Farkas, D., van Ee, R., Taranu, M., Kocsis, Z., Wimmer, M., . . . Winkler, I. (2018). Similar but separate systems underlie perceptual bistability in vision and audition. *Scientific reports*, 8. doi:10.1038/s41598-018-25587-2
- Denham, S. L., Gyimesi, K., Stefanics, G., & Winkler, I. (2013). Perceptual bistability in auditory streaming: how much do stimulus features matter? *Learning & Perception*, 5(2), 73-100. doi:DOI: 10.1556/LP.5.2013.Suppl2.6
- Denham, S. L., & Winkler, I. (2006). The role of predictive models in the formation of auditory streams. *Journal of Physiology-Paris*, 100(1-3), 154-170. doi:10.1016/j.jphysparis.2006.09.012
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64(1), 135-168. doi:10.1146/annurev-psych-113011-143750
- Díaz-Santos, M., Mauro, S., Cao, B., Yazdanbakhsh, A., Neargarder, S., & Cronin-Golomb, A. (2017). Bistable perception in normal aging: perceptual reversibility and its relation to cognition. *Aging, Neuropsychology, and Cognition*, 24(2), 115-134. doi:10.1080/13825585.2016.1173646
- Dick, A. S. (2014). The development of cognitive flexibility beyond the preschool period: An investigation using a modified Flexible Item Selection Task. *Journal of Experimental Child Psychology*, 125, 13-34. doi:10.1016/j.jecp.2014.01.021
- Dieter, K. C., & Tadin, D. (2011). Understanding Attentional Modulation of Binocular Rivalry: A Framework Based on Biased Competition. *Frontiers in Human Neuroscience*, 5(155). doi:10.3389/fnhum.2011.00155
- Doherty, M. J., Campbell, N. M., Tsuji, H., & Phillips, W. A. (2010). The Ebbinghaus illusion deceives adults but not young children. *Developmental Science*, 13(5), 714-721. doi:10.1111/j.1467-7687.2009.00931.x
- Doherty, M. J., & Mair, S. (2012). Creativity, ambiguous figures, and academic preference. *Perception*, 41(10), 1262-1266. doi:10.1068/p7350
- Doherty, M. J., & Wimmer, M. C. (2005). Children's understanding of ambiguous figures: Which cognitive developments are necessary to experience reversal? *Cognitive Development*, 20(3), 407-421. doi:10.1016/j.cogdev.2005.05.003
- Donnell, A. J., Maki, P. M., Pliskin, N. H., & Kraus, M. F. (2004). Clustering and switching on verbal fluency in traumatic brain injury. *Archives of Clinical Neuropsychology*, 19(7), 960-960.
- Ehlers, J., Struber, D., & Basar-Eroglu, C. (2015). Multistable perception in children: Prefrontal delta oscillations in the developing brain. *International Journal of Psychophysiology*, 103, 129-134.
- Farkas, D., Denham, S. L., Bendixen, A., Tóth, D., Kondo, H. M., & Winkler, I. (2016). Auditory multi-stability: idiosyncratic perceptual switching patterns, executive functions and personality traits. *Plos One*, 11(5), e0154810. doi:10.1371/journal.pone.0154810
- Frässle, S., Sommer, J., Jansen, A., Naber, M., & Einhäuser, W. (2014). Binocular rivalry: frontal activity relates to introspection and action but not to perception. *Journal of Neuroscience*, 34, 1738-1747. doi:10.1523/JNEUROSCI.4403-13.2014
- Friston, K. J., & Stephan, K. E. (2007). Free-energy and the brain. *Synthese*, 159(3), 417-458. doi:10.1007/s11229-007-9237-y
- Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, 10(4), 483-527. doi:10.1016/0885-2014(95)90024-1
- Fuster, J. M. (2002). Frontal lobe and cognitive development. *J Neurocytol*, 31(3-5), 373-385.

- Gallagher, R. M., & Arnold, D. H. (2014). Interpreting the temporal dynamics of perceptual rivalries. *Perception*, 43(11), 1239-1248. doi:10.1068/p7648
- Gelman, A. (2005). Analysis of variance—why it is more important than ever. *The annals of statistics*, 33(1), 1-53.
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel hierarchical models*. New York, NY, USA: Cambridge University Press.
- Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The Relationship between Cognition and Action - Performance of Children 31/2-7 Years Old on a Stroop-Like Day-Night Test. *Cognition*, 53(2), 129-153. doi:10.1016/0010-0277(94)90068-X
- Gonzalez, C. L. R., & Flindall, J. W. (2015). *International Encyclopedia of the Social & Behavioral Sciences*
- Gopnik, A., & Rosati, A. (2001). Duck or rabbit? Reversing ambiguous figures and understanding ambiguous representations. *Developmental Science*, 4(2), 175-183. doi:10.1111/1467-7687.00163
- Gopnik, A., & Rosati, A. (2011). Duck or rabbit? Reversing ambiguous figures and understanding ambiguous representations. *Developmental Science*, 4(2), 175-183.
- Goryo, K. (1960). The effect of past experience upon binocular rivalry. *Japanese Psychological Research*, 11(46-53).
- Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4), 404-411. doi:10.1037/h0059831
- Gray, K. L., Adams, W. J., & Garner, M. (2009). The influence of anxiety on the initial selection of emotional faces presented in binocular rivalry. *Cognition*, 113(1), 105-110.
- Guilford, J. P., Christensen, P. R., Merrifield, P. R., & Wilson, R. C. (1978). *Alternate Uses: Manual of instructions and interpretations*. Orange, CA: Sheridan Psychological Services.
- Hainline, L. (1998). The development of basic visual abilities. In A. Slater (Ed.), *Perceptual development: Visual auditory, and speech perception in infancy*. Hove, East Sussex: Psychology Press Ltd.
- Heenan, A., & Troje, N. F. (2015). The relationship between social anxiety and the perception of depth-ambiguous biological motion stimuli is mediated by inhibitory ability. *Acta Psychologica*, 157, 93-100.
- Hill, K. T., Bishop, C. W., Yadav, D., & Miller, L. M. (2011). Pattern of BOLD signal in auditory cortex relates acoustic response to perceptual streaming. *BMC neuroscience*, 12(1), 85.
- Hock, H. S., Schöner, G., & Giese, M. (2003). The dynamical foundations of motion pattern formation: Stability, selective adaptation, and perceptual continuity. *Attention, Perception, & Psychophysics*, 65(3), 429-457.
- Hohwy, J., Roepstorff, A., & Friston, K. (2008). Predictive coding explains binocular rivalry: An epistemological review. *Cognition*, 108(3), 687-701. doi:10.1016/j.cognition.2008.05.010
- Hol, K., Koene, A., & van Ee, R. (2003). Attention-biased multi-stable surface perception in three-dimensional structure-from-motion. *Journal of Vision*, 3(7), 486-498. doi:10.1167/3.7.3
- Huguet, G., Rinzel, J., & Hupé, J. M. (2014). Noise and adaptation in multistable perception: Noise drives when to switch, adaptation determines percept choice. *Journal of Vision*, 14(3), 19-19. doi:10.1167/14.3.19
- Hupé, J. M., Joffo, L. M., & Pressnitzer, D. (2008). Bistability for audiovisual stimuli: Perceptual decision is modality specific. *Journal of Vision*, 8,(7:1), 1-15.

- Hupé, J. M., & Pressnitzer, D. (2012). The initial phase of auditory and visual scene analysis. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 367(1591), 942-953. doi:10.1098/rstb.2011.0368
- Jacques, S., & Zelazo, P. D. (2001). The Flexible Item Selection Task (FIST): A measure of executive function in preschoolers. *Developmental Neuropsychology*, 20(3), 573-591. doi:10.1207/875656401753549807
- Jankowska, D. M., & Karwowski, M. (2015). Measuring creative imagery abilities. *Front Psychol*, 6(1591). doi:<http://doi.org/10.3389/fpsyg.2015.01591>
- Jastrow, J. (1900). *Fact and fable in psychology*. Boston: Houghton.
- Jastrow, J. (1971). *Fact and fable in psychology*. Freeport, N.Y.: Books for Libraries Press.
- Jung, R. E., Grazioplene, R., Caprihan, A., Chavez, R. S., & Haier, R. J. (2010). White matter integrity, creativity, and psy- chopathology: disentangling constructs with diffusion tensor imaging. *Plos One*, 5, e9818. doi:0.1371/journal.pone.0009818
- Kalarickal, G. J., & Marshall, J. A. (2000). Neural model of temporal and stochastic properties of binocular rivalry. *Neurocomputing*, 32, 843-853.
- Kanai, R., Bahrami, B., & Rees, G. (2010). Human Parietal Cortex Structure Predicts Individual Differences in Perceptual Rivalry. *Current Biology*, 20(18), 1626-1630. doi:10.1016/j.cub.2010.07.027
- Kanai, R., Carmel, D., Bahrami, B., & Rees, G. (2011). Structural and functional fractionation of right superior parietal cortex in bistable perception. *Current Biology*, 21(3), R106-R107. doi:10.1016/j.cub.2010.12.009
- Kashino, M., & Kondo, H. M. (2012). Functional brain networks underlying perceptual switching: auditory streaming and verbal transformations. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 367(1591), 977-987. doi:10.1098/rstb.2011.0370
- Kemp, B. J. (1973). Reaction time of young and elderly subjects in relation to perceptual deprivation and signal-on versus signal-off conditions. *Developmental Psychology*, 8(2), 268-272.
- Kleinschmidt, A., Sterzer, P., & Rees, G. (2012). Variability of perceptual multistability: from brain state to individual trait. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 367(1591), 988-1000. doi:10.1098/rstb.2011.0367
- Klink, P. C., van Ee, R., Nijs, M. M., Brouwer, G. J., Noest, A. J., & van Wezel, R. J. A. (2008). Early interactions between neuronal adaptation and voluntary control determine perceptual choices in bistable vision. *Journal of Vision*, 8(5). doi:10.1167/8.5.16
- Klink, P. C., van Ee, R., & van Wezel, R. J. A. (2008). General Validity of Levelt's Propositions Reveals Common Computational Mechanisms for Visual Rivalry. *Plos One*, 3(10). doi:10.1371/journal.pone.0003473
- Knapen, T., Brascamp, J., Pearson, J., van Ee, R., & Blake, R. (2011). The Role of Frontal and Parietal Brain Areas in Bistable Perception. *Journal of Neuroscience*, 31(28), 10293-10301. doi:10.1523/Jneurosci.1727-11.2011
- Kondo, H. M., & Kashino, M. (2007). Neural mechanisms of auditory awareness underlying verbal transformations. *Neuroimage*, 36(1), 123-130. doi:10.1016/j.neuroimage.2007.02.024
- Kondo, H. M., & Kashino, M. (2009). Involvement of the Thalamocortical Loop in the Spontaneous Switching of Percepts in Auditory Streaming. *Journal of Neuroscience*, 29(40), 12695-12701. doi:10.1523/Jneurosci.1549-09.2009
- Kondo, H. M., Kitagawa, N., Kitamura, M. S., Koizumi, A., Nomura, M., & Kashino, M. (2012). Separability and Commonality of Auditory and Visual Bistable Perception. *Cerebral Cortex*, 22(8), 1915-1922. doi:10.1093/cercor/bhr266

- Kondo, H. M., & Kochiyama, T. (2017). Normal aging slows spontaneous switching in auditory and visual bistability. *Neuroscience*. doi:10.1016/j.neuroscience.2017.04.040
- Kornmeier, J., Hein, C. M., & Bach, M. (2009). Multistable perception: When bottom-up and top-down coincide. *Brain and Cognition*, 69(1), 138-147. doi:10.1016/j.bandc.2008.06.005
- Kovács, I. (2000). Human development of perceptual organization. *Vision Research*, 40(10-12), 1301-1310. doi:10.1016/S0042-6989(00)00055-9
- Kowalski, C. J. (1972). On the effects of non-normality on the distribution of the sample product-moment correlation coefficient. *Applied Statistics*, 21(1), 1-12.
- Krug, K., Brunskill, E., Scarna, A., Goodwin, G. M., & Parker, A. J. (2008). Perceptual switch rates with ambiguous structure-from-motion figures in bipolar disorder. *Proceedings of the Royal Society of London B: Biological Sciences*, 275(1645), 1839-1848. doi:10.1098/rspb.2008.0043
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). R package version, 2(0).
- Langner, R., & Eickhoff, S. B. (2013). Sustaining Attention to Simple Tasks: A Meta-Analytic Review of the Neural Mechanisms of Vigilant Attention. *Psychological bulletin*, 139(4), 870-900. doi:10.1037/a0030694
- Lansbergen, M. M., Kenemans, J. L., & van Engeland, H. (2007). Stroop interference and attention-deficit/hyperactivity disorder: A review and meta-analysis. *Neuropsychology*, 21(2), 251-262. doi:10.1037/0894-4105.21.2.251
- Laukkonen, R. E., & Tangen, J. M. (2017). Can observing a Necker cube make you more insightful? *Consciousness and Cognition*, 48, 198-211. doi:doi:10.1016/j.concog.2016.11.011
- Lenth, R. V. (2017).
- Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: changing views in perception. *Trends in Cognitive Sciences*, 3(7), 254-264. doi:10.1016/S1364-6613(99)01332-7
- Levelt, W. J. (1967). Note on the distribution of dominance times in binocular rivalry. *Br J Psychol*, 58(1), 143-145.
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment*. New York, NY: Oxford University Press.
- Lindauer, M. S., & Reukauf, L. C. (1971). Introversion-extraversion and figure-ground perception. *Journal of Personality and Social Psychology*, 19(1), 107-113.
- Litovsky, R., 55. (2015). *Development of the auditory system. Handbook of clinical neurology* (Vol. 129).
- Losciuto, L. A., & Hartley, E. L. (1963). Religious affiliation and open-mindedness in binocular resolution. *Perceptual and Motor Skills*, 17(427-430).
- Lumer, E. D., Friston, K. J., & Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science*, 280(5371), 1930-1934. doi:10.1126/science.280.5371.1930
- Mamassian, P., & Goutcher, R. (2005). Temporal dynamics in bistable perception. *Journal of Vision*, 5(4), 7-7.
- Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: a model comparison perspective (2nd ed)*. Mahwah, N.J: Lawrence Erlbaum Associates.
- Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision*, 4(7), 2-2. doi:10.1167/4.7.2
- Mitroff, S. R., Sobel, D. M., & Gopnik, A. (2006). Reversing how to think about ambiguous figure reversals: Spontaneous alternating by uninformed observers. *Perception*, 35(5), 709-715. doi:10.1068/p5520

- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49-100. doi:10.1006/cogp.1999.0734
- Moreno-Bote, R., Shpiro, A., Rinzel, J., & Rubin, N. (2010). Alternation rate in perceptual bistability is maximal at and symmetric around equi-dominance. *Journal of Vision*, 10(11). doi:10.1167/10.11.1
- Nakatani, H., & van Leeuwen, C. (2005). Individual differences in perceptual switching rates; the role of occipital alpha and frontal theta band activity. *Biological Cybernetics*, 93(5), 343-354. doi:DOI 10.1007/s00422-005-0011-2
- Nayar, K., Franchak, J., Adolph, K., & Kiorpes, L. (2015). From local to global processing: The development of illusory contour perception. *Journal of Experimental Child Psychology*, 131, 38-55. doi:10.1016/j.jecp.2014.11.001
- Necker, L. A. (1832). Observations on some remarkable Optical Phenomena seen in Switzerland; and on an Optical Phenomenon which occurs on viewing a Figure of a Crystal or geometrical Solid. *The London and Edinburgh Philosophical Magazine and Journal of Science*, 1(5), 329-337.
- Ngo, T. T., Mitchell, P. B., Martin, N. G., & Miller, S. M. (2011). Psychiatric and genetic studies of binocular rivalry: an endophenotype for bipolar disorder? *Acta Neuropsychiatrica*, 23(1), 37-42.
- Noest, A. J., Van Ee, R., Nijs, M. M., & Van Wezel, R. J. A. (2007). Percept-choice sequences driven by interrupted ambiguous stimuli: a low-level neural model. *Journal of Vision*, 7(8: 10), 1-14. doi:10.1167/7.8.10
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113. doi:10.1016/0028-3932(71)90067-4
- Overgaard, M. (2015). *Behavioural methods in consciousness research*: Oxford University Press.
- Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, 66(232-258). doi:10.1016/j.cogpsych.2012.12.002
- Paffen, C. L., Alais, D., & Verstraten, F. A. (2006). Attention speeds binocular rivalry. *Psychological science*, 17(9), 752-756.
- Palmiero, M., Piccardi, L., Nori, R., Palermo, L., Salvi, C., & Guariglia, C. (2016). Creativity and mental imagery. *Front Psychol*, 7. doi:doi.org/10.3389/fpsyg.2016.01280
- Pastukhov, A., & Braun, J. (2007). Perceptual reversals need no prompting by attention. *Journal of Vision*, 7(10: 5). doi:10.1167/7.10.5
- Pastukhov, A., García-Rodríguez, P. E., Haenicke, J., Guillamon, A., Deco, G., & Braun, J. (2013). Multi-stable perception balances stability and sensitivity. *Frontiers in Computational Neuroscience*, 7(17). doi:10.3389/fncom.2013.00017
- Pettigrew, J. D., & Miller, S. M. (1998). A 'sticky' interhemispheric switch in bipolar disorder? *Proceedings of the Royal Society B: Biological Sciences*, 265(1411), 2141-2148. doi:10.1098/rspb.1998.0551
- Plude, D. J., Enns, J. T., & Brodeur, D. (1994). The development of selective attention: A life-span overview. *Acta Psychologica*, 86(2-3), 227-272. doi:10.1016/0001-6918(94)90004-3
- Pomerantz, J. R., & Kubovy, M. (1981). Perceptual organization: An overview. *Perceptual organization*, 423-456.
- Porporino, M., Iarocci, G., Shore, D. I., & Burack, J. A. (2004). A developmental change in selective attention and global form perception. *International Journal of Behavioral Development*, 28(4), 358-364. doi:10.1080/01650250444000063

- Pressnitzer, D., & Hupé, J. M. (2006). Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Current Biology*, 16(13), 1351-1357. doi:10.1016/j.cub.2006.05.054
- Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and cognition*, 13(2), 253-262.
- Quinn, H., & Arnold, D. H. (2010). Binocular rivalry and multi-stable perception: Independence and monocular channels. *Journal of Vision*, 10(10), 8-8. doi:10.1167/10.10.8
- Rankin, J., Sussman, E., & Rinzel, J. (2015). Neuromechanistic model of auditory bistability. *PLoS computational biology*, 11(11), e1004555. doi:10.1371/journal.pcbi.1004555
- Rayner, K., & Clifton, C., Jr. (2009). Language processing in reading and speech perception is fast and incremental: implications for event-related potential research. *Biol Psychol*, 80(1), 4-9. doi:10.1016/j.biopsycho.2008.05.002
- Ricci, C., & Blundo, C. (1990). Perception of Ambiguous Figures after Focal Brain-Lesions. *Neuropsychologia*, 28(11), 1163-1173. doi:10.1016/0028-3932(90)90052-P
- Rock, I. (1956). The orientation of forms on the retina and in the environment. *American Journal of Psychology*, 69(513-528).
- Rock, I., Gopnik, A., & Hall, S. (1994). Do young children reverse ambiguous figures? *Perception*, 23(6), 635-644. doi:10.1068/p230635
- Rubin, E. (1915). Synsoplevede Figurer: Studier i psykologisk Analyse. Første Del [Visually experienced figures: Studies in psychological analysis. Part one]. *Copenhagen and Christiania: Gyldendalske Boghandel*.
- Rubin, N., & Hupé, J. M. (2004). Dynamics of Perceptual Bistability: Plaids and Binocular Rivalry Compared In D. Alais & R. Blake (Eds.), *Binocular rivalry*. MIT Press.
- Sandberg, K., Blicher, J. U., Del Pin, S. H., Andersen, L. M., Rees, G., & Kanai, R. (2016). Improved estimates for the role of grey matter volume and GABA in bistable perception. *Cortex*, 83, 292-305.
- Sanders, R. D., Winston, J. S., Barnes, G. R., & Rees, G. (2018). Magnetoencephalographic Correlates of Perceptual State During Auditory Bistability. *Scientific reports*, 8(1), 976. doi:10.1038/s41598-018-19287-0
- Schielzeth, H., & Forstmeier, W. (2008). Conclusions beyond support: overconfident estimates in mixed models. *Behavioral Ecology*, 20(2), 416-420.
- Schooler, J. W., & Melcher, J. (1995). The ineffability of insight. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 97-133). Cambridge: MIT press.
- Schwartz, J. L., Grimault, N., Hupé, J. M., Moore, B. C. J., & Pressnitzer, D. (2012). Multistability in perception: binding sensory modalities, an overview. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 367(1591), 896-905. doi:10.1098/rstb.2011.0254
- Schwarzkopf, D. S., & Rees, G. (2015). Perceptual organization and consciousness. In J. Wagemans (Ed.), *The Oxford Handbook of Perceptual Organization* (pp. 799-839): OUP Oxford.
- Scocchia, L., Valsecchi, M., & Triesch, J. (2014). Top-down influences on ambiguous perception: the role of stable and transient states of the observer. *Frontiers in Human Neuroscience*, 8. doi:10.3389/fnhum.2014.00979
- Shannon, R. W., Patrick, C. J., Jiang, Y., Bernat, E., & He, S. (2011). Genes contribute to the switching dynamics of bistable perception. *Journal of Vision*, 11(3), 8-8.
- Sharma, M. A., & Singh, J. B. (2010). Use of probability distribution in rainfall analysis. *New York Science Journal*, 3(9), 40-49.

- Sheppard, B. M., & Pettigrew, J. D. (2006). Plaid motion rivalry: correlates with binocular rivalry and positive mood state. *Perception*, 35(2), 157-169.
- Silvanto, J., & Pascual-Leone, A. (2008). State-Dependency of Transcranial Magnetic Stimulation. *Brain Topography*, 21(1), 1-10. doi:10.1007/s10548-008-0067-0
- Simpson, A., & Riggs, K. J. (2005). Inhibitory and working memory demands of the day-night task in children. *British Journal of Developmental Psychology*, 23, 471-486. doi:10.1348/026151005x28712
- Singer, N., Eapen, M., Grillon, C., Ungerleider, L. G., & Hendler, T. (2012). Through the eyes of anxiety: Dissecting threat bias via emotional-binocular rivalry. *Emotion*, 12(5), 960. doi:10.1037/a0027070
- Snyder, J. S., Gregg, M. K., Weintraub, D. M., & Alain, C. (2012). Attention, awareness, and the perception of auditory scenes. *Front Psychol*, 3, 15. doi:10.3389/fpsyg.2012.00015
- Sowell, E. R., Thompson, P. M., & Toga, A. W. (2004). Mapping changes in the human cortex throughout the span of life. *The Neuroscientist*, 10(4), 372-392. doi:10.1177/1073858404263960
- Spilke, J., Piepho, H. P., & Hu, X. (2005). Analysis of unbalanced data by mixed linear models using the MIXED procedure of the SAS system. *Journal of Agronomy and Crop Science*, 191(1), 47-54.
- Sterzer, P., & Kleinschmidt, A. (2007). A neural basis for inference in perceptual ambiguity. *Proceedings of the National Academy of Sciences of the United States of America*, 104(1), 323-328.
- Sterzer, P., Kleinschmidt, A., & Rees, G. (2009). The neural bases of multistable perception. *Trends in Cognitive Sciences*, 13(7), 310-318. doi:10.1016/j.tics.2009.04.006
- Stroop, J. R. (1935). Studies of inference in serial verbal reactions. *Journal of Experimental Psychology*, 18(643-662).
- Strüber, D., Basar-Eroglu, C., Hoff, E., & Stadler, M. (2000). Reversal-rate dependent differences in the EEG gamma-band during multistable visual perception. *International Journal of Psychophysiology*, 38(3), 243-252. doi:10.1016/S0167-8760(00)00168-9
- Struber, D., & Stadler, M. (1999). Differences in top-down influences on the reversal rate of different categories of reversible figures. *Perception*, 28(10), 1185-1196. doi:10.1068/p2973
- Strüber, D., & Stadler, M. (1999). Differences in top-down influences on the reversal rate of different categories of reversible figures. *Perception*, 28(10), 1185-1196. doi:10.1068/p2973
- Sussman, E., Wong, R., Horvath, J., Winkler, I., & Wang, W. (2007). The development of the perceptual organization of sound by frequency separation in 5-11-year-old children. *Hear Res*, 225(1-2), 117-127. doi:10.1016/j.heares.2006.12.013
- Suzuki, S., & Grabowecky, M. (2007). Long-term speeding in perceptual switches mediated by attention-dependent plasticity in cortical visual processing. *Neuron*, 56(4), 741-753. doi:10.1016/j.neuron.2007.09.028
- Taddei-Ferretti, C., Radilovac, J., Musioa, C., Santilloa, S., Cibellib, E., Cotugnoa, A., & Radilc, T. (2008). The effects of pattern shape, subliminal stimulation, and voluntary control on multistable visual perception. *Brain Research*, 1225, 163-170. doi:10.1016/j.brainres.2008.04.064
- Team, R. c. d. (2015). Vienna, Austria: R Foundation for Statistical Computing Patent No.
- Thompson, P. D., Colebatch, J. G., Brown, P., Rothwell, J. C., Day, B. L., Obeso, J. A., & Marsden, C. D. (1992). Voluntary stimulus- sensitive jerks and jumps mimicking myoclonus or pathological startle syndromes. *Movement Disorders*, 7(3), 257-262.

- Toga, A. W., Thompson, P. M., & Sowell, E. R. (2006). Mapping brain maturation. *Trends in Neurosciences*, 29(3), 148-159. doi:10.1016/j.tins.2006.01.007
- Tong, F. (2001). Competing theories of binocular rivalry: a possible resolution". *Brain and Mind*, 2, 55-83.
- Tong, F., Meng, M., & Blake, R. (2006). Neural bases of binocular rivalry. *Trends in Cognitive Sciences*, 10(11), 502-511. doi:10.1016/j.tics.2006.09.003
- Toppino, T. C., & Long, G. M. (2004). Top-down and bottom-up processes in the perception of reversible figures: Toward a hybrid model. *Dynamic Cognitive Processes*, 37-58. doi:10.1007/4-431-27431-6_3
- Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology*, 11(1), 138-146. doi:10.1037//0894-4105.11.1.138
- Van de Cruys, S., de-Wit, L., Evers, K., Boets, B., & Wagemans, J. (2013). Weak priors versus overfitting of predictions in autism: Reply to Pellicano and Burr (TICS, 2012). *i-Perception*, 4(2), 95-97.
- van Ee, R. (2005). Dynamics of perceptual bi-stability for stereoscopic slant rivalry and a comparison with grating, house-face, and Necker cube rivalry. *Vision Research*, 45, 29-40. doi:10.1016/j.visres.2004.07.039
- van Ee, R. (2009). Stochastic variations in sensory awareness are driven by noisy neuronal adaptation: evidence from serial correlations in perceptual bistability. *Journal of the Optical Society of America a-Optics Image Science and Vision*, 26(12), 2612-2622.
- van Ee, R., van Dam, L. C. J., & Brouwer, G. J. (2005). Voluntary control and the dynamics of perceptual bi-stability. *Vision Research*, 45(1), 41-55. doi:10.1016/j.visres.2004.07.030
- van Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. Technical University of Eindhoven, Eindhoven. Retrieved from <http://alexandria.tue.nl/extra1/PRF2A/7707058.pdf>, Technical University of Eindhoven, Eindhoven database.
- Vernet, M., Brem, A. K., Farzan, F., & Pascual-Leone, A. (2015). Synchronous and opposite roles of the parietal and prefrontal cortices in bistable perception: A double-coil TMS-EEG study. *Cortex*, 64, 78-88. doi:10.1016/j.cortex.2014.09.021
- Vierck, E., Porter, R. J., Luty, S. E., Moor, S., Crowe, M. T., Carter, J. D., & Joyce, P. R. (2013). Further evidence for slow binocular rivalry rate as a trait marker for bipolar disorder. *Australian & New Zealand Journal of Psychiatry*, 47(4), 371-379.
- Walker, P., & Powell, D. J. (1979). The sensitivity of binocular rivalry to changes in the nondominant stimulus. *Vision Research*, 19(247-249).
- Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, 45(4), 205-217.
- Wallach, M. A., & Kogan, N. (1965). A new look at the creativity-intelligence distinction. *Journal of Personality*, 33(3), 348-369. doi:10.1111/j.1467-6494.1965.tb01391.x
- Wallis, G., & Ringelhan, S. (2013). The dynamics of perceptual rivalry in bistable and tristable perception. *Journal of Vision*, 13(2: 24), 1-21. doi:10.1167/13.2.24
- Wang, W. (2013). Identifiability of linear mixed effects models. *Electronic Journal of Statistics*, 7, 244-263.
- Warren, R. M., & Gregory, R. L. (1958). An auditory analogue of the visual reversible figure. *The American Journal of Psychology*, 71(3), 612-613. doi:10.2307/1420267
- Warren, R. M., & Warren, R. P. (1966). A comparison of speech perception in childhood, maturity and old age by means of the verbal transformation effect. *Journal of Verbal Learning and Verbal Behavior*, 5(2).

- Weilnhammer, V., Ludwig, K., Hesselmann, G., & Sterzer, P. (2013). Frontoparietal cortex mediates perceptual transitions in bistable perception. *Journal of Neuroscience*, 33(40), 16009-16015.
- Weilnhammer, V., Stuke, H., Hesselmann, G., Sterzer, P., & Schmack, K. (2017). A predictive coding account of bistable perception-a model-based fMRI study. *PLoS computational biology*, 13(5), e1005536. doi:10.1371/journal.pcbi.1005536
- Wheatstone, C. (1838). Contributions to the Physiology of Vision. Part the First. On Some Remarkable, and Hitherto Unobserved, Phenomena of Binocular Vision. *Phil. Trans. R. Soc. Lond.*(128), 371-394. doi:10.1098/rstl.1838.0019
- Wilson, H. R. (2007). Minimal physiological conditions for binocular rivalry and rivalry memory. *Vision Research*, 47(21), 2741–2750. doi:doi:10.1016/j.visres.2007.07.007
- Wimmer, M. C., & Doherty, M. J. (2011). The Development of Ambiguous Figure Perception. *Monographs of the Society for Research in Child Development*, 76(1), 1-130. doi:10.1111/j.1540-5834.2011.00589.x
- Wimmer, M. C., & Marx, C. (2014). Inhibitory processes in visual perception: A bilingual advantage. *Journal of Experimental Child Psychology*, 126, 412-419. doi:10.1016/j.jecp.2014.03.004
- Windmann, S., Wehrmann, M., Calabrese, P., & Gunturkun, O. (2006). Role of the prefrontal cortex in attentional control over bistable vision. *Journal of Cognitive Neuroscience*, 18(3), 456-471. doi:10.1162/089892906775990570
- Wiseman, R., Watt, C., Gilhooly, K., & Georgiou, G. (2011). Creativity and ease of ambiguous figural reversal. *British Journal of Psychology*, 102, 615-622. doi:10.1111/j.2044-8295.2011.02031.x
- Wolf, M., & Hochstein, S. (2011). High-level binocular rivalry effects. *Frontiers in Human Neuroscience*, 5. doi:10.3389/fnhum.2011.00129
- Wood, D., Schauer, G., Bak, T., & Carmel, D. (in preparation). Parietal TMS and binocular rivalry: Stimulus and timing dependence resolve contradictory findings.
- Wu, X., Gu, X., & Zhang, H. (2016). The Facilitative Effects of Ambiguous Figures on Creative Solution. *The Journal of Creative Behavior*, 0(0), 1-8. doi:doi:10.1002/jocb.161
- Yu, T. Y., Jacobs, R. J., Anstice, N. S., Paudel, N., Harding, J. E., & Thompson, B. (2013). Global motion perception in 2-year-old children: a method for psychophysical assessment and relationships with clinical measures of visual function. *Investigative ophthalmology & visual science*, 54(13), 8408-8419. doi:10.1167/iovs.13-13051
- Zaretskaya, N., Thielscher, A., Logothetis, N. K., & Bartels, A. (2010). Disrupting parietal function prolongs dominance durations in binocular rivalry. *Curr Biol*, 20(23), 2106-2111. doi:10.1016/j.cub.2010.10.046
- Zhou, W., & Chen, D. (2009). Binocular rivalry between the nostrils and in the cortex. *Current Biology*, 19(18), 1561-1565.
- Zhou, Y. H., Gao, J. B., White, K. D., Merk, I., & Yao, K. (2004). Perceptual dominance time distributions in multistable visual perception. *Biological Cybernetics*, 90(4), 256-263. doi:10.1007/s00422-004-0472-8
- Zuur, A. F., Ieno, E., I., Walker, N., Saveliev, A., A., & Smith, G. M. (2009). *Mixed Effects Models and Extensions in Ecology with R with R. Gail M, Krickeberg K, Samet JM, Tsiatis A, Wong W, editors*. New York, NY: Springer Science and Business Media.