Aus der Klinik für Psychiatrie und Psychotherapie

Direktor: Prof. Dr. Tilo Kircher

des Fachbereichs Medizin der Philipps-Universität Marburg

A Novel Intervention Approach Focusing on Social Communicative Functioning in Patients with Schizophrenia Spectrum Disorder: Effects of a Specific Speech-Gesture Training on Quality of Life and Neural Processing

Ein neuer Interventionsansatz mit dem Fokus auf den sozial-kommunikativen Fähigkeiten von Patient*innen mit Schizophrenie: Effekte eines spezifischen Sprach-Gestik-Trainings auf die Lebensqualität und die neurale Verarbeitung

Inaugural-Dissertation zur Erlangung des Doktorgrades der Medizinwissenschaften
dem Fachbereich Medizin der Philipps-Universität Marburg vorgelegt von

Lydia Riedl (geb. Rückert) aus Gera

Marburg, 2022

Angenommen vom Fachbereich Medizin der Philipps-Universität Marburg am: 05.04.2022

Gedruckt mit Genehmigung des Fachbereichs Medizin

Dekanin:	Prof. Dr. Denise Hilfiker-Kleiner
Referent:	Prof. Dr. Benjamin Straube
1. Korreferent:	Prof. Dr. Inge Kamp-Becker

List of Contents

List of Figures

List of Tables

List of Abbreviations

AAL	Automated Anatomical Labeling toolbox (implemented in SPM)
abs	Abstractness condition: abstract speech content
ACC	Anterior cingulate and paracingulate gyri
AC-PC	Anterior commissure posterior commissure
AFNI	Analysis of Functional NeuroImages (Program)
ANCOVA	Analysis of covariance
ANTs	Advanced Normalization Tools
aCompCor	Anatomical component-based noise correction
B_0	The constant, homogeneous magnetic field during an fMRI scan
BIDS	Brain Imaging Data Structure
BIDS App	Portable neuroimaging pipelines that understand BIDS datasets
BIDSonym	BIDSapp for the pseudo-anonymization of neuroimaging datasets
bl	Session baseline: fMRI measurement + clinical questionnaire assessment after a waiting period (TAU) and before the MSG training started
BOLD	Blood oxygenation level-dependent
CER	Cerebellum
CERCRU1	Crus I of cerebellar hemisphere
con	Abstractness condition: concrete speech content
CPZ	Chlorpromazine
CSF	Cerebrospinal fluid
CUN	Cuneus
DIM	Dimensions
DVARS	The spatial standard deviation of successive difference images
EPI	Echo planar imaging
FD	Framewise displacement
FFG	Fusiform gyrus
fMRI	functional Magnetic Resonance Imaging
fMRIPrep	A Robust Preprocessing Pipeline for fMRI Data
FoV	Field of view

FSL	FMRIB Software Library (Program)
FWHM	Full width at half maximum
G	Modality condition: gesture only (regarding the video stimuli)
GLM	General Linear Model
GM	Grey-matter
GRAPPA	GeneRalized Autocalibrating Partial Parallel Acquisition
HIP	Hippocampus
HRF	Hemodynamic response function
IFG	Inferior frontal gyrus
INS	Insula
INU	Intensity non-uniformity
ITG	Inferior temporal gyrus
k_E	Cluster size
L	Left (here: left hemisphere)
MCC	Middle cingulate and paracingulate gyri
MFG	Middle frontal gyrus
MTG	Middle temporal gyrus
MNI	Montreal Neurological Institute
MRI	Magnetic Resonance Imaging
MRIQC	Automatic Prediction of Quality and Visual Reporting of MRI Scans
MSG training	Multimodal speech-gesture training
MTG	Middle temporal gyrus
MUC	Memory Unification Control (a neurobiological model of language)
NIfTI	Neuroimaging Informatics Technology Initiative file format
PCUN	Precuneus
PHG	Parahippocampal gyrus
post	Session post: fMRI measurement + clinical questionnaire assessment after the MSG training
pre	Session pre: fMRI measurement + clinical questionnaire assessment before a waiting period (TAU) and before the MSG training started
SWLS	Satisfaction With Life Scale

tDCS	Transcranial Direct Current Stimulation
TE	Echo time
TMP	Temporal pole
TMS	Transcranial magnetic stimulation
TR	Repetition time
TrIFG	Triangular part of inferior frontal gyrus
TULIA	Test of Upper Limb Apraxia

1 Introduction

1.1 Overview

Dysfunctional social communication is one of the most stable characteristics in patients with mental disorders and especially patients with schizophrenia spectrum disorder (SSD) that also affects quality of life. Interpreting abstract speech and integrating nonverbal modalities is particularly affected. Considering the dysfunctions in sensorimotor systems in schizophrenia, recently added to the Research Domain Criteria, and the difficulty to treat communication dysfunctions with usual treatment, we investigated the possibility to improve verbal and non-verbal communication in schizophrenia by applying a multimodal speech-gesture training (MSG training).

In the MSG training intervention, we offered eight sessions (60 min each) of training. The intervention contained perceptive rating (match/mismatch of sentence and gesture) and memory tasks, imitation and productive tasks (e.g., speech-gesture fluency where the patients were asked to produce as many word-gesture pairs as possible within one minute). In addition, we offered information about gesture as meta-learning element as well as homework for reasons of transfer to everyday life as part of every session. Outcomes were measured through pre-post-fMRI and standardized psychological as well as specifically outlined questionnaires comparing two subject groups (29 patients with schizophrenia spectrum disorder and 17 healthy controls). We focused on the generalization effects of the training, mainly the effects on quality of life, and their neural correlates, to get a more comprehensive understanding of the training's effects.

At the first measurement (pre-fMRI), SSD patients showed significantly reduced quality of life compared to healthy controls but significantly improved quality of life during the training. Strikingly, this improvement correlates with MSG training related neural activation changes in middle temporal regions bilateral and (para-)hippocampal regions for the processing of abstract multimodal content.

Despite a number of limitations in the study, the overall analyses provide extraordinary promising results. Especially the transfer effects of the MSG training into the patients' everyday life provides evidence for the beneficial effects of innovative add-on treatments for patients with schizophrenia. As communication is the basis for social interaction and hence for quality of life, further research in this field is highly needed. With this study, we were able to provide implications for future implementation of specific speech and gesture intervention programs in order to improve social functioning and quality of life in SSD patients.

1.2 State of Research

1.2.1 SSD: Schizophrenia Spectrum Disorder

Schizophrenia spectrum disorder (SSD) is a chronic, often devastating disease with an estimated lifetime prevalence up to 1% (Castillejos et al., 2019; Gaebel & Wölwer, 2010; Saha et al., 2005). It affects approximately 21 million people around the globe and was the 12th most disabling disorder in the Global Burden of Disease Study 2016, accounting for 1,7% of all years lived with disability globally (Charlson et al., 2018). According to the diagnosis data of German hospitals, schizophrenic disorders (F20-F29 according to ICD-10) accounted for about 13% of the annual number of inpatient mentally ill patients and about 21% of the annual number of treatment days in 2007 (*Diagnosedaten der Patientinnen und Patienten in Krankenhäusern*, 2008), resulting in total annual costs of schizophrenic disorders for the health and social system in Germany between 4.4 and 9.2 billion euros (Kissling et al., 1999). Thus, the estimated costs for SSD amount to approximately 2% to 4% of the total costs for health care services in Germany. In terms of both direct and indirect costs (e.g., due to the patients' productivity losses), they are comparable to or even higher than those of common somatic diseases (diabetes, cardiovascular diseases) (Gaebel & Wölwer, 2010).

The schizophrenia spectrum is characterized by acute psychoses of episodic nature (often accompanied with positive symptoms such as delusions, hallucinations, disorganized speech, disorganized, or catatonic behavior) as well as variably chronic or remitting psychotic, cognitive and affective symptoms (mainly negative symptoms, e.g., apathy, anhedonia, and alogia). Life expectancy in patients suffering from SSD is reduced by 15 to 20 years (Gaebel & Wölwer, 2010). Considering the high individual and socio-economic burden of schizophrenia, further research is urgently needed to understand pathomechanisms and enable earlier diagnosis and adequate treatments.

The treatment of SSD is based on a multimodal and multi-professional concept and includes pharmacotherapy, psychotherapeutic, social, and rehabilitation therapies. The conventional medical treatment is often (Hegarty et al., 1994; Jääskeläinen et al., 2013) successful in treating positive symptoms, but especially negative symptoms and formal thought disorders concerning social-communicative skills remain relatively stable (Dollfus & Petit, 1995; Gaebel & Wölwer, 2010; Joyal et al., 2016; Lavelle et al., 2014; Wüthrich et al., 2020). On the opposite, deficits are also observed in patients without any medication (Walther, Mittal, et al., 2020), suggesting that antipsychotic medication is unlikely to account for these deficits.

1.2.2 Speech Processing

Communication is the fundamental basis of social life. According to the classic Wernicke-Lichtheim-Geschwind model, speech production is secondary to speech perception (Geschwind, 1970). In speech perception, processing abstract linguistic concepts is a particularly complex cognitive challenge in interpersonal communication. The correct interpretation of any abstract concepts requires the ability to comprehend additional meanings not directly expressed in an utterance, for example conversational implicatures (Grice, 1975). The understanding of figurative expressions like metaphors represents a special case of conversational implicature comprehension, in which it is required to go beyond the literal meaning to interpret the utterance correctly (Rossetti et al., 2018). In everyday life, metaphors are very frequently used to refer to abstract concepts such as feelings or events (e.g., “out of the blue”). According to the Conceptual Metaphor Theory, conceptual (or cognitive) metaphors refer to the understanding of one (abstract) concept in terms of another (concrete) concept (Lakoff & Johnson, 2008). Conceptual metaphors are very common in language and are mostly used unconsciously. They are very useful or, according to Lakoff and Johnson, necessary for understanding complex ideas in simple terms. E.g., to get an idea of the quite abstract concept of time, we use the cognitive metaphor TIME IS MONEY, for example in sentences like "I spent a lot of time in writing this thesis." Here, some aspects of the so called *source domain*, MONEY, are mapped to the so called *target domain*, TIME (see *figure 1*). Through this metaphor we can conclude, that time has some kind of value and that we can "buy" things with our time. So it is necessary to understand metaphors, not only because they occur very frequently in our everyday communication and we can interpret figurative speech only with the help of cognitive metaphors, but, according to Lakoff and Johnson, metaphors also conceive the very way we actually perceive and act. Lakoff and Johnson would conclude that the conceptual metaphor TIME IS MONEY could be associated with the western cultural tradition to pay employees per working hours. Because the Conceptual Metaphor Theory suggests that (in most cases) more abstract ideas can be interpreted (or *grasped*) by mapping aspects of a more concrete idea, this theory is situated in the school of thought of embodied cognition (Joue et al., 2020; Santana & De Vega, 2011), which argues that higher cognitive processes are grounded in sensorimotor experiences and therewith can influence perception and general cognition (for a review of different views, see Pecher et al. (2011)).

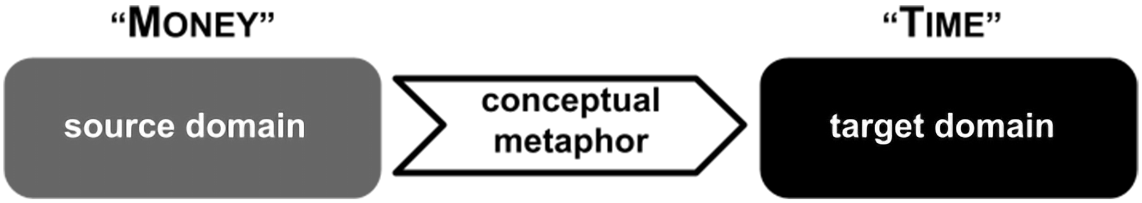


Figure 1: Illustration of the cognitive metaphor TIME IS MONEY according to Lakoff and Johnson. Some aspects of the *source domain* (here: MONEY), e.g., the concept of paying, are mapped to the *target domain* (here: TIME).

1.2.3 Gesture Processing

As a further integral feature of embodiment (Tschacher et al., 2017), gesture "serves as an outward manifestation of several interacting fundamental processes" (Walther & Mittal, 2016) including speech perception (S. D. Kelly et al., 2010; Wu & Coulson, 2010), memory (Straube et al., 2009; Straube, Green, Chatterjee, et al., 2011), and social functioning (Bucci et al., 2008; Goldin-Meadow, 2017; Straube et al., 2010). The ability to combine information from multiple sensory modalities is a key element in a person’s ability of interpersonal communication (Stevenson et al., 2011).

Gestures are not redundant in face-to-face communication but transport additional information: They are used to express information by extending and/or emphasizing a verbal utterance and by this contribute meaning in the communication process (Andric & Small, 2012; Goldin-Meadow, 2005). Thus, gestures can disambiguate speech (Driskell & Radtke, 2003; Holle & Gunter, 2007; S. D. Kelly et al., 1999, 2010), increase attention (Maricchiolo et al., 2009) and play a crucial role in 'grounding' (ensuring comprehension between conversational partners) (Bavelas et al., 2011; Nathan & Alibali, 2011) as well as in turn taking (Mondada & Oloff, 2011). Gestures therewith have a huge impact on comprehension in the listener (Hostetter, 2011). These effects seem to play a role in language acquisition (Tomasello et al., 2019), e.g., by creating 'joint actions' (Pereira, 2015). Gesturing also contributes to the gesturer's thinking and speaking: Gestures aid within the production of spoken language (Driskell & Radtke, 2003; Goldin-Meadow & Alibali, 2013) and facilitate lexical access and therewith the fluency/stream of speech (Morsella & Krauss, 2004; Pine et al., 2013; Skipper et al., 2007; Yap et al., 2011).

We can classify gestures in intrinsically meaningful gestures (e.g., *emblems*, such as the THUMBS-UP gesture and *tool-use* gestures, such as hand movements of HAMMERING) and in gestures that can only be understood in the context of verbal speech (He et al., 2015). The latter include for example iconic and metaphorical gestures. *Iconic* gestures accompany concrete spoken concepts (e.g., forming the shape of a dog's mouth with a hand while discussing a dog). *Metaphorical* gestures in contrast are hand and arm movements that accompany abstract concepts in speech (e.g., forming a cup with a hand while discussing a concept such as love) (McNeill, 1992). In the case of metaphorical gesture use, words and gestures can serve different expressive functions: While the words spoken make the *target domain* of a cognitive metaphor (in our example, TIME) explicit, the gesture illustrates a metaphorical way of thinking about this target by depicting the *source domain* (in our example, MONEY, see *figure 2*) (Cienki, 2008).

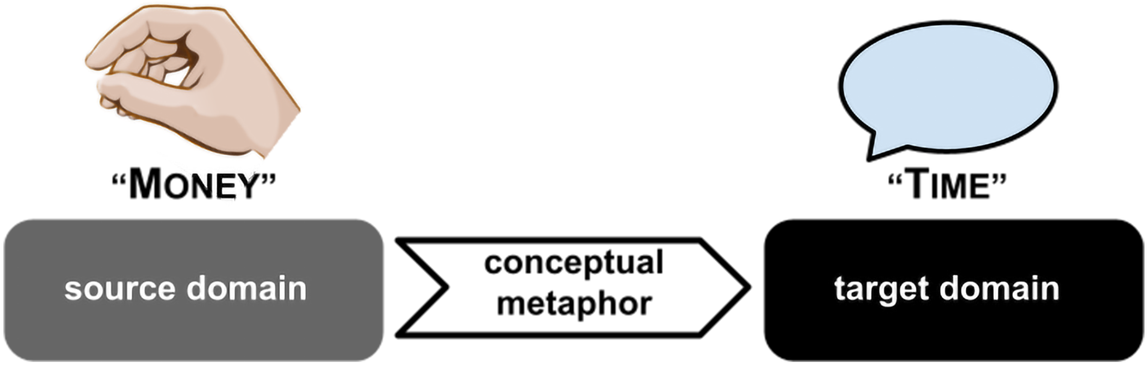


Figure 2: Illustration of how two channels of communication (speech and gesture) can serve as different expressive functions: While speech expresses the abstract *target domain* (TIME) in the conceptual metaphor TIME IS MONEY, the gesture reveals the underlying concrete *source domain* (MONEY).

1.2.4 Speech Processing in SSD

In schizophrenia, language disturbances are key signs (Rossetti et al., 2018), which are manifested at all linguistic levels. In acute phases, communication can be extremely limited. Some studies actually report parallels of schizophrenic speech and aphasia (Landre et al., 1992), e.g., patients show dysfunctional embedding, a limited lexical retrieval and neologisms (Covington et al., 2005; DeLisi, 2001; Heim et al., 2019; Marini et al., 2008). Therefore, referring to the term 'aphasia', the productive speech of SSD patients is called 'schizophasia' (Covington et al., 2005). Further linguistic impairments are for instance an incoherent discourse, less referential cohesion and loosening of associations. Aberrations in language development can already be detected in children who are diagnosed with schizophrenia later in life (Jones et al., 1994; Reichenberg et al., 2002). These symptoms of disordered speech are evidence of formal thought disorder (Cavelti et al., 2018; Hinzen & Rosselló, 2015), which can also be a symptom of other psychiatric diseases, such as depression.

Not only the production of speech is impaired in schizophrenia, but also the general understanding and interpretation of the meaning of speech and language information is severely affected. Especially, patients show difficulties in the understanding of figurative speech or abstract concepts being conveyed through metaphors, proverbs, humorous or ironic expressions, which they tend to misinterpret in a concrete way (Bergemann et al., 2008; de Bonis et al., 1997; Iakimova et al., 2010; Kircher et al., 2007; A. Rapp & Schmierer, 2010). Hence, this phenomenon is clinically termed 'concretism' (Rossetti et al., 2018). In general, patients with schizophrenia often have difficulties to inhibit the dominant, but inappropriate meaning of a word ('strong meaning response bias') (Gernsbacher et al., 1999) and to integrate an abstract utterance into in the sentence context (Langdon et al., 2002). To stay with the Conceptual Metaphor Theory (Lakoff & Johnson, 2008), patients with 'concretism' have difficulties to understand an abstract concept (e.g., TIME) in terms of another, more concrete concept (e.g., MONEY), see *figure 3*. Patients with schizophrenia thus have disturbances at all levels of language processing - with a clear focus on (abstract) concept formation (Kircher & Gauggel, 2008).

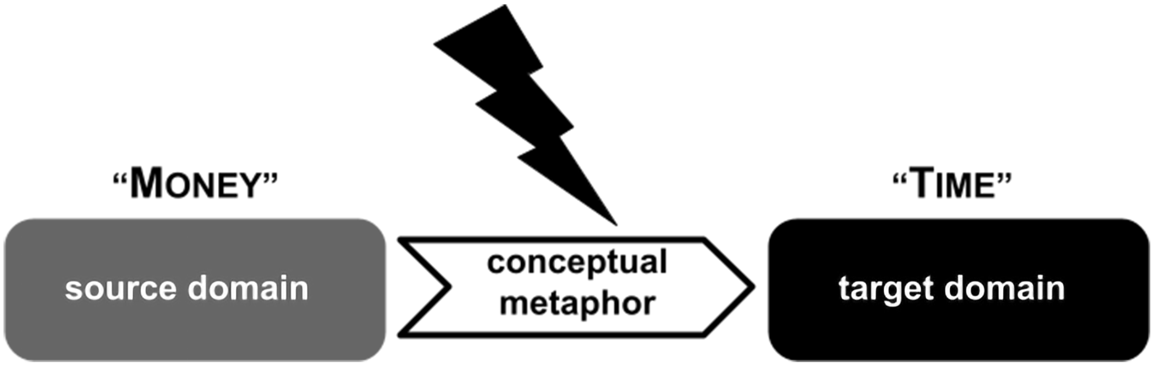


Figure 3: Illustration of disturbed conceptualization in patients with 'concretism', one of the most frequent language related symptoms in schizophrenia. In our exemplary conceptual metaphor TIME IS MONEY, aspects of the source domain (here: MONEY) cannot be transferred to the target domain (here: TIME). Thus, the metaphor might not be understood correctly in an abstract way, but the interpretation possibly remains concrete in patients with SSD.

By contrast, the understanding of concrete speech seems to be less affected in schizophrenia (He et al., 2021; Kircher et al., 2009; Straube et al., 2009; Straube, Green, Bromberger, et al., 2011a; Straube, Green, et al., 2013).

1.2.5 Gesture Processing in SSD

Dysfunction of the integration and interpretation of gesture information represents a further core feature of disordered communication processes in SSD. Aberrations in gesture processing are found across all stages of schizophrenia, in production as well as in perception and interpretation. There is converging evidence for disturbances in imitation and pantomime (Matthews et al., 2013; Walther et al., 2013a, 2013b, 2015) and incoherent, mismatching gesture use (Goss, 2011; Walther et al., 2013a). Furthermore, recognition of gestures is impaired in SSD (Berndl, Cranach, et al., 1986; Karakuła et al., 2013). When it comes to interpretation, patients tend to misinterpret gestures, often in a negative way and/or in form of referential delusions (Berndl, Cranach, et al., 1986; Bucci et al., 2008). Some authors therefore claim a general disruption of the integration of two modalities (auditory modality: speech and visual modality: gesture) in schizophrenia (Berndl, Cranach, et al., 1986; Bucci et al., 2008; P. Martin et al., 1994; Mittal et al., 2006; Troisi et al., 1998). Some studies relate working memory deficits to gestural malfunctioning (Matthews et al., 2013; Park et al., 2008). Gesture deficits in SSD are related to symptom severity, e.g., negative symptoms, hallucinations, and formal thought disorder (Bucci et al., 2008; Matthews et al., 2013; Nagels et al., 2019). Patient's gesture performance can even help predicting negative symptom progression and social functioning (Walther et al., 2016).

Patients with SSD particularly have trouble interpreting abstract meaning in gestures (Nagels et al., 2019; Straube, Green, et al., 2013; Straube et al., 2014). A recent study demonstrated aberrant eye-gaze behavior for videos with abstract co-verbal gestures already in subjects with high risk for schizophrenia (Gupta et al., 2021). Strikingly, integration skills interact with symptomatology: Nagels and colleagues found worse integration skills, reflected in the evaluation of the semantic relationship between speech and gesture, in patients with severe symptoms compared to patients with mild symptom severity in formal thought disorders (Nagels et al., 2019).

1.2.6 Speech, Gesture and Its Importance for Quality of Life

As outlined in *chapter 1.2.2 Speech Processing*, conceptual metaphors are frequently used in everyday communication. The 'concretism' symptom (see *chapter 1.2.4 Speech Processing in SSD*) thus has a tremendous impact on successful social interactions (Kircher & Gauggel, 2008; Lakoff & Johnson, 2008) and represents a real-world communication problem for patients with schizophrenia (Rossetti et al., 2018).

Considering furthermore gestures' fundamental role in social communication and functioning as discussed in *chapter 1.2.3 Gesture Processing*, integration of speech and gesture is highly important for the social inclusion of individuals (Goldin-Meadow & Alibali, 2013; Suffel et al., 2020). Particularly in SSD, impaired multimodal communication (see *chapter 1.2.5 Gesture Processing in SSD*) can have wide-ranging effects on the patients' social-communicative functioning (Walther & Mittal, 2016).

Since communication and interpersonal skills are put forward to play a crucial role in social integration, communication impairments are hence a central issue in modern psychiatry (Joyal et al., 2016) and can have a serious impact on the psychiatrist-patient communication (McCabe et al., 2013) with further negative consequences for rehabilitation: According to the German Robert Koch Institute, impairments in communication and social skills are a main factor responsible for SSD patients not being in a stable partnership and for unsuccessful educational careers with only 30% of the patients being actually employed (Gaebel & Wölwer, 2010). The limited participation on both personal and professional levels results in a dramatically reduced quality of life (Bambini et al., 2016; Falkai, 2016; Gaebel & Wölwer, 2010).

1.2.7 Speech, Gesture and Its Neural Correlates

As outlined above, decoding of abstract meaning is particularly challenging for patients with SSD in multiple modalities (see *chapters 1.2.4 Speech Processing in SSD* and *1.2.5 Gesture Processing in SSD*). Even though gestures convey meaning via the visual modality and speech is transmitted via the auditory modality, both modalities form a coupled communication system rather than two distinct processes and therefore are integrated into a unified framework (McNeill, 1992). The binding of information from multiple modalities is a complex higher order cognitive process that requires semantic integration (Green et al., 2009; He et al., 2018; Özyürek et al., 2007) as described in the 'memory unification control' (MUC) model (Hagoort, 2013; Hagoort et al., 2009; Holler & Levinson, 2019; Willems & Hagoort, 2007), shown in *figure 4*. The yellow regions (temporal cortex and angular gyrus) store fundamental information related to speech, such as phonological word forms, morphological information as well as specific syntactic templates of parts of speech. The blue regions (Broca's area and adjacent cortex) are mainly involved in so called unification processes according to the MUC model. These processes are crucial for generating larger structures in terms of semantic, syntactic or phonological unification operations. The grey regions (dorsolateral prefrontal cortex) are responsible for executive control (e.g., selecting the correct language or turn taking). They are also relevant for paying attention to the most relevant information in the input (selection processes). Besides these regions, midline structures including the anterior cingulate cortex and parts of parietal cortex are involved in attention (not shown in the figure).

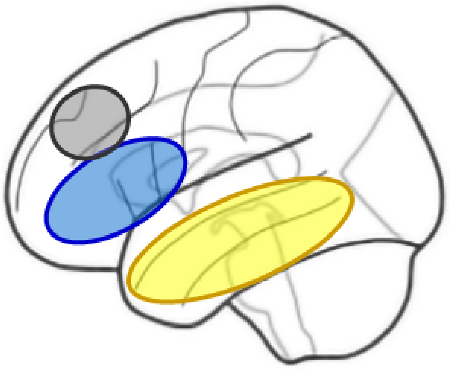


Figure 4: The three components of the MUC model of language. The figure displays a lateral view of the left hemisphere: Memory (yellow) in left temporal cortex, Unification (blue) in l.IFG, and Control (grey) in dorsolateral prefrontal cortex. The anterior cingulate cortex (part of the Control component) is not shown.

Figure based on Hagoort (2005).

l.IFG = left inferior frontal gyrus.

These regions play an important role in complex integration processes: Independent from modality, temporal regions, specifically the superior temporal sulcus (STS) and middle temporal gyrus (MTG), show activation for general integration processes and perceptual-matching processes (Dick et al., 2012; Green et al., 2009; Joue et al., 2020; Kircher et al., 2009; Straube, Green, Bromberger, et al., 2011a; Straube, Green, et al., 2013; Straube et al., 2014) and the inferior frontal gyrus (IFG) is involved in complex higher-order processes such as the unification of disparate semantic information in abstract speech-gesture pairs (Dick et al., 2009, 2012; He et al., 2015, 2018; Kircher et al., 2009; Straube, Green, Bromberger, et al., 2011a; Straube, Green, et al., 2013).

Compared to healthy controls, patients with SSD show abnormal activation, mostly in fronto-temporal regions, during the interpretation of metaphorical speech (Kircher et al., 2007; Rossetti et al., 2018). During the integration of metaphoric gestures, patients with schizophrenia also show abnormal activation of and connectivity between specific regions, especially in the IFG and middle and superior temporal areas (Straube, Green, et al., 2013; Straube et al., 2014) as well as globally in the praxis network (Viher et al., 2020). Strikingly, aberrant neural activation as well as reduced behavioral performance in communication tasks have been found in medicated patients with chronic schizophrenia in most of the studies and hence seem to be stable despite medication.

On the other hand, more recent findings suggest that the gesture-speech integration network (involving the posterior STS, visual occipital cortex, and auditory cortex) is functional in patients with schizophrenia. Only the verbal pathway between left MTG and STS seems to be impeded in SSD patients, which correlates with concretism ratings and negative symptom rating scores (Wroblewski et al., 2020).

1.2.8 Current Treatments for Patients With SSD

Symptoms in SSD concerning social-communicative skills remain relatively stable despite medication (Dollfus & Petit, 1995b; Gaebel & Wölwer, 2010a; Joyal et al., 2016a; Lavelle et al., 2014a; Wüthrich et al., 2020a). Nowadays, several therapeutic programs exist which can efficiently complement the therapy of patients with schizophrenia (Falkai, 2016), including psychotherapy (e.g., cognitive behavioral therapy), occupational therapy (Buchain et al., 2003; Liberman et al., 1998), physical therapy (Vancampfort et al., 2012), dance and movement therapy (L. A. L. Martin et al., 2016), art therapy (Crawford & Patterson, 2007) and music therapy (Gold et al., 2009; Peng et al., 2010; Tang et al., 1994).

However, despite ample evidence of dysfunctional social communication in schizophrenia and its association with negative outcomes for patients' quality of life, speech language therapy is not yet a systematical part of a comprehensive intervention (Joyal et al., 2016). Nevertheless, first studies show positive effects of speech related tasks in rehabilitation approaches on speech production in patients with symptoms such as alogia or delusional speech (Allen et al., 1978; Clegg et al., 2007; Foxx et al., 1988; Kramer et al., 2001). Further studies have found positive effects for discourse production (score of intelligibility, appropriateness and elaboration of responses) (Baker, 1971), verbal fluency (Ojeda et al., 2012; Vianin et al., 2014; Wykes, 1998) and naming (Kondel et al., 2006). Furthermore, results from first exploratory studies on specific speech language therapy in patients with SSD actually give reason to assume beneficial effects of speech language interventions (Bosco et al., 2016; Santos et al., 2021). On the other hand, some studies could neither find a benefit of speech language therapy for sentence understanding, repetition and naming (Man et al., 2012), nor for verbal fluency (Blairy et al., 2008; Ojeda et al., 2012) and pragmatic non-verbal skills (Kawakubo et al., 2007).

Nonverbal trainings are rare up to nonexistent in the field of psychiatry. First studies show beneficial effects of stimulation techniques on gesture integration and production without training (Schülke & Straube, 2019; Walther, Kunz, et al., 2020).

A particular challenge with this patient group is furthermore the feasibility of specific intervention programs: previous intervention studies report a dropout rate of approximately 15% (Gordon et al., 2018).

Considering the potential impact of communication on social life and therefore life quality in general (Bambini et al., 2016), it is particularly important to focus on the development of a well feasible specific speech language training intervention (Heim, 2020; Joyal et al., 2016; Riedl et al., 2020; Rossetti et al., 2018).

1.3 Aim and Research Questions

Here we describe the development and implementation of a novel multimodal speech-gesture training for patients with schizophrenia spectrum disorder. For the first time, we conducted such a training with SSD patients. Hence, we focused on the feasibility of the training program and on initial evidence of efficiency in changing dysfunctional neural processing and transfer effects on quality of life.

In order to be able to make statements about the feasibility of the training program, we analyzed the dropout rate of SSD patients during the MSG training. In order to get more detailed impressions of how the patients experienced the training program, we exploratory investigated satisfaction with training through specifically outlined questionnaires.

As communication is the basis for social interaction and hence for quality of life, we compared quality of life scores before and after training with a control group through a standardized psychological questionnaire (Satisfaction With Life Scale: SWLS).

To examine the relationships of training, quality of life, and neural processing, we furthermore measured neural patterns of processing abstract speech-gesture episodes through fMRI using specific experiments with iconic and metaphoric gesture materials (Straube, Green, Bromberger, et al., 2011a) and correlated these neural patterns with changes in quality of life.

To get a more comprehensive understanding of the potential training effects, we also exploratory examined suitable behavioral outcomes from the training tasks and subjective impressions of the patients and their relatives using specifically outlined questionnaires about nonverbal communication and social life.

1.4 Hypotheses

(1) The MSG training is well feasible with patients with SSD, which is expressed by a low dropout rate of four dropouts maximum (lower than 15% dropout rate, considering dropout rates in comparable studies (Gordon et al., 2018)).

(2) Patients with schizophrenia spectrum disorder (SSD group) show significantly reduced quality of life (measured by SWLS score) in comparison to the control group before training (ses-pre).

(3) The quality of life (measured by SWLS score) increases significantly over the MSG training (ses-pre versus ses-post) in the SSD group.

(4) Neural activation in abstract multimodal conditions changes significantly over the MSG training (ses-pre versus ses-post) in the SSD group. In particular, we expect changes in bilateral temporal regions and left inferior frontal gyrus.

(5) Hence, neural activation patterns are more similar between SSD patients and control group after (ses-post) than before (ses-pre) the MSG training.

(6) The neural activation changes in multimodal conditions with abstract content are specific to the MSG training (for the waiting-first group): We expect no significant changes during a no treatment period (ses-pre versus ses-bl), but significant pre-post MSG training changes (ses-bl versus ses-post).

(7) Changes in the SSD group in quality of life (measured by SWLS score) correlate with the pre-post training changes in neural activation in multimodal abstract conditions.

2 Methods

2.1 Participants

2.1.1 Inclusion of Participants

The study was carried out in accordance with recommendations of the local ethics committee (Philipps-University Marburg, Department of Medicine, Deanery/Ethics Committee, Reference: R1, Study 01/17) and has been registered at the German Clinical Trials Register, DRKS (DRKS00015118, DRKS.de). All subjects gave written informed consent in accordance with the Declaration of Helsinki and were compensated for participation. The protocol was approved by the local ethics committee. All information collected is kept confidential, stored securely and archived in accordance with the research governance policy of the university. Participant anonymity is retained by allocating a unique identification number for the trial and any identifiable information stored separately from this.

In a telephone screening, volunteers were called to make sure that both their status of physical health and capability to participate in a fMRI study met our conditions. Inclusion and exclusion criteria were part of this screening as well as a questionnaire about the progress, the individual core symptoms of schizophrenia and medication.

Subjects *were eligible* for study entry if they met the following criteria:

- Aged between 18 and 62 years[1]
- Capability to give informed consent

Additional criteria, patients only:

- Diagnosis of schizophrenia spectrum disorder according to ICD-10 criteria
- Relatively stable symptoms (no acute psychosis)

---

[1] Here, we deviated by 2 years from the original criteria described in the study protocol (Riedl et al., 2020) to ensure accurate matching by age in the two subject groups (SSD group and control group).

To evaluate the stadium and type of symptoms of the patients as well as for safety reasons, additionally a pre-scan interview with the patients including questions concerning the illness and medication as well as standardized psychological questionnaires for schizophrenia (SAPS (Andreasen, 1984)/SANS (Andreasen, 1983)) were conducted.

Subjects *were not eligible* if any of the following criteria were present:

- No capability to give informed consent
- Risk of suicide necessitating hospitalization
- Physical illnesses that interfere with the planned measurements
- Medical contraindication against fMRI measurements
- Pregnancy
- Contraception per intrauterine contraceptive device

Additional criterion for exclusion, control subjects only:

- Diagnosis of mental diseases

If volunteers met the criteria, they met with a member of the research team to go through the study information. Inclusion and exclusion criteria were confirmed along with the assessment of positive and negative symptoms of schizophrenia as well as gesture performance (for patients), social activity and quality of life (all participants).

All participants were free of visual and auditory deficits, additional neurological and medical impairments as well as any cerebral abnormality, as assessed by a T1-weighted MRI. All participants reported German to be their primary language. For an overview of the demographics and the equivalents of antipsychotic medication in the patients, see *table 1*.

Participants were informed that they were free to withdraw at any time without giving reasons and without prejudicing any further treatment.

2.1.2 SSD Group

The original sample encompassed 30 patients diagnosed with schizophrenia spectrum disorder (SSD). One of the patients dropped out for private reasons, so our final sample of patients comprised n=29. From these, 20 patients were allocated to the wait-first group and 9 to the training-first group. The patients were diagnosed by clinicians according to the International Classification of Diseases (World Health Organization, 1992) with schizophrenia (F20.0, n = 13; F20.1, n = 2; F20.3, n = 1; F20.6, n = 1), schizoaffective disorder (F25, n = 9), acute and transient psychotic disorders (F23.0, n = 1; F23.1, n = 1; F23.2, n = 2), substance-induced psychosis (F1x.50, n = 3) or schizoid personality disorder (F60.1, n = 1).[2] Twenty-two of the patients were treated with atypical antipsychotic medication; six patients received (additional) antidepressants or other psychiatric medical treatment. Five patients were not medicated. Nineteen patients were under psychological treatment throughout the time of the study. Before participation and in order to characterize the severity of schizophrenia symptoms, the German adaptation of the Scale for the Assessment of Negative Symptoms (SANS) (Andreasen, 1983) and the Scale for the Assessment of Positive Symptoms (SAPS) (Andreasen, 1984) were assessed. On average, the patients had a total SANS score of 22.52 (SD = 14.8) and a total SAPS score of 24.54 (SD = 19.62). Eight of the patients were women, mean age was 35.48, ranging from 22 to 62 years. In average, they held a General Certificate of Secondary Education, which is equivalent to 10 years of school in Germany. Since the training was cognitively demanding and furthermore focused on the stable symptoms, only patients who were deemed clinically stable were invited.

2.1.3 Control Group

In the control group, initially 27 participants without any psychiatric diagnosis were included. 10 dropped out: 2 aborted for private reasons, 4 had to be excluded for safety reasons and/or incidental findings of neural structure during the first assessment and 4 had to be excluded or quit the experiment, because the waiting time

---

[2] The divergent total number of diagnoses from the number of participating patients results from multiple diagnoses in individual patients.

between the measurements was too long[3] due to lockdown in the Covid-19 pandemic. The final sample in the control group comprised n=17, so we did not split them into two different treatment groups as we did with the patients (see also section *2.2.2 Modifications of the Study Design*). Four participants from the final sample were women, mean age was 36.53, ranging from 22 to 62 years. Just like the SSD group, in average they held a General Certificate of Secondary Education.

2.1.4 Matching of Groups

The groups were matched for sex, age and education, where they did not significantly differ (see *table 1*).

	Patients with SSD		**Controls**		**Difference (*p*)**
Age (years)	35.48	±11.07	36.53	±11.77	0.774
Sex male/female	22/7		13/4		0.964
Education (codes)*	2.41	±0.67	2.35	±0.68	0.776
SAPS (total)	24.54	±19.26	-	-	-
SANS (total)	22.52	±14.8	-	-	-
TULIA	11.36	±0.93	-	-	-
CPZ Equivalent	275.03	±330.27	-	-	-

Table 1: Demographics, medication and neuropsychological measures. Values are presented as mean ± standard deviation.

CPZ: chlorpromazine; SAPS/SANS: scale for the Assessment of Positive/Negative Symptoms; SSD: schizophrenia spectrum disorder; TULIA: Apraxia Screen of TULIA (AST).

* Education codes: 1: Certificate of secondary education; 2: General certificate of secondary education; 3: General qualification for university entrance

2.2 Study Design

2.2.1 Experimental Procedure

This was a single-center randomized controlled trial of intensive single speech-gesture training versus wait-list control being conducted at Philipps-University Marburg, Department of Psychiatry and Psychotherapy. In this institution, patients diagnosed with SSD (see *chapter 2.1.2 SSD Group*) and healthy controls (see *chap-*

---

[3] Too long in terms of being not longer comparable to the intervals between the measurements/assessments in the SSD group.

ter 2.1.3 Control Group) were recruited. Outcomes were measured through pre-post-fMRI and standardized psychological questionnaires comparing the two subject groups.

To ensure quality as well as transparency of our methods (as it is postulated specifically in the fields of psychiatry (Bell, 2017) and neuroimaging (Gau et al., 2021) and promoted by the European Commission (The Lisbon Council, 2019)), we described our methods, including the study's design and the MSG training program, in detail in advance the completion of data acquisition (Riedl et al., 2020).

2.2.1 Intervention Groups

SSD patients were randomly allocated (computerized random numbers) to one of two groups - waiting-first or training-first. In the waiting-first group, a first baseline fMRI measurement (ses-pre) was conducted, followed by a waiting time (treatment as usual: TAU) and a second baseline fMRI measurement (ses-bl). Then, patients performed the MSG training program and concluded the study with a post fMRI measurement (ses-post). In the training-first group, the baseline fMRI measurement was directly followed by the MSG training program. Then followed the second fMRI measurement, a waiting period (TAU) and then the final fMRI measurement.

Thus, patients in our study participated in both – the MSG training program and the TAU period. A first benefit from this approach is the comparability of both treatment groups (perfect matching of the subjects from the MSG training and the TAU). Secondly, we wanted to offer our training program to all patients, which is possible with our design where all patients could benefit from the program. Due to a lack of former similar studies' effect sizes for power calculation, the comparison of three measurements provides further evidence for possible training specific effects through comparing intra-individual repetition and training effects.

To assess normal functioning on behavioral and neural level, healthy subjects were additionally thought to be involved as control group in our training procedure, but they received no training and were therefore measured only twice (ses-pre and ses-post with TAU in between).

With this design, altogether, we were able to compare patients and control subjects pre-to-post with an equal number of repetitions as well as data from MSG training program of all patients. For further details, see Riedl et al. (2020).

2.2.2 Modifications of the Study Design

Due to recruitment delays and dropouts in the control group during the Covid-19 pandemic (as clarified in *chapter 2.1.2 SSD Group*), control subjects were not divided into two intervention groups as originally planned (Riedl et al., 2020) and thus did not conduct the training program. Therefore, we were not able to compare training success between the subject groups (SSD group versus control group). Nevertheless, the modified design allowed us to make inferences about subject group differences before the MSG training program as well as the progress in training task performance in the SSD group.

2.3 MSG: The Multimodal Speech-Gesture Training Program

2.3.1 Development of the MSG Training Intervention

Since specific intervention approaches focusing on multimodal communication skills in patients with SSD are lacking so far, we developed a completely new multimodal speech-gesture (MSG) training program. Every session had a similar sequence of exercises and was developed following best practice in therapy intervention.

According to the classic Wernicke-Lichtheim-Geschwind model, speech production is secondary to perception (Geschwind, 1970), see *chapter 1.2.2 Speech Processing*. Hence, in speech language therapy, for example in patients with aphasia after stroke, speech perception is treated first or at least parallel to the production of speech. Due to the fact that in schizophrenia understanding language is frequently affected (see *chapter 1.2.4 Speech Processing in SSD*), a training of speech perception and interpretation of meaning is appropriate. Results from first exploratory studies actually give reason to assume beneficial effects of speech language interventions also for speech perception in patients with schizophrenia (Bosco et al., 2016; Santos et al., 2021), see *chapter 1.2.8 Current Treatments for Patients With SSD*.

Because some of the core communication problems of patients with schizophrenia concern the integration into the sentence context (Kostova et al., 2003), it is important to not only take into account isolated words or phrases. The integration of words into context is strongly associated with working memory capacities, as Kintsch and van Dijk claim in their model of text comprehension and production (Kintsch & van Dijk, 1978). Hence, working memory and language processing seem to be thoroughly connected. This is also true for working memory and the integration of gesture (Rudner, 2018; Straube et al., 2009; Straube, Green, Chatterjee, et al., 2011), see *chapter 1.2.3 Gesture Processing*, which seems to be dysfunctional in SSD, see *chapter 1.2.5 Gesture Processing in SSD*. There are approaches of working memory training for patients with schizophrenia already (Bor et al., 2011; Haut et al., 2010; Penadés et al., 2013; Subramaniam et al., 2014;

Wykes et al., 2002). For these reasons, specifically outlined working memory tasks should be taken into account as an important part of a speech language training program.

Social communication is not only based on speech itself but also on the integration of nonverbal information such as gesture (Suffel et al., 2020), see *chapter 1.2.3 Gesture Processing*. Thus, a multimodal training program including gesture training tasks might help to develop, reactivate or promote communication resources in patients with schizophrenia. Nonverbal trainings are rare up to nonexistent in the field of psychiatry, see *chapter 1.2.8 Current Treatments for Patients With SSD*. One study showed positive effects of transcranial direct current stimulation (tDCS) on gesture integration in a group of patients with SSD without training (Schülke & Straube, 2019). Another study could show a benefit from single session transcranial magnetic stimulation (TMS) also for gesture production in patients with schizophrenia (Walther, Kunz, et al., 2020). This effect could be potentially increased and prolonged in combination with an adequate gesture training program. Despite the difficulties in processing multimodal input, patients with schizophrenia seem to benefit at least partially from gestures accompanying speech compared to processing unimodal input (Cuevas et al., 2021; He et al., 2021). Thus, gestures that visualize concrete concepts of the source domain could help patients with SSD to understand the abstract meaning in a cognitive metaphor and integrate it into the sentence context.

To solve the open questions regarding the efficiency of a specific training in patients with schizophrenia, considering perceptive speech processing, working memory functions and the integration of nonverbal communication into abstract sentence content, we developed a specific multimodal speech-gesture training (MSG training) program.

The influence of speech-gesture training has scarcely been examined so far, but given the previous findings, it seems to be promising to implement such a training procedure for patients with schizophrenia (Heim, 2020; Heim et al., 2019; Joyal et al., 2016; Riedl et al., 2020; White et al., 2016; Wüthrich et al., 2020).

2.3.2 Implementation of the MSG Training Intervention

Examiners

Speech language therapists and medical students with their focus on psychiatric disorders were involved as examiners. They were trained in detail before they conducted parts of the study (e.g., the MSG training) with patients. Examiners did not alternate between patients to ensure steadiness. During MSG training sessions, at least one examiner was present. Examiners were unaware about our specific neural and behavioral hypotheses regarding the MSG training effects on the different measures.

Setting

The training took place in an individual setting, taking into account that exercises including gesture performance were possibly hard to execute when other participants would have been present. We offered eight sessions (60 minutes each) of high frequent training, considering the dose-effect relationship that suggests eight training sessions to be the minimum for patients to benefit (Howard et al., 1986), as well as recommendations of high frequent interventions in speech language therapy for SSD patients (Joyal et al., 2016).

MSG Training Procedure

Every session began with a short personal intro to establish a respectful relationship between patient and examiner, including a discussion about the homework prepared for the current session. Thereafter, patients executed four exercises which we developed and selected to increase complexity: First two perceptual tasks (one relatedness (Choudhury et al., 2021; Schülke & Straube, 2019) and one working memory (Rudner, 2018; Straube et al., 2009; Straube, Green, Chatterjee, et al., 2011) task), then a productive (imitation/mime) task. As in the fMRI measurements, video material was used for these three tasks where an actor produces sentences with a concrete or abstract meaning and accompanies these sentences with gesture. In the relatedness task, videos with related and videos with unrelated speech-gesture pairs were presented for the patients to rate. In the working memory task, patients were asked to decide whether a video was new or presented for the second time (n-back task). The videos in the working memory task contained unimodal (speech or gesture only) and multimodal (speech accompanied by gesture) input. The videos differed from the stimuli used in the fMRI measurements.

As the training increased in difficulty, we developed and selected as the last exercise a free productive speech gesture fluency (SG fluency) task. Similar to verbal fluency tasks (Rosenkranz et al., 2019; Wende et al., 2012), the patients were instructed to generate as many words with accompanying suitable gesture as possible for each of the three semantic fields per session, resulting in a total of 24 categories. Time was limited to one minute per semantic field. The stimulus material in the SG fluency task was selected with a primary focus on eligibility for therapeutic intervention and controlled[4] regarding psycholinguistic properties that are known to have an effect on language processing, including typicality (Glauer et al., 2007) and the effect of word frequency[5] (Brysbaert et al., 2011).

At the end of each training session, patients were provided with some interesting background information about gesture and how it is related to language and communication (“what it is good for”) that was supposed to motivate them to attend to and talk about gestures in everyday life situations. Handouts summarizing the content of this information and an explanation of the new homework were given to patients at the end of each session to encourage transfer of the training effects to daily life routine.

Protocols were written during the training and productive tasks were video recorded for later independent analysis. After the last training sessions, patients and their relatives were asked for feedback regarding the MSG training and its impact on the patients' daily life using specifically outlined questionnaires.

---

[4] A one-way analysis of variance (ANOVA) did not reveal significant differences with respect to word frequency ($F(7.16) = 0.80$; $p = 0.598$) and typicality between the categories of each session ($F(7.16) = 3.52$; $p = 0.092$).

[5] To control for word frequency, dlexDB frequencies (DlexDB, n.d.) were used, which seem to be more reliable for German than the commonly used CELEX frequencies (Brysbaert et al., 2011).

2.3.3 TAU: Treatment as Usual

In addition to the fMRI before and after the MSG training program, we conducted a further fMRI measurement (ses-bl) after a period of waiting (TAU) in the SSD group. The duration of MSG training and TAU period were the same within a patient. This allowed us to compare the outcomes from the training to a period of no training. This approach is similar to a training-TAU-design with two groups (one group undergoing the training and a control group undergoing the TAU, which means no training in the present study).

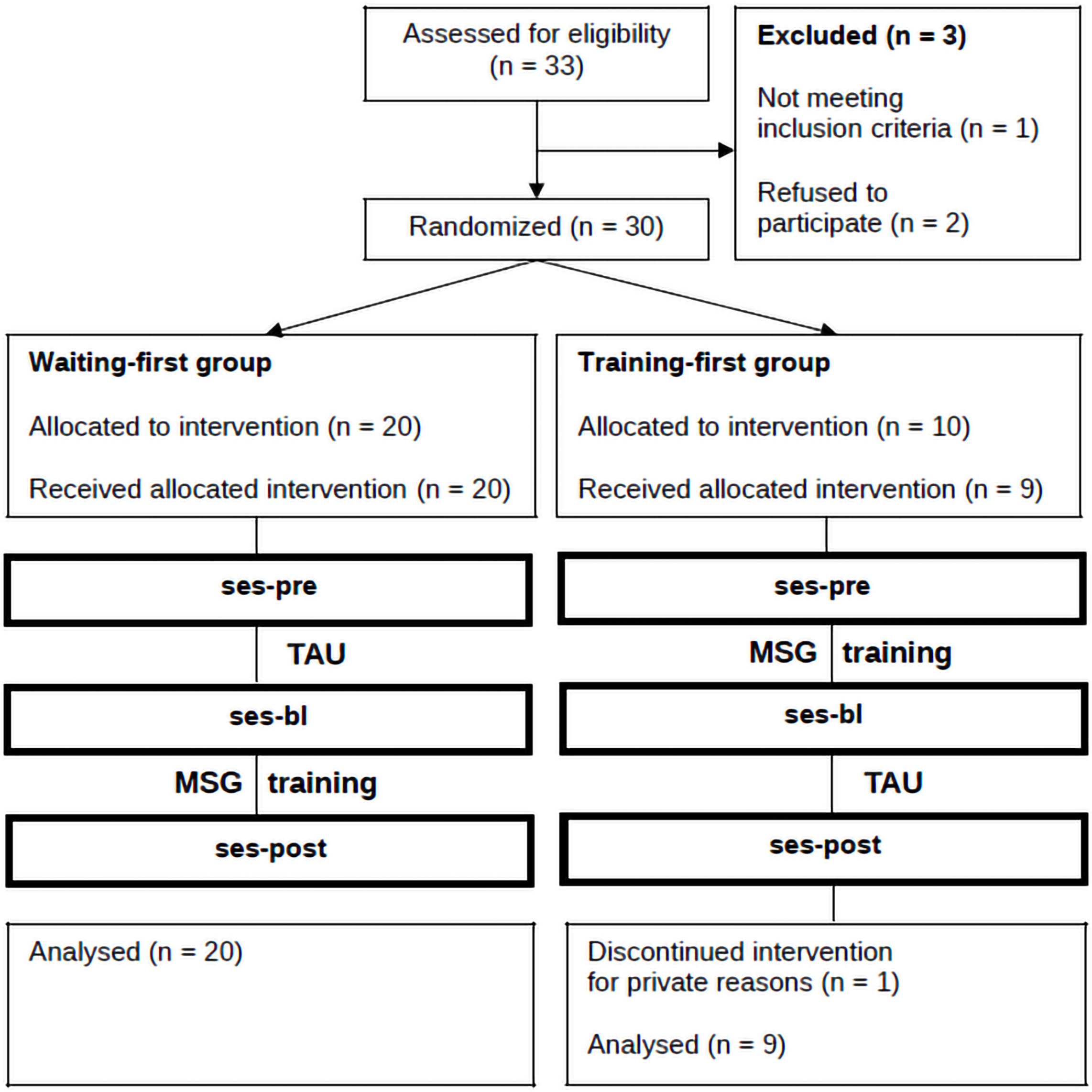


Figure 5: Simplified design of the study after CONSORT guidelines.

Abbreviated illustration of the experimental design and treatment groups in the patients after CONSORT guidelines (Boutron et al., 2017).

TAU, treatment as usual: waiting period without training; MSG training, multimodal training program (patients only).

2.4 Measurements

2.4.1 Feasibility Assessment

In order to be able to make a statement about the feasibility of the training, the number of dropouts during the training was collected.

Furthermore, in order to obtain detailed information that could be useful for evaluating the feasibility of the MSG training program, we exploratory analyzed two specifically outlined post training questionnaires. One of the questionnaires was intended to represent the impression about the training of the patients, the other one the impression of a close relative.

Patients' Questionnaire about the MSG Training

At the end of the very last training session, the patients were asked for their feedback regarding the MSG training and its impact on their everyday life using a standardized questionnaire.

Statements about the MSG training to be evaluated:

1. I got along well with the training management.
2. I enjoyed the tasks in the training (motivation).
3. I found the training useful.
4. It was easy for me to work on the tasks in the training.
5. I made a lot of effort to work on the tasks in the training.
6. I've put a lot of effort into working through the application exercises
7. I also benefited from the training in everyday life.
8. I understood the tasks in the training.

The questionnaire's scale ranged from 1 (very bad) to 10 (very good).

Relatives' Questionnaire about the MSG Training

After the training, the patients' relatives were contacted to to ask them what impact the MSG training had on the patient's communication and social functioning from their perspective as a close person and therewith a frequent communication partner.

Questions:

Compared to the state before training:

1. How often does she/he have social contacts now?

Can the study participant do the following, compared to the state before training:

2. Overall, adequately understand what the interlocutor is saying?

3. Make her-/hisself understandable only by gestures (e.g., in the case of loud ambient noises)?

4. Draw attention to her-/himself with gestures?

5. Make it clear in the conversation using gestures whether she/he is listening / bored / wants to speak for her-/himself (turn taking)?

6. Use gestures while speaking?

7. Use gestures that match her/his language/statement?

8. How often does the participant use gestures?

9. Communicate in an understandable way (e.g., express her-/himself clearly)?

10. Express emotions/emotional states using gestures/facial expressions?

11. Interpreting the other party's gestures correctly (appropriate reaction)?

12. Use gestures appropriate to the situation?

The questionnaire's scale ranged from -5 (much worse than before) to 5 (much better than before).

2.4.2 Quality of Life Assessment

To assess the effect of the MSG training on social functioning and quality of life, SSD patients as well as control subjects were asked to complete a psychological questionnaire about quality of life (German version of SWLS – Satisfaction with Life Scale (Glaesmer et al., 2011): satisfying validity and reliability with cronbach alpha = 0.87 (Pavot & Diener, 2009)) before and after the training program. Outcomes were checked for implausible values.

2.4.3 FMRI Measurements

Detailed manuals were developed and provided to the examiners who conducted the fMRI measurements. Furthermore, examiners were trained in detail before they conducted parts of the study with participants. During the fMRI measurements, at least two examiners were present. Protocols were written for fMRI measurements.

Before we started with the measurement, the participants were informed about the setting of the study. A member of the research team went through the study information and answered any questions.

Stimuli

In order to gain an advanced understanding in the engine of integration mechanisms and their neural correlates in patients with schizophrenia, participants underwent fMRI measurements watching videos on a screen[6] before and after a waiting period (TAU) as well as after the actual MSG training. These videos differed from the ones presented during the MSG training procedure. The videos had been recorded with an actor expressing concrete (con) or abstract (abs) sentences (for an explanation of abstractness see section *1.2.2 Speech Processing*) accompanied by gesture (SG), see *figure 6*. Additionally, abstract and concrete unimodal control conditions (speech only: S, or gesture only: G) were presented. The videos were standardized, extensively evaluated and had been successfully applied in a large number of fMRI (Green et al., 2009; Kircher et al., 2009; Straube et al., 2009; Straube, Green, Bromberger, et al., 2011a), EEG (He et al., 2018), tDCS (Schülke

---

[6] In the MRI scanner, an additional mirror had to be used for presentation.

& Straube, 2019) and patient studies including patients with schizophrenia (Straube, Green, et al., 2013; Straube et al., 2014). The video sequences had a mean duration of five seconds.

The experimental paradigm was implemented in the software *Presentation*® (Version 18.3, (Presentation (Software) Documentation, n.d.) Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com). The experiment comprised 30 video sequences per condition, 180 video sequences in total, presented in a pseudo-randomized order in three blocks of data acquisition (functional runs) with a duration of ~9 minutes each (27 minutes in total). Each video sequence trial was followed by a fixation cross presented in the center of a black screen with an average jitter of 8215 ms (6682 – 9818 ms) between the onsets of two succeeding videos. The participants were pseudo-randomly assigned to one of eight different counterbalanced versions of stimuli presentation to avoid sequence effects. No sentence content was presented more than once in the same or another condition to the participants in the same measurement session.

Figure 6:

Example of stimuli.

Illustration of the stimulus conditions:

Above: sentences with abstract content accompanied with gesture (absSG);

Below: sentences with concrete content accompanied with gesture (conSG).

Task

The videos were presented to the participants who had been instructed to watch the videos and tap with their left forefinger on the buttons of a response box that was fixated at their left leg to confirm that they attentively watched (passive perception task). This allowed us to investigate implicit speech and gesture processing in concrete and abstract conditions (due to 'concretism' in schizophrenia) and different modalities as well as multimodal integration (Straube, Green, Bromberger, et al., 2011a; Straube, Green, et al., 2013).

FMRI Paradigm

Imaging data were collected with a 3 T whole body MRI system (SIEMENS MAGNETOM TrioTim syngo MR B17) equipped with a standard head coil. Structural image acquisition consisted of 176 T1 weighted sagittal slices (slice thickness = 1.0 mm; FoV = 256 mm; TR = 1900 ms; TE = 2.26 s). To measure BOLD changes in brain activity during acquisition, T2* weighted gradient echo planar imaging (EPI) with 34 slices covering the whole brain were used (voxel size = 3 x 3 x 4 mm; descending slice acquisition; slice thickness = 4.0 mm; TR = 1650 ms; TE = 25 ms; flip angle = 70°; FoV = 192 mm; GRAPPA = 2). Slices were adjusted after the anterior commissure posterior commissure (AC- PC) line. In the three measurements (se-pre, ses-bl and ses-post), 936 functional images were acquired during acquisition phase. A gradient echo field map sequence was measured before the functional runs to get information for unwarping B_0 distortions.

2.4.4 Exploratory Assessment of Performance During MSG Training

To get a more comprehensive understanding of the potential training effects, we also exploratory examined behavioral outcomes from the training tasks. The speech-gesture (SG) fluency task was a fixed component in every session of the MSG training and therefore suitable to explore training related changes across training sessions. The performance in this task was evaluated via the videos recorded during the MSG training. Two experienced raters transcribed and rated the individual performances in the SG fluency task independently based on the video recordings. Reliability of the ratings was calculated between the two raters using the intraclass correlation coefficient (McGraw & Wong, 1996) implemented in the *R* (R Core Team, 2013) package *irr* (Gamer et al., 2012). Based on a one-way random effects model, there was an excellent agreement (Koo & Li, 2016) between the two raters (kappa = 0.99; $p < 0.001$). In general, produced speech-gesture pairs were classified as correct, if the word was a member of the semantic field and accompanied by a matching adequate gesture. The total number of correctly produced words with accompanying gesture was calculated for each participant and each session as a sum of the productions in the three semantic fields.

2.5 Data Analysis

2.5.1 Feasibility Analysis

In order to be able to make a statement about the feasibility of the training (*hypothesis (1)*), the number of dropouts during the training was analyzed and compared with dropout rates of previous comparable studies (Gordon et al., 2018).

2.5.2 Quality of Life Score Analysis

For difference between SSD group and control group in SWLS score (*hypothesis (2)*), a Welch t-test and a student's t-test for session comparisons within SSD group (*hypothesis (3)*) were performed using the *ggstatsplot package* (Patil, 2021) for *R* (R Core Team, 2013).

2.5.3 FMRI Data Analysis

Defacing

As a first preprocessing step, the structural images were defaced and personal data were deleted from NIfTI headers using *BIDSonym* (Herholz, n.d.; Herholz et al., 2021) in order to ensure anonymity of participants during further analyses.

Quality Control

For quality control, the BIDS[7] app *MRIQC* was used based on different parameters like co-registration, motion and temporal signal-to-noise (Esteban et al., 2017). No data had to be excluded.

Modification of Preprocessing Strategy

Diverging from the planned preprocessing pipeline described in (Riedl et al., 2020), for the reasons of newly developed preprocessing pipelines, we decided to work with standardized BIDS apps in form of *singularity* containers[8] to ensure quality of data, transparency of the preprocessing process and possibility of replication.

---

[7] BIDS: Brain Imaging Data Structure (K. J. Gorgolewski et al., 2017).

Results included in this manuscript come from preprocessing performed using *fMRIPrep* 20.2.0rc0 (Esteban et al., 2018; Esteban, Markiewicz, Goncalves, et al., 2020) (RRID:SCR_016216), which is based on *Nipype* 1.5.1 (Esteban, Markiewicz, Burns, et al., 2020; K. Gorgolewski et al., 2011) (RRID:SCR_002502).

Anatomical Data Preprocessing

The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with *N4BiasFieldCorrection* (Tustison et al., 2010), distributed with ANTs 2.2.0 (Avants et al., 2008) (RRID:SCR_004757), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a *Nipype* implementation of the *antsBrainExtraction.sh* workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using *fast* (Zhang et al., 2001) (FSL 5.0.9, RRID:SCR_002823). Brain surfaces were reconstructed using *recon-all* (Dale et al., 1999) (FreeSurfer 6.0.1, RRID:SCR_001847), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (Klein et al., 2017) (RRID:SCR_002438). Volume-based spatial normalization to two standard spaces (MNI152NLin6Asym, MNI152NLin2009cAsym) was performed through nonlinear registration with *antsRegistration* (ANTs 2.2.0), using brain-extracted versions of both T1w reference and the T1w template. The following templates were selected for spatial normalization: *FSL's MNI ICBM 152 non-linear 6th Generation Asymmetric Average Brain Stereotaxic Registration Model* [(Evans et al., 2012), RRID:SCR_002823; TemplateFlow ID: MNI152NLin6Asym], *ICBM 152 Nonlinear Asymmetrical template version 2009c* [(Fonov et al., 2009), RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym].

---

[8] Singularity: A computer program that performs operating-system-level virtualization also known as containerization. One of the main uses of *Singularity* is to bring containers and reproducibility to scientific computing.

Functional Data Preprocessing

The experiment was divided in 3 blocks of data acquisition (functional runs) per measurement session. For each of these functional runs, the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Susceptibility distortion correction (SDC) was omitted. The BOLD reference was then co-registered to the T1w reference using *bbregister* (FreeSurfer) which implements boundary-based registration (Greve & Fischl, 2009). Co-registration was configured with six degrees of freedom. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using *mcflirt* (FSL 5.0.9) (Jenkinson et al., 2002). The functional runs were slice-time corrected using *3dTshift* from AFNI 20160207 (Cox & Hyde, 1997) (RRID:SCR_005927). The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying the transforms to correct for head-motion. These resampled BOLD time-series will be referred to as *preprocessed BOLD in original space*, or just *preprocessed BOLD*. The BOLD time-series were resampled into several standard spaces, correspondingly generating the following *spatially-normalized, preprocessed functional runs*: MNI152NLin6Asym, MNI152NLin2009cAsym. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, (Power et al., 2014)) and Jenkinson (relative root mean square displacement between affines, (Jenkinson et al., 2002)). FD and DVARS were calculated for each functional run, both using their implementations in *Nipype* (following the definitions by (Power et al., 2014)). The three global signals were extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (*CompCor*, (Behzadi et al., 2007)). Principal components were estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components were then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) were generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by 2 pixels on BOLD space, the aCompCor masks are subtracted a mask of pixels that likely contain a volume fraction of GM. This mask is obtained by dilating a GM mask extracted from the FreeSurfer's *aseg* segmentation, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks were resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components were also calculated separately within the WM and CSF masks. For each CompCor decomposition, the *k* components with the largest singular values were retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components were dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al., 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with *a single interpolation step* by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using *antsApplyTransforms* (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos, 1964). Non-gridded (surface) resamplings were performed using *mri_vol2surf* (FreeSurfer).

In order to remain agnostic to any possible subsequent analysis, *fMRIPrep* does not perform spatial smoothing. For this reason, functional images were smoothed using a Gaussian kernel of 6mm full width at the half maximum in all directions to adjust for inter-subject brain anatomical variations using *Statistical Parametric Mapping* (*SPM12* software package (*SPM12*, 2018); Wellcome Trust Centre Human Neuroimaging, London; http://www.fil.ion.ucl.ac.uk, RRID: SCR_007037) implemented in *MATLAB* R2017a (version 9.2.0 (*MATLAB*, 2017); MathWorks). For this purpose, a standardized *MATLAB* script was created that transfers *fMRIPrep* output to files that are ready to be analyzed by *SPM* (*fmriprep2spm*, https://github.com/LydiaRiedl/fmriprep2spm (Riedl, 2021/2021)).

Statistical Analysis

FMRI data were analyzed using standard routines for first and second level analyses in *SPM* (*SPM12* software package (*SPM12*, 2018); Wellcome Trust Centre Human Neuroimaging, London; http://www.fil.ion.ucl.ac.uk, RRID: SCR_007037) implemented in *MATLAB* R2017a (version 9.2.0 (*MATLAB*, 2017); MathWorks).

Single Subject Level Analysis

On the first level, single subjects' voxel-wise BOLD activity was modeled by a General Linear Model (GLM) (Friston et al., 1995; Worsley & Friston, 1995). In the context of experiments in which a particular comparison of conditions at multiple points in time are examined, effects are tested using an analysis of covariance (ANCOVA). The use of ANCOVA in the context of the GLM allows examining effects of interest while removing the effects of other possibly confounding variables such as head motion parameters (Poldrack, 2000). The six realignment parameters extracted by *fMRIPrep* (head motion) were modeled as regressors of no interest. The hemodynamic response was modeled in terms of the canonical hemodynamic response function (HRF) with parameter estimate images calculated for each participant and condition in the three functional runs (see section *2.4.3 FMRI Measurements*) separately. The onset of each event was defined as the integration point (the time when the stroke of the gesture coincides with the keyword of the sentence (Green et al., 2009). All events were modeled with a duration of 1 s and assigned to one of the conditions.

Group Level Analysis

The single subject level parameter estimates were deployed for a full factorial analysis. The main effect of condition in all three functional runs were defined as contrasts of interest for the second level analysis, resulting in baseline contrasts for the two conditions of interest (absSG, conSG) and the four unimodal control conditions (absS, absG, conS, conG). In the statistical model, group (SSD group, control group) was considered as a between-subject factor, session (ses-pre, ses-post) as within-subject factor, content (abs, con) as within-subjects factor and modality (SG, S, G) also as within-subjects factor, resulting in a 2 x 2 x 2 x 3 design.

A Monte-Carlo-Simulation was performed (acquisition matrix: $x = 64$, $y = 64$; slices: 34; DIM: $xy = 3$ mm, $z = 4$ mm; FWHM = 10.3 mm; DIM resampled = 2 mm; no mask; iterations: 1000) to calculate the minimum voxel contiguity threshold needed to correct for multiple comparisons at $p < 0.05$, assuming an individual voxel type I error of $p < 0.01$ (Slotnick, 2017; Slotnick et al., 2003). A cluster extent threshold of 221 contiguous resampled voxels at $p < 0.05$ (whole-brain analysis) was used for all contrasts of interest.

Voxel coordinates reported are located in the Montreal Neurological Institute (MNI) brain space. For anatomical location, functional data were referenced to the Automated Anatomical Labeling toolbox (*AAL*) in *SPM12* (Rolls et al., 2015; Tzourio-Mazoyer et al., 2002).

For further statistical analyses of neural and behavioral data, *R Studi*o (RStudio, 2021) was utilized.

Contrasts of Interest

To clarify *neural changes for the processing of abstract multimodal videos in the SSD group* as predicted in *hypothesis (4)*, we calculated a session X abstractness interaction for the processing of multimodal input:

(1) **Session comparison of abstract multimodal processing**

(ses-pre(absSG>conSG)) < (ses-post(absSG>conSG))

To test *hypothesis (5)*, we tested *group similarities for the processing of abstract multimodal videos* by applying conjunction 0 (minimum t-statistics) to explore if the predicted neural changes after the MSG training led to more commonalities between SSD patients and control group:

(2) **Group comparison of abstract multimodal processing: ses-pre effect**

(control group ses-pre (absSG > conSG)) ∩ (SSD group ses-post (absSG > conSG))

(3) **Group comparison of abstract multimodal processing: ses-post effect**

(control group ses-pre (absSG > conSG)) ∩ (SSD group ses-post (absSG > conSG))

To investigate *specific training effects in the processing of abstract multimodal videos (hypothesis (6))*, we furthermore calculated a three-sessions-interaction X abstractness (F-test), comparing all three sessions (ses-pre, ses-bl, ses-post) in the SSD wait-first group (n=20):

(4) **Intervention comparison of abstract multimodal processing**

2.5.4 Exploratory Analysis of Performance During MSG Training

For the exploratory analysis of the performance in the SG fluency task across all eight training sessions, a Fisher's one-way ANOVA was calculated using the *ggstatsplot package* (Patil, 2021) for *R* (R Core Team, 2013).

2.5.5 Correlations of Neural Activation, Performance and Quality of Life

In order to examine the relationship of changes in quality of life (SWLS score changes) and neural activation changes (*hypothesis (7)*), eigenvariates of the multimodal condition regressors (absSG and conSG) were extracted from the significant clusters in the *patients' interaction session X abstractness* (*contrast (3)*) and the *three-sessions interaction session X abstractness* (*contrast (4)*). Activation changes (ses-post - ses-pre) in the neural activation of the absSG condition and changes in the SWLS score (also ses-post - ses-pre) and activation changes (ses-post - ses-pre) in the neural activation of the conSG condition and improvement in SG fluency task (first - last training session) were correlated with *ggstatsplot package (Patil, 2021)* for *R* (R Core Team, 2013) via Pearson’s r.

Similarly, these neural activation changes were exploratory correlated with the outcomes from the SG fluency training task.

3 Results

3.1 Feasibility of the MSG Training

Analysis of Dropout Rate

In order to be able to make a statement about the feasibility of the training, we analyzed the number of dropouts during the training. In *hypothesis (1)* we suggested a dropout rate of less than 15% (considering reported dropout rates in comparable studies with SSD patients (e.g., (Gordon et al., 2018)) to be an indication for the feasibility of the MSG training program. For our study, a dropout rate of less than 15% would be equivalent to 4 (out of N=30 enrolled SSD patients) dropouts maximum. In our study, only one patient dropped out for private reasons[9].

Exploration of Satisfaction With the Training

To get a better impression about the satisfaction with the training and the subjective impressions about the effects of the training on everyday life (especially community functions), we additionally analyzed exploratory specifically outlined post training questionnaires. By majority, patients as well as their relatives were not only satisfied with the training but 94% of the patients rated the MSG training also as useful, 53% of the patients' relatives reported an increase of the patients' social contacts and 74% reported a gesture related improvement of perceptive, 66% of expressive communication skills (see *figures 7* and *8*).

Inconsistent numbers of observations result from withheld information from the patients or their relatives.

---

[9] This patient moved for work reasons. We assume that he would have continued the training otherwise because he reported high satisfaction with the training.

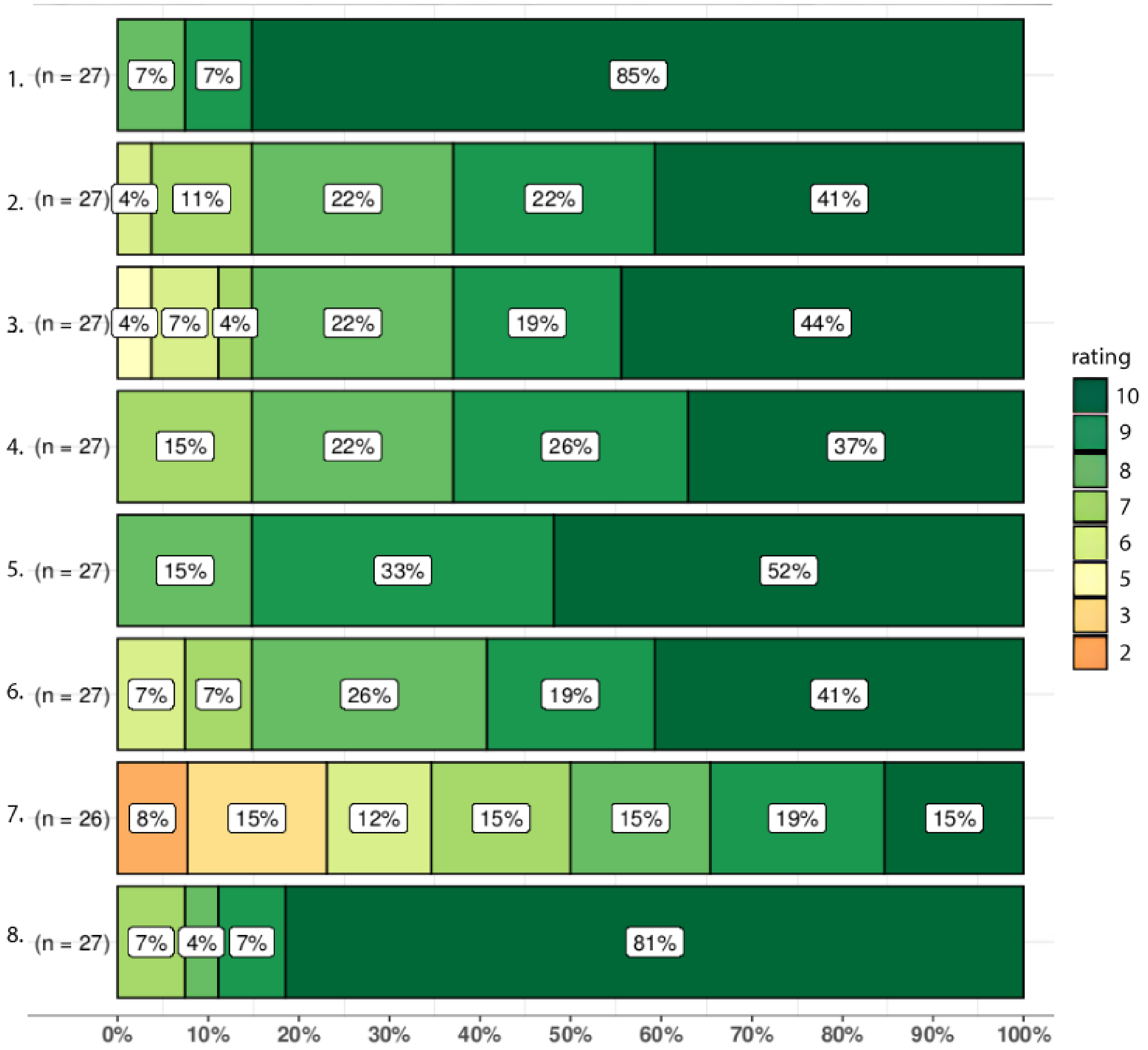


Figure 7: Patients' ratings of the MSG training. The questionnaire's scale ranging from 1 (very bad) to 10 (very good).

Statements about the MSG training to be evaluated:

1. I got along well with the training management.
2. I enjoyed the tasks in the training (motivation).
3. I found the training useful.
4. It was easy for me to work on the tasks in the training.
5. I made a lot of effort to work on the tasks in the training.
6. I've put a lot of effort into working through the application exercises
7. I also benefited from the training in everyday life.
8. I understood the tasks in the training.

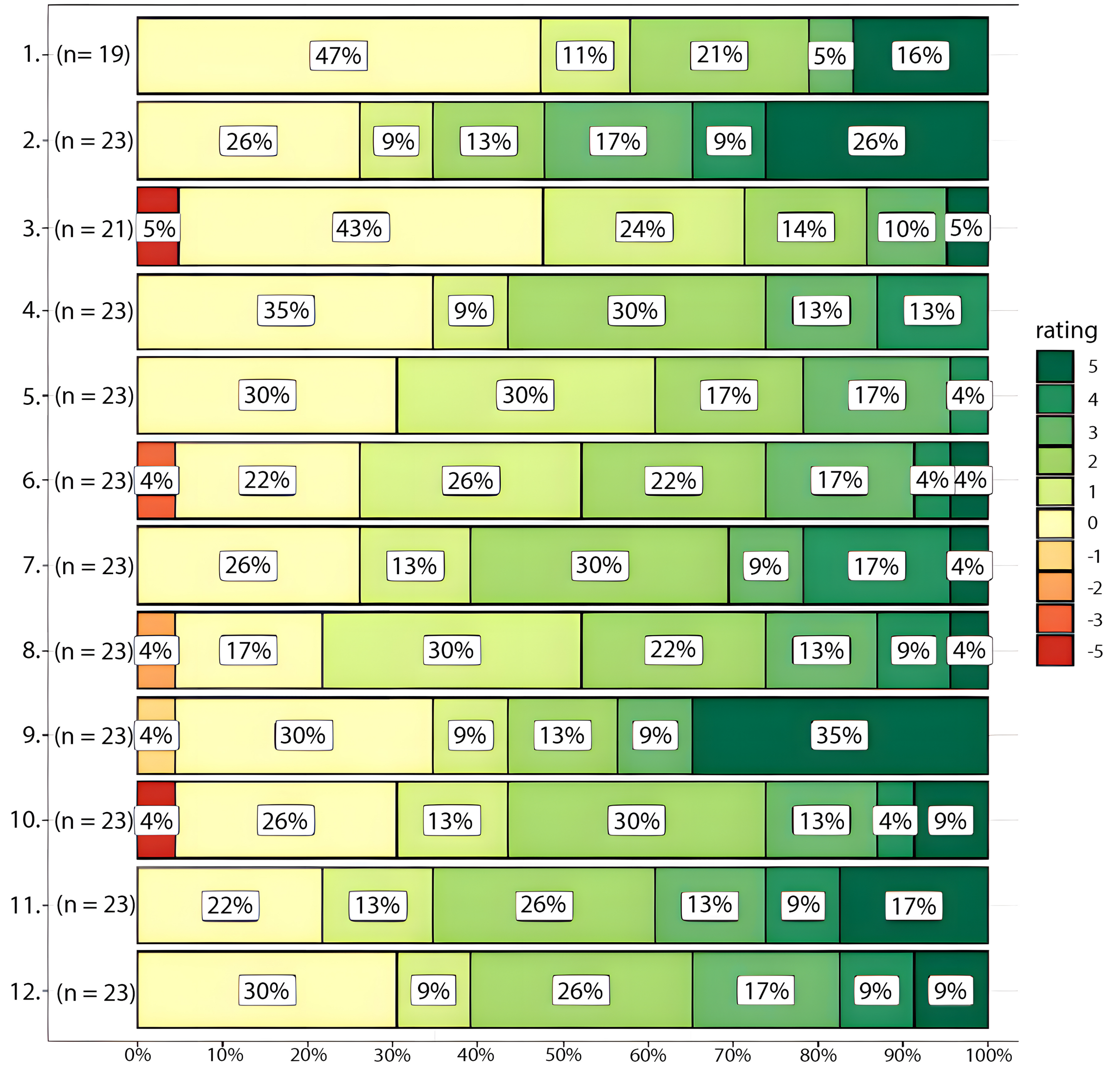


Figure 8: Patients' relatives' ratings of the training's impact on daily life. The questionnaire's scale ranging from -5 (much worse than before) to 5 (much better than before).

Questions:

1. How often does she/he have social contacts now compared to the state before training?
Can the study participant do the following, compared to the state before training:
2. Overall, adequately understand what the interlocutor is saying?
3. Make her-/hisself understandable only by gestures (e.g., in the case of loud ambient noises)?
4. Draw attention to her-/himself with gestures?
5. Make it clear in the conversation using gestures whether she/he is listening/bored/wants to speak for her-/himself (turn taking)?
6. Use gestures while speaking?
7. Use gestures that match her/his language/statement?
8. How often does the participant use gestures?
9. Communicate in an understandable way (e.g., express her-/himself clearly)?
10. Express emotions/emotional states using gestures/facial expressions?
11. Interpreting the other party's gestures correctly (appropriate reaction)?
12. Use gestures appropriate to the situation?

3.2 Changes in Quality of Life After MSG Training

Satisfaction With Life Score Before and After Training

In order to examine possible transfer effects of the MSG training on daily life, subjective quality of life of the subjects was investigated through a standardized psychological questionnaire (SWLS) and a specifically outlined questionnaire about nonverbal communication and social life before (ses-pre) and after (ses-post) the MSG training (note: SWLS in the control group was only measured in ses-pre). To test *hypothesis (2)*, in which we predicted a reduced overall quality of life in SSD patients, the summarized SWLS score was compared between groups (control group vs. SSD group) in ses-pre using a between-subjects t-test (*figure 9*).

To test the improvement in quality of life in the SSD group between measurements (ses-pre vs. ses-post) predicted in *hypothesis (3)*, a within-subjects t-test was calculated (*figure 10*).

Outcomes (overall SWLS scores) were checked for implausible values. Data from two patients[10] had to be excluded from further analysis, due to implausible information in the questionnaire (e.g., a massive reported increase in wellbeing or worsening of negative symptoms accompanied by temporary discontinuation of medication during the training period). In ses-pre (before the MSG training started), two patients and in ses-post (after the MSG training), five patients withheld information regarding satisfaction with life or filled in the SWLS questionnaire incompletely.

---

[10] One of the excluded participants reported an exceptional increase (by 10 points) and another an exceptional decrease (by -9 points) pre to post training in the SWLS questionnaire.

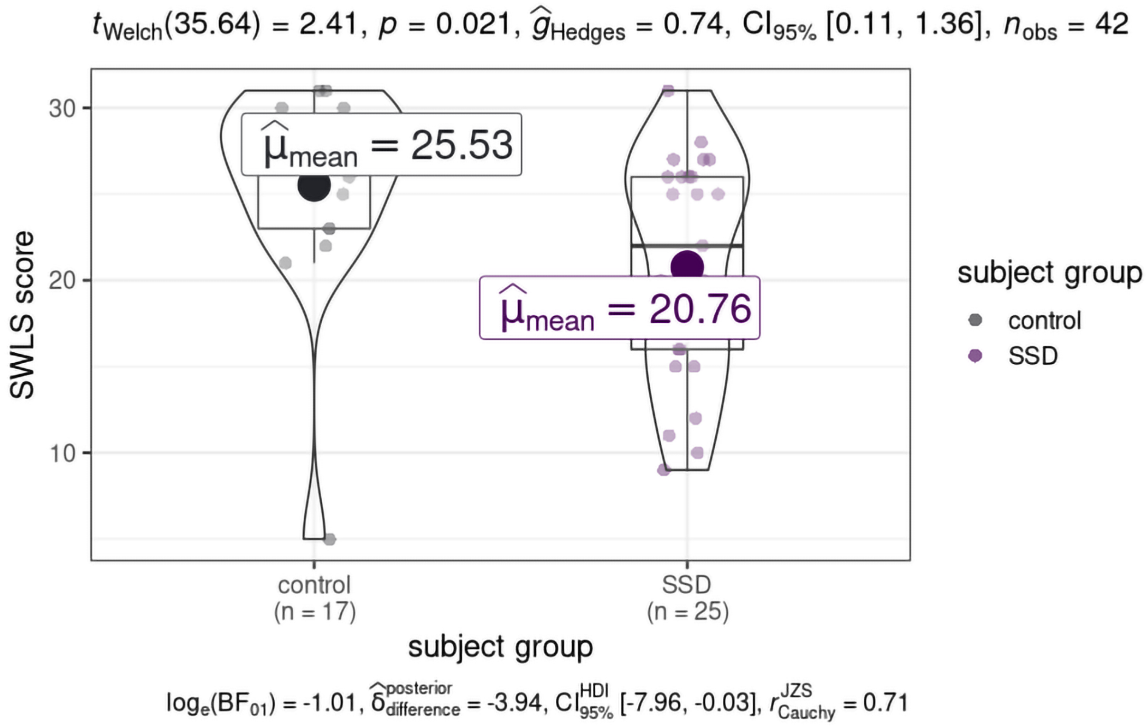


Figure 9: Group comparison (control group versus SSD group) of quality of life before the MSG training (ses-pre), examined using the SWLS score. The score ranges from a minimum of 5 points, which corresponds to the worst possible quality of life, to a maximum of 35 points, which corresponds to the highest possible quality of life.

SWLS = Satisfaction With Life Scale.

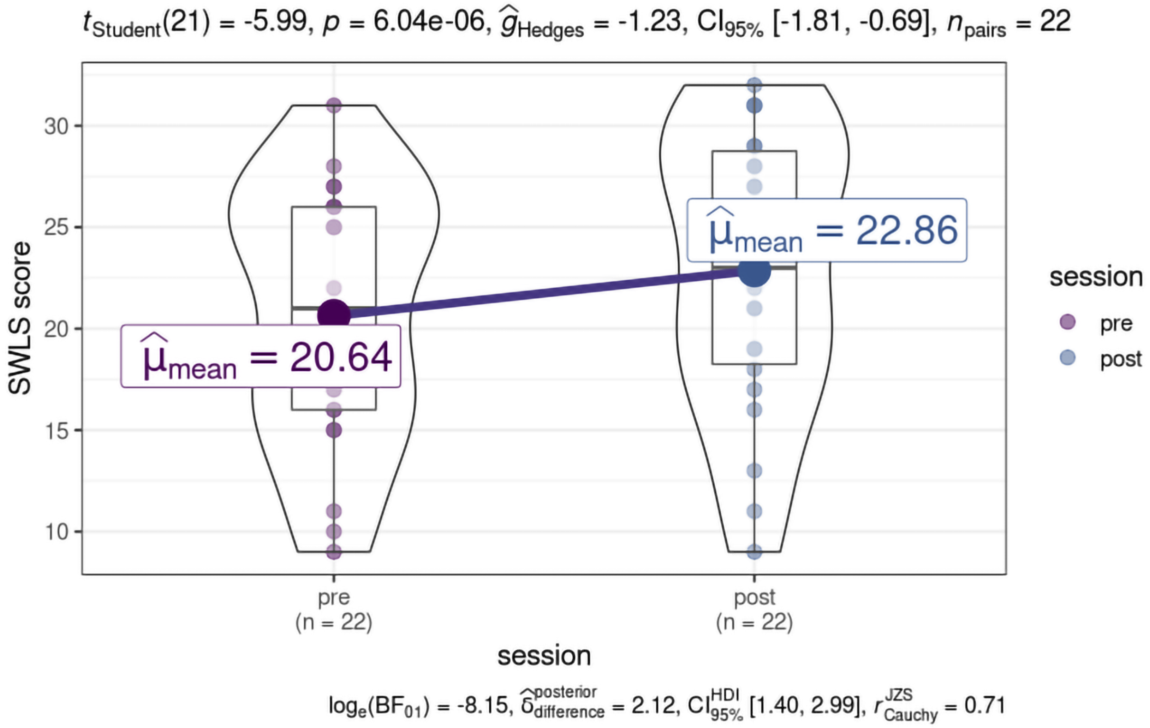


Figure 10: Pre-post MSG training comparison of quality of life in the SSD group, examined using the SWLS score. The score ranges from a minimum of 5 points, which corresponds to the worst possible quality of life, to a maximum of 35 points, which corresponds to the highest possible quality of life.

SWLS = Satisfaction With Life Scale.

The group comparison of the SWLS score in the control group versus SSD group (*figure 9*) confirmed *hypothesis (2)*, that patients suffer from a significant reduction in quality of life ($t_{Welch}(35.64) = 2.41$; $p = 0.021$; $\hat{g}_{Hedges} = 0.74$).

The pre-post training comparison in the SSD group comparing the SWLS score in ses-pre versus ses-post (*figure 10)* revealed a significant increase of life quality ($t_{Student}(21) = -5.99$; $p = 6.04e-06$, $\hat{g}_{Hedges} = -1.23$) in the patients after the MSG training, as predicted in *hypothesis (3)*.

3.3 Changes in Neural Activation After MSG Training

The investigation of conventional fMRI brain responses was based on a simple perception task of video sequences showing an actor performing gestures that accompany either abstract (absSG) or concrete (conSG) speech. Additionally, abstract and concrete unimodal conditions (speech only, S, or gesture only, G) were presented.

BOLD activation patterns were explored in whole-brain group level fMRI analyses using a 2 x 2 x 2 x 3 full factorial design (see *Group Level Analysis* in *chapter 2.5.3 FMRI Data Analysis*). All results are shown for an individual voxel type I error of $p < 0.01$ with a minimum of 221 contiguous resampled voxels, calculated per Monte-Carlo-Simulation (Slotnick, 2017; Slotnick et al., 2003). BOLD activation patterns are shown in the following figures, corresponding statistical parameters are listed in the tables with cluster coordinates based on the MNI template. Anatomical region names are derived from the *AAL* toolbox implemented in *SPM12* (Rolls et al., 2015; Tzourio-Mazoyer et al., 2002).

3.3.1 Session Comparison of Abstract Multimodal Processing

To clarify *neural changes for the processing of abstract multimodal videos in the SSD group* as predicted in *hypothesis (4)*, we calculated a session X abstractness interaction for the processing of multimodal input:

(1) SSD group interaction session X abstractness in the multimodal condition SG: *(ses-pre(absSG>conSG))* < *(ses-post(absSG>conSG))*

For the interaction of session (ses-pre, ses-post) X multimodal abstractness condition (absSG, conSG) in the SSD group, significant activation was found in three clusters including left parahippocampal regions, middle frontal and bilateral temporal regions as well as superior frontal regions, see *figures 11* and *12*. Anatomical regions, cluster extend, coordinates (MNI), t-values and no. of voxels of this contrast are summarized in *table 2*.

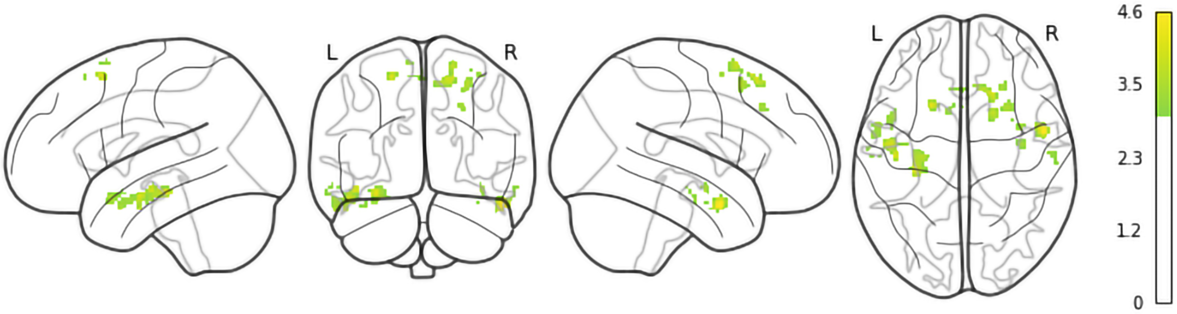


Figure 11: Glass brain visualization of neural activation for the interaction of session X abstractness in the multimodal condition in the SSD patients (*contrast (1)*): *(ses-pre(absSG>conSG))* < *(ses-post(absSG>conSG))*. Activation clusters were thresholded at $p < 0.01$, with a minimum cluster size of 221 voxels. Crosshair point at the right MTG [x = 52, y = 2, z = -26]; left PHG [x = -30, y = -24, z = -18] and left SFG [x = -20, y = 18, z = 58].

MTG = middle temporal gyrus; PHG = parahippocampal gyrus; SFG = superior frontal gyrus; L = left; R = right.

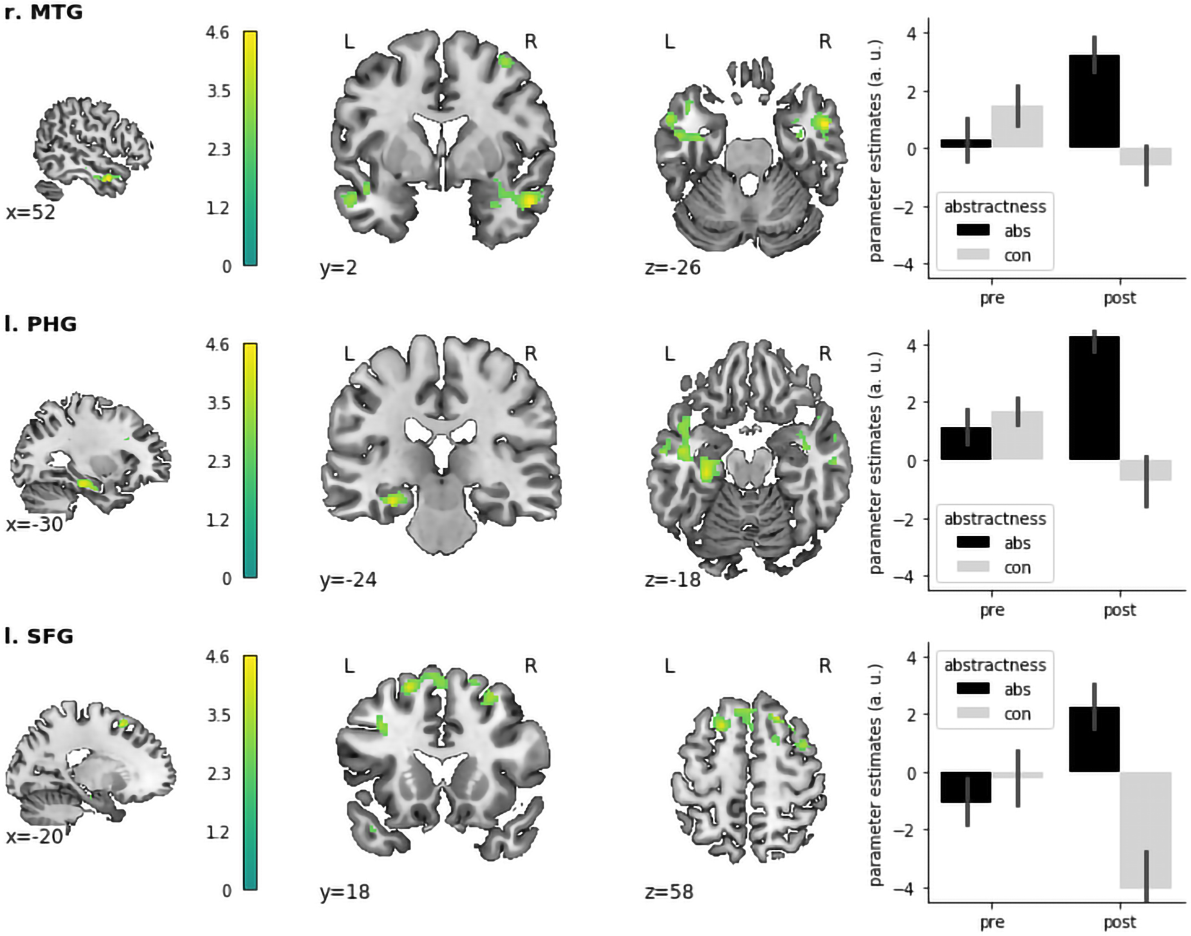


Figure 12: Neural activation clusters in slice view visualization for the interaction of session X abstractness in the multimodal condition in the SSD patients (*contrast (1)*): *(ses-pre(absSG>conSG)) < (ses-post(absSG>conSG))*. Activation clusters were thresholded at $p < 0.01$, with a minimum cluster size of 221 voxels, cluster level corrected at $p > 0.05$. Crosshair point at the right MTG [x = 52, y = 2, z = -26]; left PHG [x = -30, y = -24, z = -18] and left SFG [x = -20, y = 18, z = 58].

MTG = middle temporal gyrus; PHG = parahippocampal gyrus; SFG = superior frontal gyrus; L = left; R = right; abs = abstract multimodal condition (absSG); con = concrete multimodal condition (conSG); pre = session pre, first measurement (before no treatment period and before training); post = session post, last measurement (after training).

Anatomical region of peak	Cluster extend	MNI coordinates			t-Value	No. of voxels
		x	y	z		
Right MTG	TMP, ITG	52	2	-26	4.61	414
	HIP, INS, STG	38	-8	-14	3.48	
	MTG, ITG, STG	58	-14	-22	3.24	
Left PHG	HIP, FFG	-30	-24	-18	4.30	723
	MTG, STG, ITG	-46	-8	-20	4.06	
	MTG, ITG, STG	-44	-16	-16	4.01	
Left SFG	Left MFG, SMA	-20	18	58	4.18	727
	Right SFG, SMA, SFGmedial	20	24	56	4.01	
	Right SFG, SMA, MFG	20	12	64	3.82	

Table 2: Anatomical regions, cluster extend, coordinates (MNI), t-values and no. of voxels of the interaction of session X abstractness in the SSD group (*contrast (1)*): *(ses-pre(absSG>conSG)) < ses-post(absSG>conSG)*. Cluster level corrected at p < 0.05.

FFG, fusiform gyrus; HIP, Hippocampus; INS, Insula; ITG, inferior temporal gyrus; MFG, middle frontal gyrus; MTG, middle temporal gyrus; PHG, parahippocampal gyrus; SFG, superior frontal gyrus; SFGmedial, Superior frontal gyrus, medial; SMA, supplementary motor area; STG, superior temporal gyrus; TMP, temporal pole.

3.3.2 Group Comparison of Abstract Multimodal Processing

To test *hypothesis (5)*, we tested *group similarities for the processing of abstract multimodal videos* by applying conjunction 0 (minimum t-statistics) to explore if the predicted neural changes after the MSG training led to more commonalities between SSD patients and control group:

(2) overlap of abstractness effects in both groups, **ses-pre**: *(control group ses-pre (absSG > conSG)) ∩ (SSD group ses-pre (absSG > conSG))*

For the conjunction of groups (control group ∩ SSD group) in the multimodal abstractness contrast (absSG > conSG) in ses-pre, no significant activation was found, see *figure 13*.

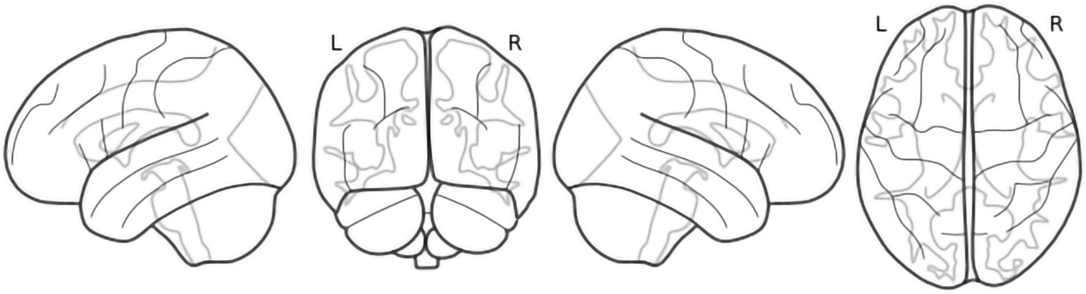


Figure 13: Glass brain visualization of neural activation for the conjunction of groups in the multimodal abstractness contrast (absSG > conSG) in the first fMRI measurement (ses-pre) (*contrast (2)*). Activation clusters were thresholded at $p < 0.01$, with a minimum cluster size of 221 voxels. No significant activation was found above the reported threshold.

L = left; R = right.

(3) overlap of abstractness effects in both groups, **ses-post**: *(control group ses-pre (absSG > conSG)) ∩ (SSD group ses-post (absSG > conSG))*

For the conjunction of groups (control group ∩ SSD group) in the multimodal abstractness contrast (absSG > conSG) in ses-post, significant activation was found in a left middle temporal cluster, see *figure 14*.

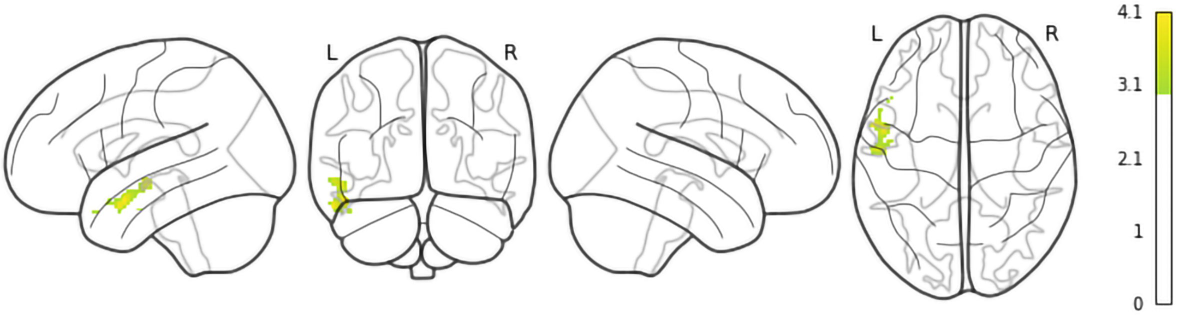


Figure 14: Glass brain visualization of neural activation for the conjunction of groups in the multimodal abstractness contrast (absSG > conSG) in the SSD group's after training fMRI measurement (ses-post) (*contrast (3)*). Activation clusters were thresholded at $p < 0.01$, with a minimum cluster size of 221 voxels. Crosshair point at the left MTG [x = -50, y = 2, z = -22].

MTG = middle temporal gyrus; L = left; R = right; ses-pre = before training measurement; ses-post = after training measurement.

3.3.3 Intervention Comparison of Abstract Multimodal Processing

To distinguish *specific MSG training effects in the processing of abstract multimodal videos (hypothesis (6)) from potential repetition effects during TAU*, we furthermore calculated a three-sessions-interaction X abstractness (F-test), comparing all three sessions (ses-pre, ses-bl, ses-post) in the SSD wait-first group (n=20):

(4) interaction session X abstractness for the multimodal condition (SG) in the SSD wait-first group

For the interaction (F-test) of session (ses-pre, ses-bl, ses-post) X multimodal abstractness conditions (absSG; conSG) in the SSD group, significant activation was found in a large left middle temporal cluster as well as in a smaller cluster in right temporal gyrus as well as a smaller cluster in the cerebellum, see *figures 15* and *16*. Anatomical regions, cluster extend, coordinates (MNI), t-values and no. of voxels of this contrast are summarized in *table 3*.

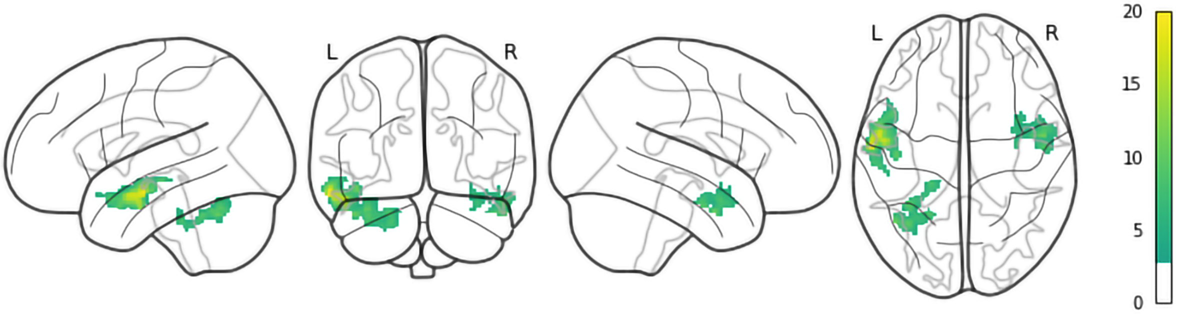


Figure 15: Glass brain visualization of neural activation for the interaction (F-test) of all three sessions (ses-pre, ses-bl, ses-post) X abstractness in SSD patients, wait-first group (*contrast (4)*). Activation clusters were thresholded at $p < 0.01$, with a minimum cluster size of 221 voxels. Crosshair point at the left MTG [x = -60, y = -10, z = -18], right MTG [x = 52, y = -2, z = -30] and left CERCRU1 [x = -44, y = -58, z = -32].

CERCRU1 = Crus I of cerebellar hemisphere; MTG = middle temporal gyrus; L = left; R = right; abs = abstract multimodal condition (absSG); con = concrete multimodal condition (conSG); pre = session pre, first baseline measurement (before no treatment period, TAU); bl = session bl, second baseline measurement (after Tau and before training); post = session post, last measurement (after training).

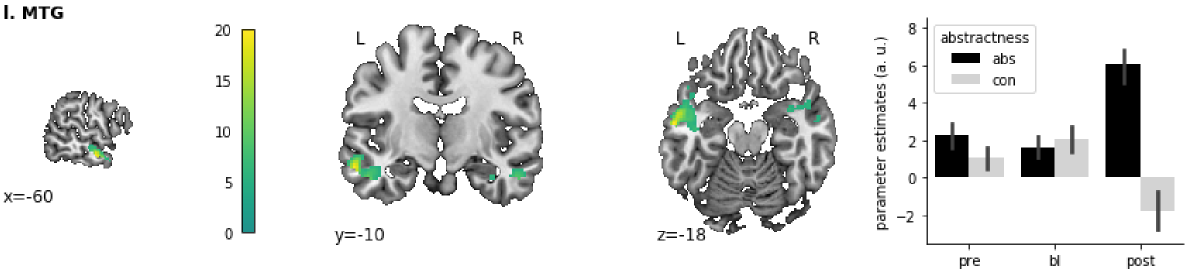


Figure 16: Neural activation clusters in slice view for the interaction (F-test) of all three sessions (ses-pre, ses-bl, ses-post) X abstractness in SSD patients, wait-first group (*contrast (4)*). Activation clusters were thresholded at p < 0.01, with a minimum cluster size of 221 voxels, cluster level corrected at p > 0.05. Crosshair point at the left MTG [x = -60, y = -10, z = -18], right MTG [x = 52, y = -2, z = -30] and left CERCRU1 [x = -44, y = -58, z = -32].

CERCRU1 = Crus I of cerebellar hemisphere; MTG = middle temporal gyrus; L = left; R = right; abs = abstract multimodal condition (absSG); con = concrete multimodal condition (conSG); pre = session pre, first baseline measurement (before no treatment period, TAU); bl = session bl, second baseline measurement (after TAU and before training); post = session post, last measurement (after training).

Anatomical region of peak	Cluster extend	MNI coordinates			F-Value	No. of voxels
		x	y	z		
Left MTG	ITG, STG	-60	-10	-18	19.94	652
	MTG, ITG, TPOsup	-50	-6	-22	16.41	
	MTG, ITG, STG	-42	-16	-16	10.69	
Right MTG	ITG, STG	52	-2	-30	11.02	336
	ITG, MTG, FFG	54	-10	-26	8.93	
	MTG, TPOmid, TPOsup	54	2	-22	8.55	
Left CERCRU1	CER6, ITG	-44	-58	-32	10.65	335
	CERCRU1, CER6, FFG	-36	-60	-28	9.99	
	CER6, CER4_5, CERCRU1	-22	-54	-30	9.32	

Table 3: Anatomical regions, cluster extend, coordinates (MNI), F-values and no. of voxels of the iteraction (F-test) of all three sessions (ses-pre, ses-bl, ses-post) X abstractness in SSD patients, wait-first group (*contrast (4)*). Cluster level corrected at $p < 0.05$.

CERCRU1 = Crus I of cerebellar hemisphere; CER4_5 = Lobule IV, V of cerebellar hemisphere; CER6 = lobule VI of cerebellar hemisphere; FFG, fusiform gyrus; ITG, inferior temporal gyrus; MFG, middle frontal gyrus; MTG, middle temporal gyrus; PHG, parahippocampal gyrus; SFG, superior frontal gyrus; STG, superior temporal gyrus; TPOmid, temporal pole: middle temporal gyrus; TPOsup, temporal pole: superior temporal gyrus.

3.4 Quality of Life Changes After Training and Their Neural Correlates

As a last step, we examined whether the activation in the peak clusters of the reported main interactions are related to the change in quality of life during the MSG training period in the patient group. For this purpose, eigenvariates from the *contrasts (1) SSD group interaction session X abstractness* and *(4) the three-sessions interaction session X abstractness* were extracted from the significant clusters. Thus, to test *hypothesis (7)*, activation changes (absSG ses-post - absSG ses-pre) in the neural activation and changes in the SWLS score (ses-post - ses-pre) were correlated via Pearson's r.

For *contrast (1)*, we found neural changes in the right temporal regions to be associated with the changes in the SWLS score (with $t_{Student}(20) = 3.04$, $p = 0.006$, $r_{Pearson} = 0.56$ for the right MTG). Changes in left and frontal regions seem to be less associated with changes in the SWLS score (with $t_{Student}(20) = 2.06$, $p = 0.053$, $r_{Pearson} = 0.42$ for the left PHG and $t_{Student}(20) = -0.38$, $p = 0.705$, $r_{Pearson} = -0.09$ for the SFG). Bonferroni-correction was applied for multiple comparisons, resulting in a corrected significance level of $p < 0.0167$ (α was corrected for the three clusters of neural activation $p < 0.05$). Results for correlations in this contrast are illustrated in *figure 17.*

For *contrast (4)*, we found neural changes in left temporal regions to be associated with changes in the SWLS score (with $t_{Student}(13) = 2.62$, $p = 0.021$, $r_{Pearson} = 0.59$), see *figure 18*.

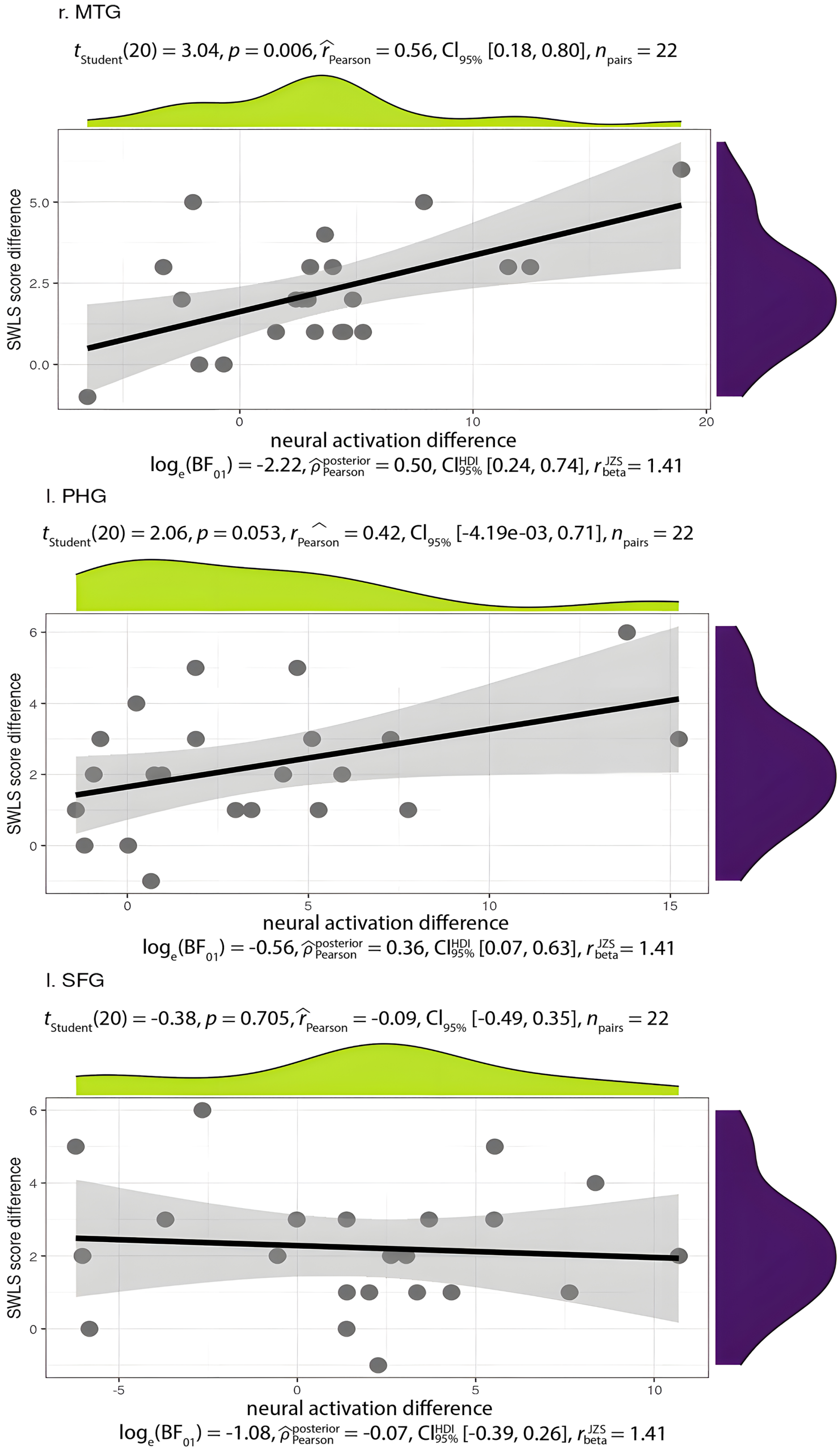


Figure 17:

Correlation of *contrast (1)* neural activation changes (ses-post - ses-pre, see *figure 12*) and quality of life (SWLS) score changes (ses-post - ses-pre, see *figure 10*) pre-post MSG training in SSD patients.

MTG = middle temporal gyrus;

PHG = parahippocampal gyrus;

SFG = superior frontal gyrus;

absSG = abstract multimodal condition;

SWLS = Satisfaction With Life Scale;

ses-pre = before training measurement;

ses-post = after training measurement.

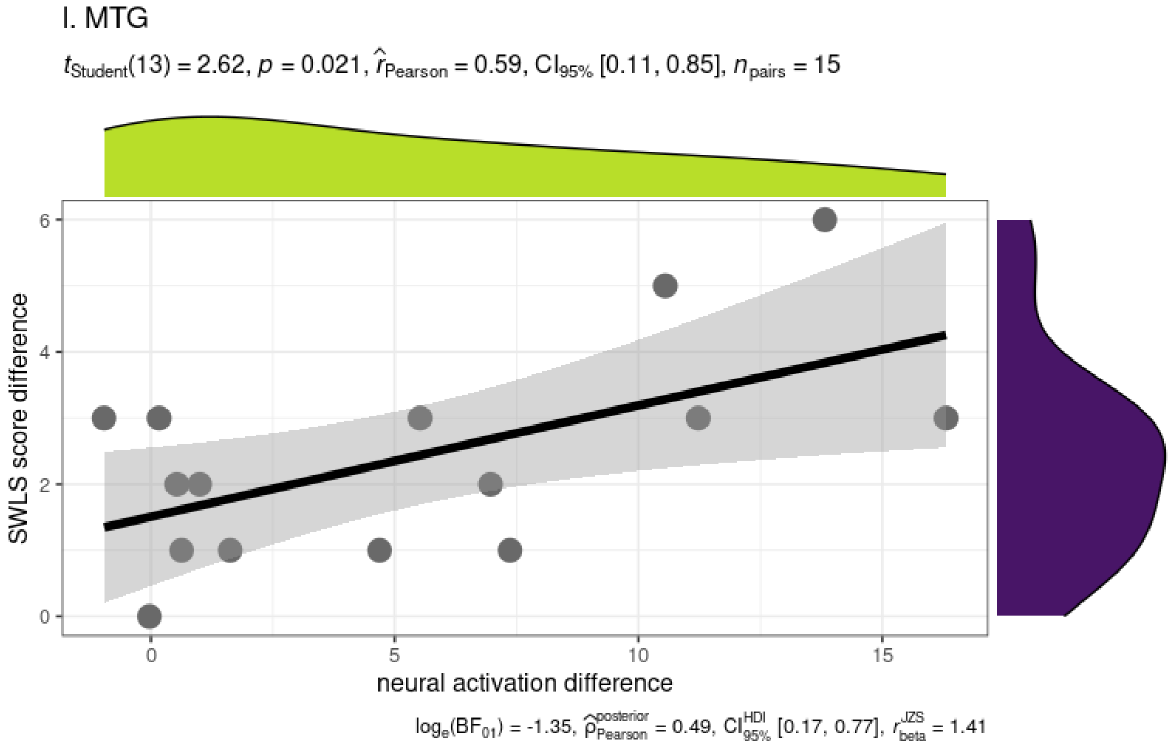


Figure 18:

Correlation of *contrast (4)* neural activation changes (ses-post - ses-pre, see *figure 16*) and quality of life (SWLS) score changes (ses-post - ses-pre) pre-post MSG training in SSD patients.

MTG = middle temporal gyrus; absSG = abstract multimodal condition; SWLS = Satisfaction With Life Scale.

3.5 Exploratory Analysis of Performance During Training

Exploration of Performance During MSG Training

The training process in the SG fluency task was analyzed exploratory.

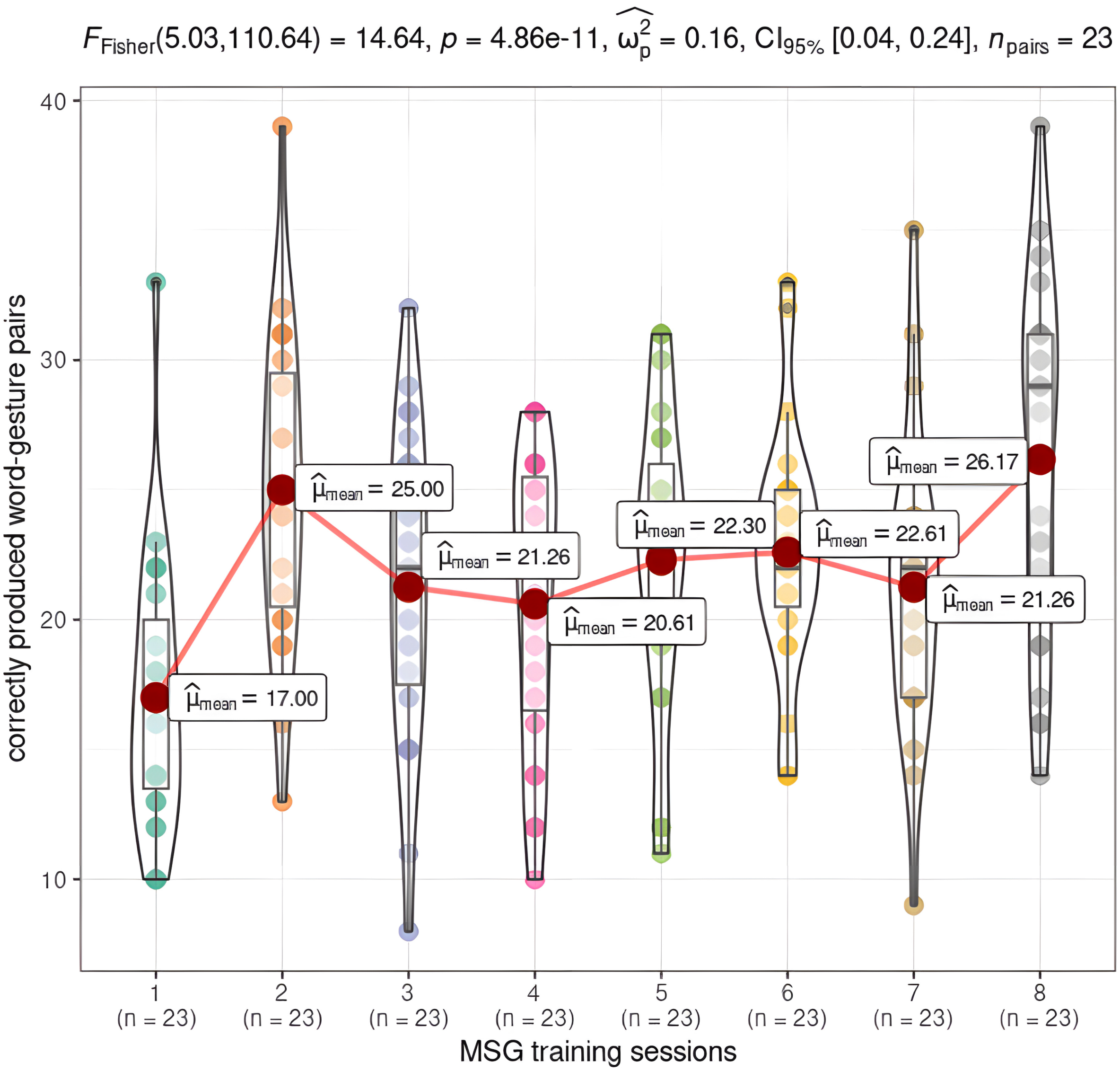


Figure 19: Performance of SSD group in the SG fluency training task over all eight MSG training sessions. SG fluency was measured counting correctly produced word-gesture pairs for each of the three semantic fields per session during one minute per semantic field.

SG fluency = speech-gesture fluency.

During the MSG training, the patients showed an increase in SG fluency performance (see *figure 19*), despite of the increasing complexity during the training.

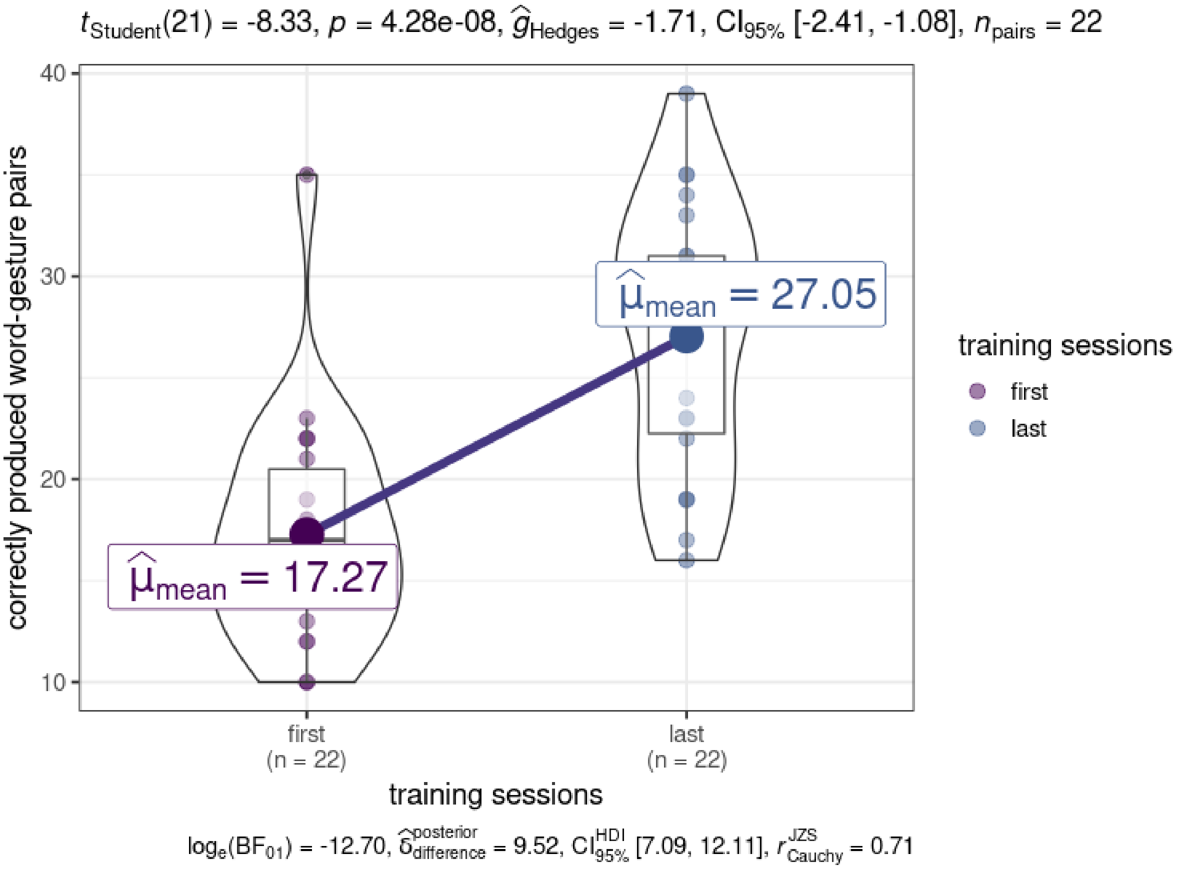


Figure 20: Performance of SSD group in the SG fluency training task in first versus last MSG training session. SG fluency was measured counting correctly produced word-gesture pairs for each of the three semantic fields per session during one minute per semantic field.

SG fluency = speech-gesture fluency.

A comparison of the performance from the first to the last MSG training session showed a significant increase in SG fluency task performance ($t_{Student}(21) = -8.33$; $p = 4.28e-08$; $\hat{g}_{Hedges} = -1.71$), see *figure 20.*

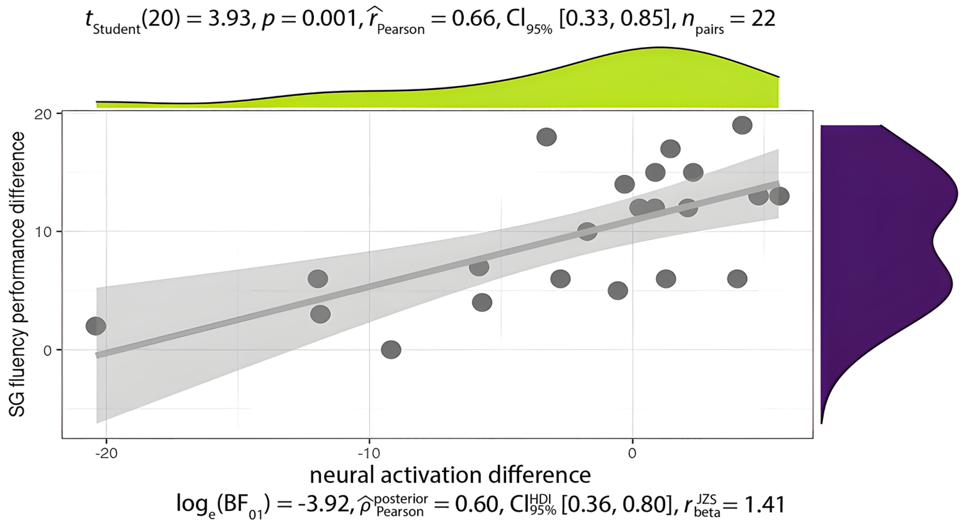
Exploration of Improvement During Training and Its Neural Correlates

As a secondary outcome, neural activation changes (conSG ses-post - conSG ses-pre) and the patients' improvement in the SG fluency task over training (last - first MSG training session) were correlated for exploration via Pearson’s r. As described in *chapter 3.4 Quality of Life Changes After Training and Their Neural Correlates*, again eigenvariates from the *contrasts (1)* and *(4)* were extracted from the significant clusters.

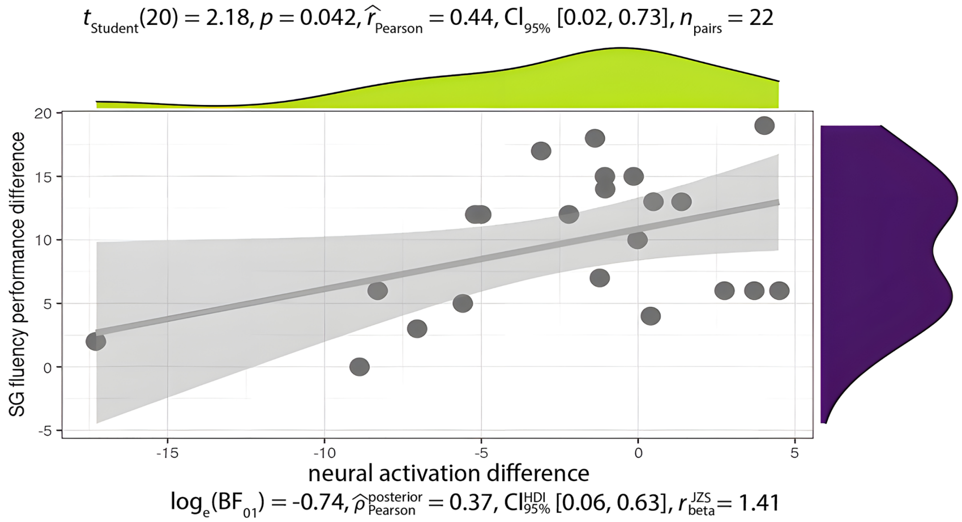
For *contrast (1)*, neural changes in right (with $t_{Student}(20) = 3.93$, $p = 0.001$, $r_{Pearson} = 0.66$ for the right MTG) and left temporal regions ($t_{Student}(20) = 2.70$, $p = 0.014$, $r_{Pearson} = 0.52$ for the SFG) and less in frontal regions (with $t_{Student}(20) = 2.18$, $p = 0.042$, $r_{Pearson} = 0.44$ for the left PHG) are associated with the patients' improvement over training. Bonferroni-correction was applied for multiple comparisons, resulting in a corrected significance level of $p < 0.0167$ (α was corrected for the three clusters of neural activation $p < 0.05$). Results for correlations in this contrast are also illustrated in *figure 21*.

For contrast *(4) the three-sessions interaction session X abstractness*, we found neural changes in left temporal regions to be associated with the patients' improvement in the SG fluency task over training (with $t_{Student}(12) = 2.61$, $p = 0.023$, $r_{Pearson} = 0.60$), see *figure 22*.

r. MTG



l. PHG



l. SFG

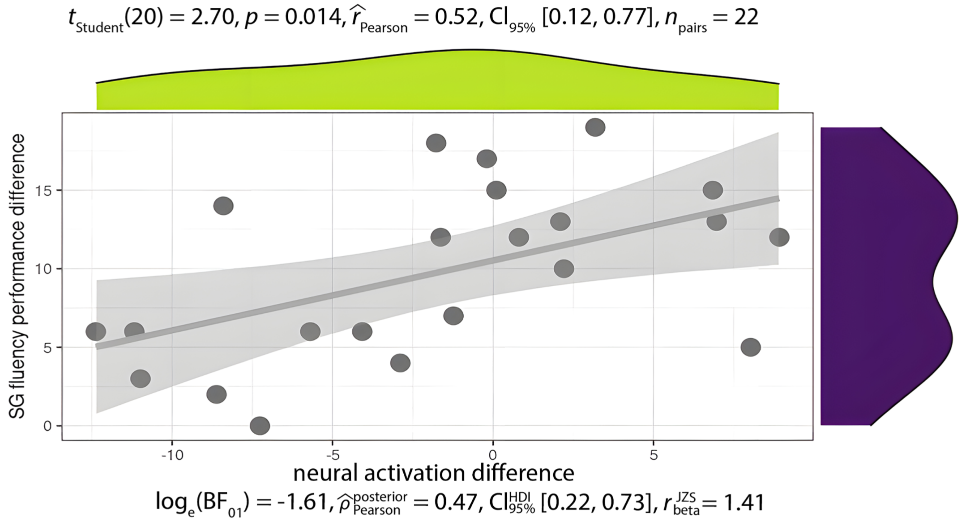


Figure 21:

Correlation of *contrast (1)* neural activation changes (ses-post - ses-pre, see *figure 12*) and SG fluency training task performance changes (last - first MSG training session, see *figure 20*) in SSD patients.

MTG = middle temporal gyrus;

PHG = parahippocampal gyrus;

SFG = superior frontal gyrus;

conSG = concrete multimodal condition;

SG fluency = speech gesture fluency;

ses-pre = before training measurement;

ses-post = after training measurement.

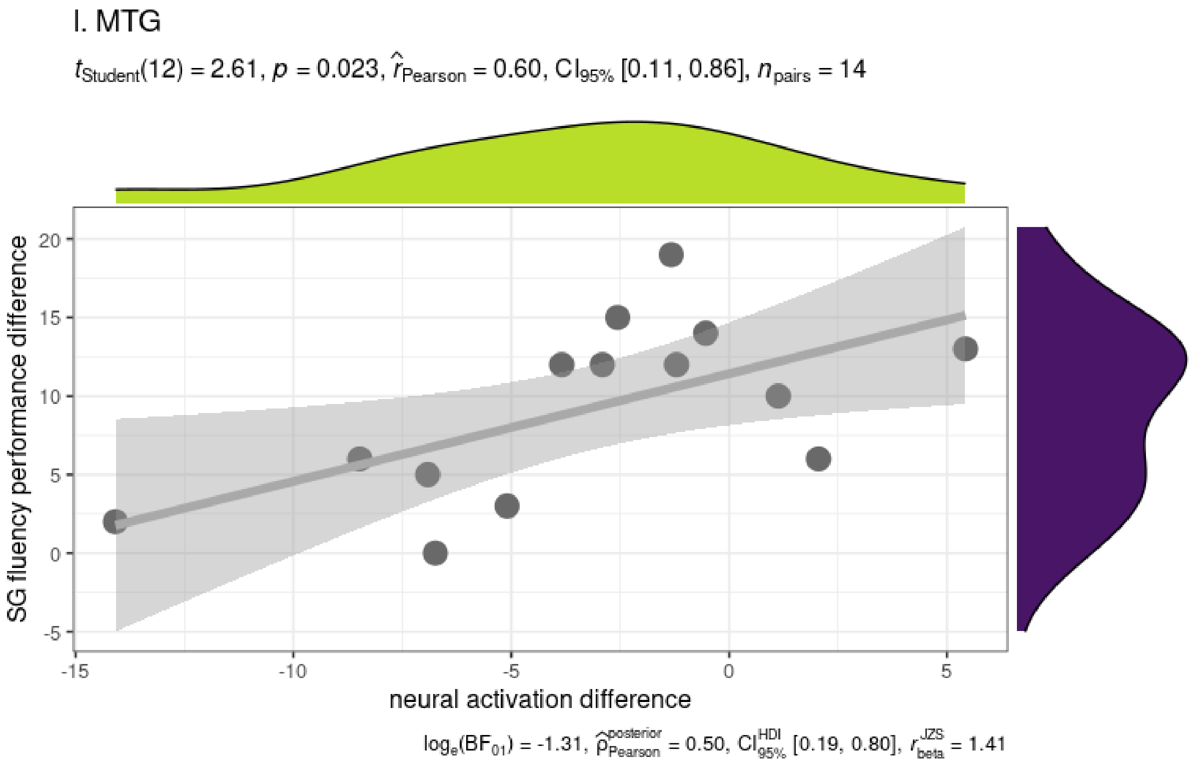


Figure 22:

Correlation of *contrast (4)* neural activation changes (ses-post - ses-pre, see *figure 16*) and SG fluency training task performance changes (last - first MSG training session, see *figure 20*) in SSD patients.

MTG = middle temporal gyrus; conSG = concrete multimodal condition; SG fluency = speech gesture fluency.

4 Discussion

In the present study, we examined the possible impact of a multimodal speech-gesture (MSG) training on the quality of life in patients suffering from schizophrenia spectrum disorder (SSD) and, further, its neural correlates.

SSD patients have been reported to exhibit impairments in the interpretation of abstract speech ("concretism", see *chapter 1.2.4 Speech Processing in SSD*) as well as in speech-gesture integration (see *chapter 1.2.5 Gesture Processing in SSD*), which is also reflected in dysfunctional neural activation (see *chapter 1.2.7 Speech, Gesture and Its Neural Correlates*). Abstract speech, such as cognitive metaphors, has a tremendous impact on interpersonal communication (see *chapter 1.2.2 Speech Processing*). Likewise, gestures play a fundamental role in communication processes (see *chapter 1.2.3 Gesture Processing*). A dysfunctional speech-gesture integration in abstract context as it is often reported for SSD patients (Nagels et al., 2019; Straube, Green, et al., 2013; Straube et al., 2014) thus affects integral elements of social skills which may lead to social isolation and therewith to a reduced quality of life (see *chapter 1.2.6 Speech, Gesture and Its Importance for Quality of Life*).

Although social-communicative dysfunctions are considered to be one of the core symptoms in schizophrenia (Kircher & Gauggel, 2008), especially symptoms concerning communication remain relatively stable under medication (Dollfus & Petit, 1995; Gaebel & Wölwer, 2010; Lavelle et al., 2014; Wüthrich et al., 2020). Studies in the last 20 to 30 years have proved psychotherapeutic intervention to be an effective treatment of schizophrenia and at least support the medical treatment (Falkai, 2016), but despite the exceeding evidence on specific speech and gesture dysfunctions in schizophrenia, these impairments are yet not explicitly targeted in conventional therapeutic interventions (Heim, 2020; Heim et al., 2019; Joyal et al., 2016; Riedl et al., 2020; Wüthrich et al., 2020), see *chapter 1.2.8 Current Treatments for Patients With SSD*. Considering gesture being an integral feature of interpersonal direct nonverbal communication that serves as an outward manifestation of several interacting fundamental processes, we therefore developed a novel

MSG training which focuses on the possible beneficial effects of gesture in speech-gesture integration processes, specifically designed for the needs of patients with SSD.

Here, we provide first promising evidence on the feasibility as well as on behavioral, social and neural outcomes of the MSG training program. The MSG training intervention contained perceptive rating and memory tasks, imitation and free productive tasks (e.g., speech-gesture (SG) fluency). Concerning the data that we collected, one of the current study's strengths was the brain imaging techniques that allowed us to investigate MSG training effects on adaptive processes in functional neural activation in patients with schizophrenia. Furthermore, we focused on cognitive aspects of gesture to facilitate social-cognitive functioning in everyday life. Due to the possible impact dysfunctional communication has on quality of life, transfer in everyday life social functioning was one of our main objectives. Hypotheses included a significant change in neural activation due to the training and a significant impact of the MSG training on quality of life. Secondary outcomes concern the association of performance increases during the MSG training (specifically, in the SG fluency task) with neural activation changes (see *chapter 1.3 Aim and Research Questions*). In the following, these results and their implications for future therapeutic approaches in psychiatry are discussed.

4.1 Feasibility of the MSG Training

First of all, in *hypothesis (1),* we suggested a dropout rate of less than 15% to be an indication for the feasibility of the MSG training program (compared with other studies involving SSD patients, e.g., (Gordon et al., 2018), see *chapter 1.2.8 Current Treatments for Patients With SSD*). With only one dropout (out of N=30 enrolled SSD patients) our rate is far below that. Considering furthermore that the only patient who dropped out was not able to continue the training because of work reasons, we assume that he would have probably continued and thus most likely did not drop out for feasibility reasons.

Regarding furthermore the overall positive ratings from the overwhelming majority of the patients and their relatives (see *figures 7* and *8* in *chapter 3.1 Feasibility of* *the MSG Training*) that we analyzed exploratory, the subjective impressions about the effects of the MSG training seem to be satisfying.

Taken together the low dropout rate and the high level of satisfaction with the MSG training, we evaluated our design and training as well feasible for this group of patients.

4.2 Changes in Quality of Life After MSG Training

In order to examine possible transfer effects of the MSG training on daily life, subjective quality of life of the SSD patients and control subjects was investigated through the standardized SWLS (German version of the Satisfaction With Life Scale (Glaesmer et al., 2011)) before (ses-pre) and after (ses-post) the MSG training (for the results, see *chapter 3.2 Changes in Quality of Life After MSG Training*).

As predicted in *hypothesis (2)*, quality of life in the patients was significantly reduced compared to the control group (see *figure 9*). Considering that only patients with stable symptoms were included in the experiment, who already were under medication and additionally had access to psychotherapy, the anyhow dramatically reduced quality of life suggests a strong impact of the remaining symptoms on the patients' well-being, which has already been discussed in *chapter 1.2.6 Speech, Gesture and Its Importance for Quality of Life*. Psychopathology in general (Eack & Newhill, 2007) and anxiety in particular (Huppert et al., 2001; Karow et al., 2005), as well as depression and negative symptoms, including the severe communication problems (Narvaez et al., 2008) and a lack of social interaction (Galuppi et al., 2010; Laliberte-Rudman et al., 2000; Solanki et al., 2008), are discussed to be responsible for the poor outcome in quality of life in (already medicated) patients.

As predicted in *hypothesis (3)*, quality of life in the SSD patients increased significantly after the MSG training (see *figure 10*), confirming the idea that severe communication problems in schizophrenia could be at least in parts be responsible for reduced quality of life and that therefore patients with schizophrenia might benefit from a specific speech-gesture training in their social-communicative functioning (Heim, 2020b; Joyal et al., 2016a).

These results are consistent with the exploratory investigated subjective evaluation of the training by the patients and their relatives (see *chapter 3.1 Feasibility of the MSG Training*): In a specifically outlined post training questionnaire (see *figure 7*), the overwhelming majority of the patients rated the MSG training as useful. Overall, the patients were very satisfied with the training in general. In another post training questionnaire (see *figure 8*), the patients' impressions have been confirmed by their relatives, who reported an increase of the patients' social contacts, improved perceptive communicative and improved expressive communicative skills.

Since communication and interpersonal skills are put forward to play a crucial role in social integration, communication impairments associated with a diagnostic of schizophrenia are hence a central issue regarding the patients' quality of life and functioning in society on both a personal and professional level (Gaebel & Wölwer, 2010; Joyal et al., 2016). The beneficial effects of a specific MSG training on quality of life in patients with schizophrenia are in line with the positive impressions of the training reported by the patients and their relatives. These results provide first evidence, that specific speech-language therapy focusing on the processing of abstract speech and nonverbal skills might be an appropriate approach to treat social-communicative deficits observed in patients with schizophrenia.

4.3 Changes in Neural Activation After MSG Training

Previous evidence on training of communicative and pragmatic skills in schizophrenia suggest that it may be an acceptable and tolerable intervention to conduct an intensive MSG training (Joyal et al., 2016), see *chapters 1.2.8 Current Treatments for Patients With SSD* and *2.3.1 Development of the MSG Training Intervention*.

Former studies have successfully demonstrated the effect of other cognitive training programs on neural activation in patients with schizophrenia (Bor et al., 2011; Edwards et al., 2010; Haut et al., 2010; Hooker et al., 2012; Penadés et al., 2013, 2020; Subramaniam et al., 2014; Vianin et al., 2014; Wexler et al., 2000; Wykes, 1998; Wykes et al., 2002). Therefore, adaptive processes of neural activity due to the MSG training could provide further evidence of the training program. With the help of pre-post-fMRI, we thus sought to investigate further implications from the MSG training program.

Session Effects of Neural Processing in SSD Group

In line with our *hypothesis (4)*, the interaction of session X abstractness of the within-subjects effects revealed a bilateral activation in temporal regions, including (para-)hippocampal and middle as well as superior frontal regions (*contrast (1)*, see *figures 11* and *12* in *chapter 3.3.1 Session Comparison of Abstract Multimodal Processing*).

Since Wernicke's area lies next to the cortical representation of hearing in the left temporal lobe, the Wernicke-Lichtheim-Geschwind model of language (see *chapter 1.2.2 Speech Processing*) suggests this region to be involved in the recognition of spoken language patterns or comprehension of spoken language (Geschwind, 1970). Nowadays, the temporal lobe is seen as less specific region but still is reported to be involved in supporting the retrieval of phonological forms (Binder, 2017), auditory word-form recognition (Ardila et al., 2016; DeWitt & Rauschecker, 2013) and language associations/integration processes (Ardila et al., 2016). According to the 'memory unification control' (MUC) model (Hagoort, 2013; Hagoort et al., 2009; Holler & Levinson, 2019; Willems & Hagoort, 2007), temporal regions are involved in fundamental integration processes and thus contribute crucially in the binding of information from multiple modalities like speech and gesture (for an explanation regarding the MUC model, see *chapter 1.2.7 Speech, Gesture and Its Neural Correlates*). Temporal regions were also reported in former studies to be active during speech-gesture integration (Dick et al., 2014; Green et al., 2009; Joue et al., 2020; Kircher et al., 2009; Straube, Green, Bromberger, et al., 2011a; Straube, Green, et al., 2013; Straube, He, et al., 2013; Straube et al., 2014) with the right hemisphere aiding in the integration of speech and gesture information, e.g., through 'coarse semantic coding' (Beeman, 1998; Kircher et al., 2009; Straube, Green, Bromberger, et al., 2011a). In our study, also the SSD group showed strong activation in temporal regions for the integration of speech and gesture, suggesting relatively normal multimodal integration processes (e.g., Wroblewski et al. (2020), see *chapter 1.2.7 Speech, Gesture and Its Neural Correlates*).

Neural activation changed during the MSG training not only in temporal regions, but also in parahippocampal regions, in the hippocampus and in a frontal cluster. The hippocampus has long been understood to be important for memory function, with left hippocampus contributing to verbal/narrative memory (Frisk & Milner, 1990). In schizophrenia, visual and verbal memory impairments have been reported being associated with abnormal activation in the hippocampus (Hanlon et al., 2011; Harrison, 2004; Kuperberg & Heckers, 2000; Straube, Green, et al., 2013; Tamminga et al., 2010) and this abnormal activation being connected to temporal dysfunctions (Cirillo & Seidman, 2003; Saykin et al., 1991, 1994; Shenton et al., 2001; Wible, 2012). The results from our study also imply a strong connection between temporal and hippocampal activation in patients with schizophrenia. Changes due to a cognitive training have already been reported to affect hippocampal structures in patients with schizophrenia (Falkai et al., 2017; Morimoto et al., 2018). In our MSG training, specific working memory tasks that were included in the training program might have had an influence. But since hippocampus activation furthermore seems to play a role in the successful binding of gesture information into an abstract sentence context (Straube et al., 2009), it is suggested that the increase of activation reflects training effects for speech-gesture binding processes.

Frontal regions seem to play a role as an important part in abstract content perception (A. M. Rapp et al., 2012) and in production (Beaty et al., 2017; Benedek et al., 2014), together with temporal and hippocampal regions, among others. Furthermore, superior and medial frontal activation could be explained by differences in social-emotional content between abstract and concrete conditions, which have been often found for social functioning, social cognition, theory of mind, or mentalizing, and are relevant for pragmatic comprehension of a communicative content (Krach et al., 2009; Straube et al., 2010; H. Uchiyama et al., 2006; H. T. Uchiyama et al., 2012). In our study, patients demonstrated stronger bilateral frontal activation after the MSG training, which might reflect a compensation strategy to integrate speech and gesture information on an abstract level.

In all three clusters, neural activation shows a specific increase (from ses-pre to ses-post) in the abstract multimodal condition (absSG). This pattern provides further evidence for some kind of adaptive processes toward the patterns observed in healthy control subjects, as it is reported for cognitive remediation therapy (Penadés et al., 2013, 2017). The changes in neural responses might reflect a specific training effect on brain regions relevant for metaphor comprehension (impaired in concretism) or the integration of abstract speech-gesture combinations (Bergemann et al., 2008; de Bonis et al., 1997; Iakimova et al., 2010; Kircher et al., 2007; Nagels et al., 2019; A. Rapp & Schmierer, 2010; Straube, Green, Bromberger, et al., 2011a; Straube, Green, et al., 2013; Straube et al., 2014). On the other hand, we found a decrease in neural activation for comparatively unimpaired processes like the processing of concrete concepts (conSG), which might express increased neural efficiency (A. M. C. Kelly & Garavan, 2005).

Commonalities of Neural Activation Between Groups

The finding of neural activation changes revealed by the session X abstractness interaction (*contrast (1)*) seems to be due to an increase in predominantly temporal regions in the SSD group: To investigate the commonalities of neural activation in SSD patients and controls before and after the MSG training program, we calculated conjunctions of within-subjects effects (see *chapter 3.3.2 Group Comparison of Abstract Multimodal Processing*). The conjunction of both groups (SSD group and control group) of the within-subjects effect of abstractness in the multimodal condition shows no significant activation pre MSG training (*contrast (2)*, see *figure 13*), but a cluster in left superior temporal gyrus post MSG training (*contrast (3)*, see *figure 14*). This increase in activation seems to lead to more similar activation in SSD patients and the control group after MSG training, as we predicted in *hypothesis (5)*. These similarities in neural activation in the abstract multimodal condition (absSG) after the training might provide evidence for some kind of adaptive processes toward the patterns observed in healthy control subjects, as it is reported for cognitive remediation therapy for patients with schizophrenia (Penadés et al., 2013).

Comparison of All Three Measurement Sessions in SSD Group

By testing *hypothesis (6)* we wanted to prove that the reported changes in neural activation arise from effects during the MSG training period, we calculated an interaction (F-test) of all three measurements (ses-pre, ses-bl and ses-post) (*contrast (4)*, see *figures 15* and *16* in *chapter 3.3.3 Intervention Comparison of Abstract Multimodal Processing*). Again, clusters in bilateral temporal regions and in the cerebellum with a peak in the left middle temporal gyrus were detected. Comparing the two sessions before the MSG training started (ses-pre and ses-bl with TAU in between), the pattern of activation in the peak cluster in left middle temporal gyrus (see *figure 16*) shows no significant differences, neither in the abstract multimodal condition (absSG), nor in the concrete multimodal condition (conSG), suggesting repetition effects being less relevant for the effects. In contrast, significant differences arise over the MSG training period (between ses-bl and ses-post) in the abstract mutlimodal (absSG) as well as the concrete multimodal (conSG) condition in the same direction (increase of activation in absSG and decrease in conSG), proving evidence for the specific effects during the MSG training period on the neural processing of multimodal integration of abstract content in patients with schizophrenia.

Summary of Neural Functioning

Taken these findings together, temporal regions (including hippocampus) seem to play a key role in the processing of multimodal input in an abstract content, as expected (Dick et al., 2009; He et al., 2018; Joue et al., 2020; Kircher et al., 2009; Rossetti et al., 2018; Straube, Green, Bromberger, et al., 2011a; Straube, Green, et al., 2013). According to the 'memory unification control' (MUC) model (Hagoort, 2013; Hagoort et al., 2009; Holler & Levinson, 2019; Willems & Hagoort, 2007), the binding of information from multiple modalities is a complex higher order cognitive process that requires semantic integration (Green et al., 2009; He et al., 2018; Özyürek et al., 2007) with general integration processes and fundamental information related to speech being processed in predominantly left temporal regions (see also *chapter 1.2.7 Speech, Gesture and Its Neural Correlates*). Common neural activation in temporal regions in SSD patients and controls suggests at least some unimpaired neural mechanisms in speech-gesture integration processes in SSD patients (Choudhury et al., 2021) and furthermore enable an adjustment of neural activation in these regions due to a specific MSG training intervention.

Taking into account the MUC model's assumption of complex higher-order integration processes of abstract multimodal input (unification, see *chapter 1.2.7 Speech, Gesture and Its Neural Correlates*) being processed in left inferior frontal gyrus (IFG) (Dick et al., 2009; He et al., 2015; Straube, Green, Bromberger, et al., 2011) and findings from former studies, we expected inferior frontal regions besides temporal regions to play a crucial role in 'concretism' (Rossetti et al., 2018), see also chapter *1.2.4 Speech Processing in SSD*, as well as in the speech-gesture integration problems in SSD patients (Straube, Green, Bromberger, et al., 2011a; Straube, Green, et al., 2013), see also chapter *1.2.5 Gesture Processing in SSD*. We furthermore expected changes in this region due to MSG training effects, because in former studies, neural activation was also found to modify in IFG due to remediation therapy (Penadés et al., 2017) and verbal working memory training (Wexler et al., 2000). Other than expected, we could not show changes in these regions. Indeed, we could not find differences in neural activation in the inferior frontal gyrus in patients compared to the controls before the MSG training started. The IFG is sensitive to mismatch manipulation of speech and gesture (Choudhury et al., 2021; Dick et al., 2009; Green et al., 2009; Steines et al., 2021; Straube et al., 2009; Willems et al., 2007). The absence of differences in neural activation of inferior frontal regions thus might reflect the relatedness of our speech-gesture pairs, also in the abstract condition.

To conclude, neural activation patterns in patients with schizophrenia change over the MSG training (*hypothesis (4)*, see *figures 11* and *12* in *chapter 3.3.1 Session Comparison of Abstract Multimodal Processing*) and are thus more similar compared with neural activation patterns in healthy controls after the MSG training as hypothesized (*hypothesis (5)*), but only in temporal and (para-)hippocampal regions, but not in the left IFG. Nevertheless, the changes of activation in left temporal regions seem to be a specific to the MSG training period, because comparing all measurement sessions in the SSD group (see *figures 15* and *16* in *chapter 3.3.3 Intervention Comparison of Abstract Multimodal Processing*), changes in neural activation were found for the MSG training period (comparing ses-bl and ses-post), but no significant changes were found for the waiting period (TAU period between ses-pre and ses-bl, *hypothesis (6)*). Finally, our results are to a considerable degree consistent with former studies on neural correlates of speech-gesture integration processes in schizophrenia and provide evidence for adaptive processes in neural activation during a specific MSG training intervention.

4.4 Quality of Life Changes After Training and Their Neural Correlates

If components of the MSG training and the changes at the neural level are responsible for the improvement of the quality of life, then these changes should also correlate in our design. Therefore, we correlated these effects to examine the relationship of the MSG training, outcomes in quality of life and neural activation changes.

Strikingly, when we correlated the neural changes in right temporal regions of *contrast (1)* (see *figures 11* and *12*) and the SWLS score changes (see *figure 10*), we indeed found associations in the abstract multimodal condition (see *figure 17* in *chapter 3.4 Quality of Life Changes After Training and Their Neural Correlates* for right MTG). The same is true for neural changes in left temporal regions of *contrast*

(4) (see *figure 18* in *chapter 3.4 Quality of Life Changes After Training and Their Neural Correlates*). With this finding, we proved our last and main *hypothesis (7)*, in which we predicted associations between changes in neural activation and quality-of-life improvements in patients with SSD.

For *contrast (1)*, diverging from our findings in right temporal regions, neural activation changes in left parahippocampal and frontal regions did not correlate with the changes in quality of life (see *figure 17* for left PHG and left SFG). This could be due to the wide spread distribution of activation, especially in the frontal cluster. Also the neural activation in frontal regions might show some other effects than the MSG training's effect on quality of life. Possible effects are a matter of speculation so far.

Additionally, we exploratory correlated the neural changes in the reported regions of *contrasts (1)* and *(4)* with the performance increase over the MSG training (see *figure 20*) and found associations in the concrete multimodal condition (see *figure 21* for *contrast (1)* and *figure 22* for *contrast (4)* in *chapter 3.5 Exploratory Analysis of Performance During Training*). These associations seem plausible, since the SG fluency training task contained a very concrete instruction for action: The patients accordingly produced appropriate gestures with which they could "*grasp*" (see *chapter 1.2.3 Gesture Processing*) the concrete as well as the abstract concepts in this task. These exploratory results need further investigation, but provide interesting insights into potential resources of integration processes that are reported to be relatively unimpaired in SSD (He et al., 2021; Kircher et al., 2009; Straube et al., 2009; Straube, Green, Bromberger, et al., 2011a; Straube, Green, et al., 2013).

Taken together, these results reveal beneficial MSG training effects on quality of life being associated with neural activation changes involved in abstract speech-gesture integration in patients with schizophrenia. This links the potential benefits of a specific training of social-communicative skills with improvement in everyday life in patients with SSD. Reports of relatives support this finding, indicating improved communication skills and increase of social contacts after MSG training in the majority of patients (see *figure 8*).

5 Limitations

Given the study design, some methodological limitations need to be considered for a valid interpretation of the results.

The first set of limitations addresses the sample. Due to recruitment problems and dropouts in the control group, we had to modify the number of participants in the control group and thus the study design, as outlined in section *2.2.2 Modifications of the Study Design*. As a consequence, a direct comparison of MSG training effects in both subject groups (SSD patients versus controls) was not possible. We were only able to compare patients and controls before and after the patients (but not the control subjects) conducted the MSG training. It would have also been interesting, if controls would have benefited from the training and if this potential benefit would have also been associated with outcomes in quality of life and if the intervention groups would have had different outcomes in patients and healthy control subjects. Problems might also arise from the SSD group: Because our inclusion criteria were very strict in terms of medical contraindications against fMRI measurements and the requirement of relatively stable symptoms in the patient group, our sample might be an imperfect representative for patients with SSD in general. In consequence, our results might not be transferable to other patients with the same or a similar diagnosis, particularly to patients suffering from an acute episode. After this successful results using a waiting list controlled pilot trial, comparing two active treatment arms would be important to demonstrate the unique contribution of the MSG training intervention.

A second set of limitations considers the analysis of the diverse tasks and items we used in the MSG training. It would be of great interest, which of these tasks and which conditions of the training items show particularly beneficial effects, being associated with the neural changes and changes in quality of life. Furthermore, the possible influence of the trainers' interpersonal style should be taken into consideration. Due to concerns regarding the patient's acceptance of the MSG training, we had decided for a time limit of 60 minutes per training session. Furthermore, the items we used increased in complexity from session to session in order to facilitate the training effect. For these reasons, some of the tasks do have a lack of items or might not be suitable for a sufficient and detailed analysis. Future studies should investigate the most efficient aspects of the MSG training in terms of the setting, the training tasks and the items (conditions etc) in order to clarify how exactly gesture helps SSD patients through disclosing the concrete concepts of the source domain in a cognitive metaphor (see also *chapters 1.2.3 Gesture Processing* and *1.2.5 Gesture Processing in SSD*). Of particular interest could be the novel SG fluency task exclusively created for the MSG training. As this task is the most suitable task for studying training effects including abstract and concrete production of speech accompanied by gesture, the outcomes of this task could demonstrate even small between-subject as well as within-subject differences.

Importantly, the results of this exploratory study require further validation in independent studies. For this purpose, we provide detailed descriptions of the study design, fMRI paradigms as well as training material on our public project on the open science framework (OSF: DOI 10.17605/OSF.IO/UH4F9), which is freely available under CCBy Attribution 4.0 International license.

6 Conclusions and Outlook

Social-communicative dysfunctions in schizophrenia have received increased interest from the field of clinical neuroscience. SSD patients have been reported to exhibit impairments in the interpretation of abstract speech ('concretism') (Bergemann et al., 2008; de Bonis et al., 1997; Iakimova et al., 2010; Kircher et al., 2007; A. Rapp & Schmierer, 2010; Rossetti et al., 2018) as well as in speech-gesture integration (Berndl, Cranach, et al., 1986; Berndl, Grüsser, et al., 1986; Bucci et al., 2008; Choudhury et al., 2021; Goss, 2011; Green et al., 2008; Karakuła et al., 2013; Lavelle et al., 2013; P. Martin et al., 1994; Matthews et al., 2013; Mittal et al., 2006; Nagels et al., 2019; Park et al., 2008; Troisi et al., 1998; Walther et al., 2013a, 2013b, 2015; Walther, Mittal, et al., 2020; Walther & Mittal, 2016; Wüthrich et al., 2020). Despite ample evidence of dysfunctional social communication in schizophrenia and its association with poor outcomes for patients' quality of life (Bambini et al., 2016; Falkai, 2016; Gaebel & Wölwer, 2010), there have been only few studies on speech therapy for patients suffering from schizophrenia (Heim, 2020; Heim et al., 2019; Joyal et al., 2016). Therefore, development of innovative add-on treatments with a focus on quality of life is highly needed (Falkai, 2016; Falkai et al., 2017; Heim, 2020; Joyal et al., 2016). Considering the tremendous impact of figurative speech (Kircher & Gauggel, 2008; Lakoff & Johnson, 2008) and gesture's fundamental role in social communication and functioning (Goldin-Meadow & Alibali, 2013; Suffel et al., 2020; Walther & Mittal, 2016), we developed a specific multimodal speech-gesture (MSG) training to address social-communicative dysfunctions and thereby overall quality of life in patients with schizophrenia.

Despite the number of limitations in the present study, the overall analyses evaluating the MSG training provide evidence for the potentially benefits of a multimodal training program. Particularly noteworthy are the patients' significantly increased quality of life scores after the MSG training program. With the help of fMRI measurements, we furthermore proved MSG training effects on neural adaptive processes: The data support the assumption of temporal regions, including (para-)hippocampal regions, being a relevant area for abstract speech-gesture integration in patients with schizophrenia and neural activation in these regions being possibly influenced by a specific MSG training program. Importantly, changes in neural activation and beneficial training effects on the quality of life in patients are closely connected. This is an outstanding argument for the chances that come with novel add-on treatments for patients suffering from schizophrenia. A specific training program which addresses gesture and speech integration may thus offer the opportunity to complement the currently recommended treatments and enable the patients to broaden their experiences.

Former studies reported speech-gesture integration deficits not only for patients with schizophrenia, but also for other diseases, including autism spectrum disorder (Eigsti, 2013; Fourie et al., 2020; Hubbard et al., 2012; MacNeil & Mostofsky, 2012; Redcay, 2008; Silverman et al., 2010; Stieglitz Ham et al., 2008) and depression (Annen et al., 2012; Fiquer et al., 2018; Gupta et al., 2021; Segrin, 1990). It was suggested that aberrations in functional brain activation during multimodal integration processes, specifically aberrations in temporal regions, provide a possible common etiology for integration deficits in several clinical populations (Fourie et al., 2020; Redcay, 2008; Stevenson et al., 2011; Suffel et al., 2020). Thus, a MSG training might be a promising therapy approach not only for patients with schizophrenia, but also for patients suffering from other psychiatric diseases.

The overall analyses evaluating the MSG training provide extraordinarily promising results which should be validated and extended in further independent studies. Especially the subjectively reported transfer effects of the MSG training, the changes in patients' quality of life and the associated neural changes in temporal regions provide evidence for the possible beneficial effects of innovative add-on treatments. Future studies should investigate combinations of speech-gesture training with neural stimulation techniques like transcranial magnetic stimulation (Walther, Kunz, et al., 2020) or transcranial direct current stimulation (Schülke & Straube, 2019), which seem to be further promising approaches (Cavelti et al., 2018).

Summary

Dysfunctional social communication is one of the most stable characteristics in patients with schizophrenia that severely affects quality of life. Interpreting abstract speech and integrating nonverbal information is particularly affected.

Considering the difficulty to treat communication dysfunctions with usual intervention, we investigated the possibility to improve quality of life and co-verbal gesture processing in patients with schizophrenia by applying a multimodal speech-gesture (MSG) training.

In the MSG training, we offered eight sessions (60 min each) including perceptive and expressive tasks as well as meta-learning elements and transfer exercises to 29 patients with schizophrenia spectrum disorder (SSD). Patients were randomized to a waiting-first group (N=20) or a training-first group (N=9), and were compared to healthy controls (N=17). Outcomes were quality of life and related changes in the neural processing of abstract speech-gesture information, which were measured pre-post training through standardized psychological questionnaires and functional Magnetic Resonance Imaging, respectively.

Pre-training, patients showed reduced quality of life as compared to controls but improved significantly during the training. Strikingly, this improvement was correlated with neural activation changes in the middle temporal gyrus for the processing of abstract multimodal content. Improvement during training, self-report measures and ratings of relatives confirmed the MSG-related changes.

Together, we provide first promising results of a novel multimodal speech-gesture training for patients with schizophrenia. We could link training induced changes in speech-gesture processing to changes in quality of life, demonstrating the relevance of intact communication skills and gesture processing for well-being.

Clinical Trial Registration: DRKS.de, identifier DRKS00015118

Study Protocol: Riedl, L., Nagels, A., Sammer, G., & Straube, B. (2020). A multimodal speech-gesture training intervention for patients with schizophrenia and its neural underpinnings–the study protocol of a randomized controlled pilot trial. *Frontiers in psychiatry*, *11*, 110.

Zusammenfassung

Dysfunktionale sozial-kommunikative Fähigkeiten sind eines der stabilsten Merkmale bei Patient*innen mit Schizophrenie, die die Lebensqualität stark beeinträchtigen. Die Interpretation abstrakter Sprache und die Integration nonverbaler Informationen sind hierbei besonders betroffen.

Da Kommunikationsstörungen bei Schizophrenie von herkömmlichen Therapien bisher wenig beeinflusst werden, haben wir ein spezifisches Sprach-Gestik-Training (MSG-Training) entwickelt, um die multimodale Verarbeitung und damit auch die sozial-kommunikativen Fähigkeiten sowie die Lebensqualität der Patient*innen zu verbessern.

Im Rahmen des MSG-Trainings führten wir mit 29 Patient*innen mit Schizophrenie-Spektrum-Störung (SSD) in acht Sitzungen (je 60 Minuten) ein Training durch, das perzeptive und expressive Aufgaben sowie Elemente des Meta-Lernens und Transferübungen umfasste. Die Patient*innen wurden nach dem Zufallsprinzip einer Wartegruppe (N=20) oder einer Trainingsgruppe (N=9) zugeteilt und mit gesunden Kontrollpersonen (N=17) verglichen. Untersucht wurden die Lebensqualität und die damit assoziierten Veränderungen in der neuronalen Verarbeitung abstrakter Sprach-Gestik-Informationen, die vor und nach dem Training mittels standardisierter psychologischer Fragebögen bzw. mit funktioneller Magnetresonanztomographie gemessen wurden.

Vor dem Training wiesen die Patient*innen im Vergleich zu den Kontrollpersonen eine geringere Lebensqualität auf, die sich jedoch während des Trainings deutlich verbesserte. Interessanterweise waren diese Lebensqualitätssteigerungen mit neuralen Aktivierungsänderungen im mittleren temporalen Gyrus für die Verarbeitung abstrakter multimodaler Inhalte korreliert. Die ebenfalls gemessene Leistungssteigerung während des Trainings sowie Selbsteinschätzungen und Bewertungen von Angehörigen untermauern diese Befunde.

Insgesamt liefern wir erste vielversprechende Ergebnisse zur Wirksamkeit des neu entwickelten MSG-Trainings für Patient*innen mit Schizophrenie. Wir konnten trai-

ningsinduzierte Veränderungen in der Sprach-Gestik-Verarbeitung mit Steigerungen der Lebensqualität in Verbindung bringen und damit die Bedeutung intakter sozial-kommunikativer Fähigkeiten für das allgemeine Wohlbefinden aufzeigen.

Appendices

Proof of Concept Contrasts of Neural Activation

Exploratory proof of concept contrasts were calculated:

For the gesture contrasts SG > G and SG > S and for the contrast of abs > con conditions, conjunctions 0 (minimum t-statistics) were calculated to examine overall *group similarities.*

The conjunction of group contrasts for effects of speech (*contrast (I)*, see *figure 23*) and for effects of gesture (*contrast (II)*, see *figure 24*) over both groups and measurements show strong bilateral temporal activation and strong bilateral occipital and superior parietal activation, respectively, which can be rated as a proof of principle.

The conjunction of group contrasts, *contrast (III)*, shows furthermore a common activation in the left middle temporal gyrus (MTG) for the effects of abstractness (see *figure 25*), suggesting that the temporal lobe is similarly engaged in control subjects and SSD patients. We have previously shown similar common neural activation for the perception of metaphoric gestures in healthy subjects and SSD patients (Choudhury et al., 2021). These findings suggest at least some unimpaired neural mechanisms that are relevant for abstractness processing and audiovisual integration, which may provide the basis for a successful MSG training intervention focusing on the potentially disrupted processing of abstract context ('concretism') in schizophrenia.

In healthy subjects, in addition to predominantly left temporal regions, also the inferior frontal gyrus (IFG) was found to play a crucial role in the processing of abstract compared to concrete speech content (A. M. Rapp et al., 2012). Former studies found aberrant inferior frontal activation in SSD patients for abstractness processing (Jeong et al., 2009; Kircher et al., 2007; Rossetti et al., 2018) and gesture integration in an abstract content (Straube et al., 2014; Straube, Green, et al., 2013).

The lack of common activation in the IFG in controls and SSD patients is thus no exceptional finding.

(I) Effects of speech: *(control group ses-pre (SG > G))* ∩ *(SSD group ses-pre (SG > G))* ∩ *(SSD group ses-post (SG > G))*

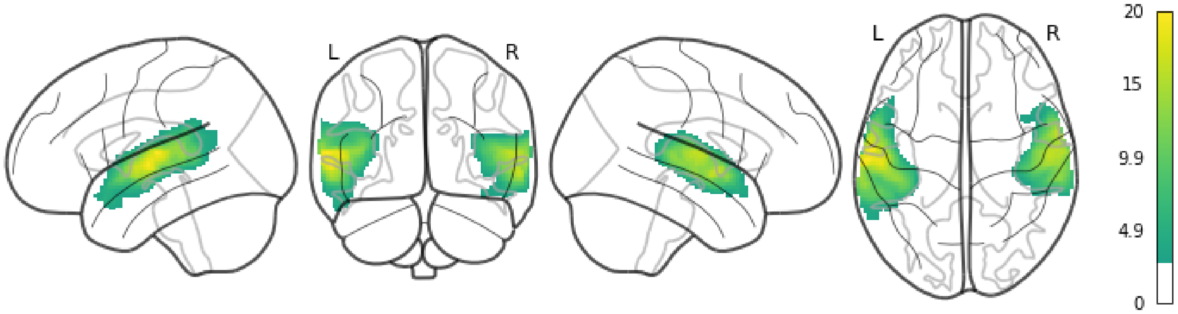


Figure 23: Glass brain visualization of neural activation for verbal contrast (SG > G) in integration (conjunction 0, minimum t-statistics) of groups and sessions. Activation clusters were thresholded at $p < 0.01$, with a minimum cluster size of 221 voxels. Crosshair points at the left STG [x = -60, y = -12, z = 0], at the right STG [x = 62, y = 0, z = -4].

STG = superior temporal gyrus; L = left; R = right.

For the conjunction of groups (control group ∩ SSD group) and sessions (ses-pre ∩ ses-post) over the verbal contrast (SG > G), significant activation was found bilaterally in a cluster with a peak in the superior temporal gyrus.

(II) Effects of gesture: *(control group ses-pre (SG > S)) ∩ (SSD group ses-pre (SG > S)) ∩ (SSD group ses-post (SG > S))*

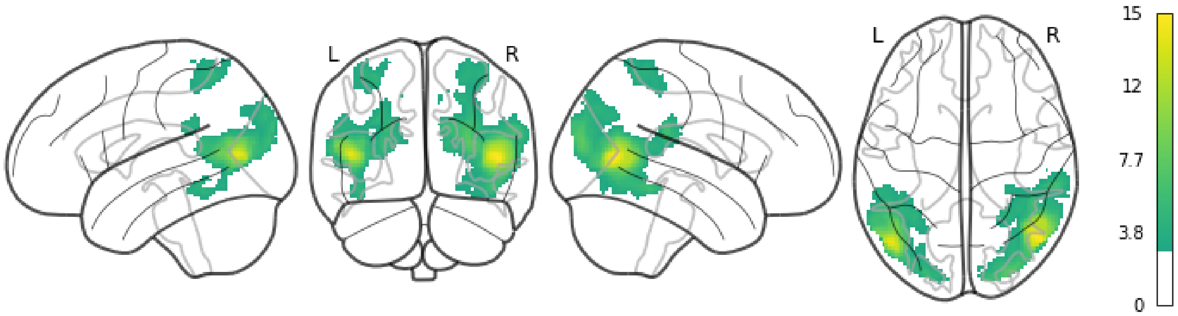


Figure 24: Glass brain visualization of neural activation for gestural contrast (SG > S) in integration (conjunction 0, minimum t-statistics) of groups and sessions. Activation clusters were thresholded at $p < 0.01$, with a minimum cluster size of 221 voxels. Crosshair points at the right MTG [x = 48, y = -66, z = 6], at the left MTG [x = -48, y = -70, z = 8], left SPG [x = -32, y = -54, z = 60] and at the right SPG [x = 28, y = -52, z = 58].

MTG = middle temporal gyrus; SPG = superior parietal gyrus; L = left; R = right.

For the conjunction of groups (control group ∩ SSD group) and sessions (ses-pre ∩ ses-post) over the gestural contrast (SG > S), significant activation was found bilaterally in a cluster extending from middle temporal regions to parietal and occipital gyrus.

(III) Effects of abstractness: *(control group ses-pre (abs > con))* ∩ *(SSD group ses-pre (abs > con))* ∩ *(SSD group ses-post (abs > con))*

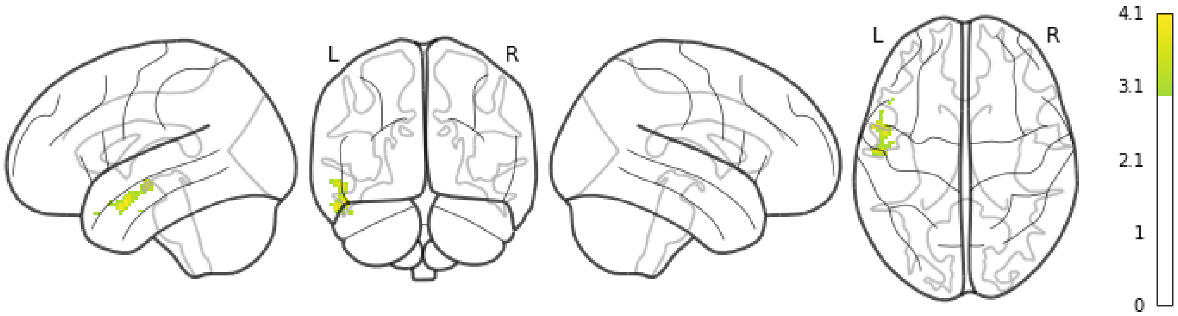


Figure 25: Glass brain visualization of neural activation for abstractness contrast (abs > con) in integration (conjunction 0, minimum t-statistics) of groups and sessions. Activation clusters were thresholded at $p < 0.01$, with a minimum cluster size of 221 voxels. Crosshair point at the left MTG [x = -52, y = -12, z = -10].

MTG = middle temporal gyrus; L = left; R = right.

For the conjunction of groups (control group ∩ SSD group) and sessions (ses-pre ∩ ses-post) over the abstractness contrast (abs > con), significant activation was found in left middle temporal gyrus.

Within-Subjects Effects

To get an overview of the single *abstractness effects in the multimodal condition* in the individual groups at the different measurement sessions, simple within-subjects effects were calculated.

The within-subjects effect of abstractness in the control group pre MSG training in the multimodal condition showed significant activation in left precentral and temporal regions as well as in right middle cingulate gyrus (*contrast (IV)*, see *figure 26*). As outlined above, a left lateralized network of neural activation, particularly in temporal and frontal regions, has been reported for the processing of abstract speech in healthy individuals (A. M. Rapp et al., 2012). Although our sample in the control group is small, we can already find similar activation patterns in the multimodal processing of abstractness.

Unlike in the control group, we found increased activation in SSD patients in several frontal, temporal and parietal, postcentral and occipital regions as well as in cuneus and thalamus and in parts of the cerebellum for the processing of abstract compared to concrete speech and gesture pre MSG training (*contrast (V)*, see *figure 27*). Hyperactivation in general is reported in previous studies for dynamic social, emotional and attentional gesture perception (Wible, 2012) and at least partially being associated with formal though disorders in general (Cavelti et al., 2018). Thus, the hyperactivation we also in our subjects might reflect an increased integration effort for abstract content in patients with schizophrenia (Green et al., 2009; Straube, Green, et al., 2013; Willems et al., 2007). Post MSG training (*contrast (VI)*, see *figure 28*), the patients show even higher activation in several clusters, but a particular increase in the left temporal lobe.

(IV) control group ses-pre in SG condition, effect of abstractness: *absSG > conSG*

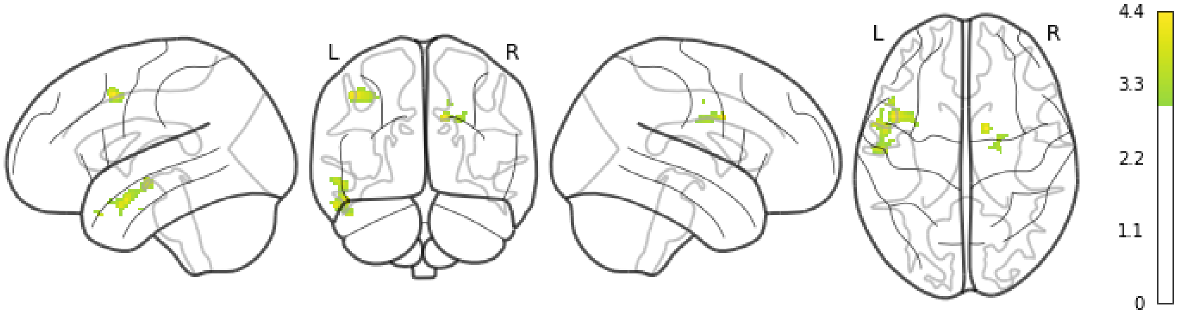


Figure 26: Glass brain visualization of neural activation for the effect of abstractness in multimodal condition (absSG > conSG) in the control group's first fMRI measurement (ses-pre). Activation clusters were thresholded at $p < 0.01$, with a minimum cluster size of 221 voxels. Crosshair points at the left PreCG [x = -44, y = 10, z = 46], right MCC [x = 14, y = 2, y = 32], left MTG [x = -50, y = 2, z = -22].

PreCG = precentral gyrus; MCC = middle cingulate and paracingulate gyri; MTG = middle temporal gyrus; L = left; R = right.

For the control group in ses-pre, abstractness contrast (absSG > conSG), significant activation was found in several clusters, including left frontal to precentral and temporal regions and a right hemispheric cluster in middle cingulate gyrus.

(V) SSD group ses-pre in SG condition, effect of abstractness: *absSG* > *conSG*

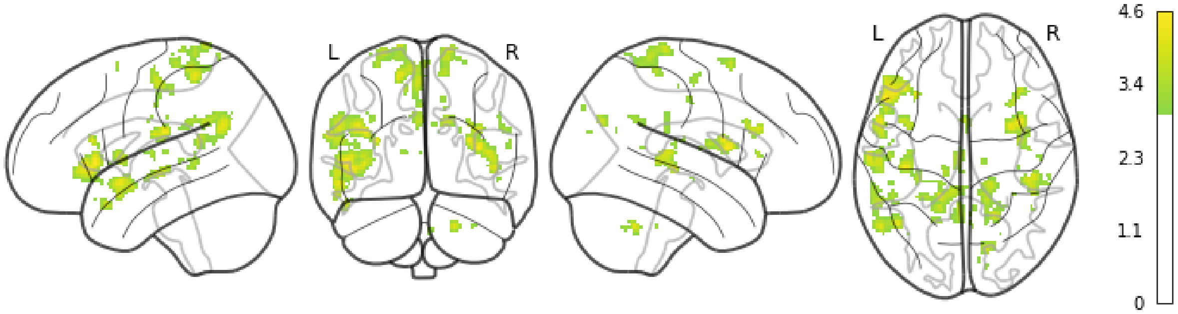


Figure 27: Glass brain visualization of neural activation for the effect of abstractness in multimodal condition (absSG > conSG) in the SSD group's first fMRI measurement (ses-pre). Activation clusters were thresholded at $p < 0.01$, with a minimum cluster size of 221 voxels. Crosshair points at the left IFGtriang [x = -50, y = 22, z = 2], right INS [x = 36, y = 6, z = 12], left PCUN [x = -16, y = -42, y = 58], right CER [x = 20, y = -54, z = -38], right STG [x = 46, y = -30, z = -2], right CUN [x = 14, y = -76, z = 28] and left Thalamus [x = -14, y = -24, z = 8].

IFGtriang = triangular part of inferior frontal gyrus; INS = insula; PCUN = precuneus; CER = cerebellum; STG = superior temporal gyrus; CUN = cuneus; L = left; R = right.

For the SSD group in ses-pre, abstractness contrast (absSG > conSG), significant activation was found in large clusters covering bilateral frontal and temporal regions and insula, postcentral regions, precuneus, cuneus, thalamus and parts of the cerebellum.

(VI) SSD group ses-post in SG condition, effect of abstractness: *absSG > conSG*

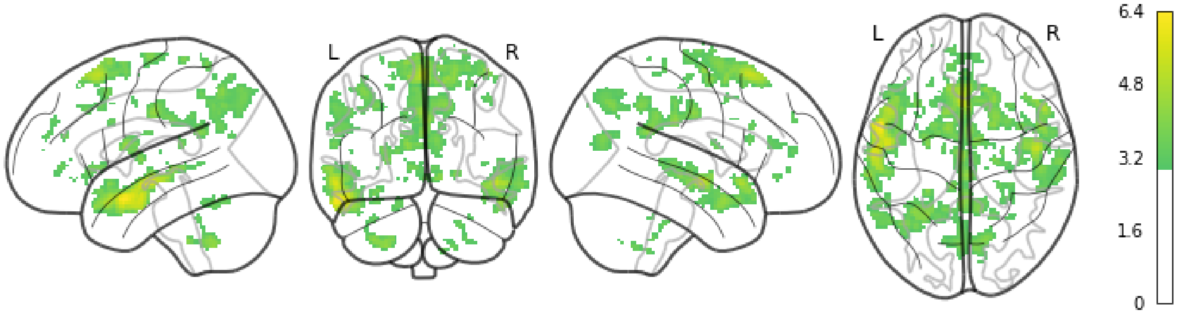


Figure 28: Glass brain visualization of neural activation for the effect of abstractness in multimodal condition (absSG > conSG) in the SSD group's after training fMRI measurement (ses-post). Activation clusters were thresholded at p < 0.01, with a minimum cluster size of 221 voxels. Crosshair points at the left MTG [x = -54, y = 4, z = -22], left SMA [x = 0, y = 20, z = 58], right STG [x = 48, y = -10, y = -10], left Caudate [x = -10, y = 2, z = 14], left C ER [x = -32, y = -52, z = -48], right Thalamus [x = 10, y = -4, z = -6], left MFG [x = -40, y = 14, z = 46], left ACC [x = -8, y = 44, z = 0], left CER [x = -30, y = -60, -22] and right CER [x = 12, y = -46, z = -54].

MTG = middle temporal gyrus; SMA = supplementary motor area; STG = superior temporal gyrus; CER = cerebelum; MFG = middle frontal gyrus; ACC = anterior cingulate and paracingulate gyri; L = left; R = right.

For the SSD in ses-post, abstractness contrast (absSG > conSG), significant activation was found in a large cluster over the whole brain with a peak activation in left middle temporal regions.

Study Data and Training Material Availability

Anonymized behavioral and clinical data as well as MSG training material and detailed information about our assessments are freely available under CCBy Attribution 4.0 International license on the Open Science Framework: DOI 10.17605/OSF.IO/UH4F9.

Lebenslauf

Die Seite 93 (Lebenslauf) enthält persönliche Daten. Sie ist deshalb nicht Bestandteil der Online-Veröffentlichung.

Verzeichnis der akademischen Lehrer/-innen

Meine akademischen Lehrenden waren

in Jena:
Honegger, T.
Schwarz, M.

in Leipzig:
Obleser, J.
Scharinger, M.

in Marburg:
Albert, R.
Alday, P.
Bachmann
Becker
Behr
Berger, R.
Bornkessel-Schlesewsky, I.
Carl
Depboylu
Dodel
Domahs, F.
Domahs, U.
Dulmage, D.
Eggert
Heyde-Zybatow, T.
Fleischer, W.
Franz, K.
Frobel, F.
Girnth, H.
Grond
Haberhausen
Handke
Hannken-Iljes, K.
Heilmann, C.
Hirsch
Hofer, M.
Holzapfel, A.
Hundertmark, L.
Kamp-Becker
Kappus
Kasper, S.
Kauschke, C.
König
Könitz, F.
Kuhnt
Künzel, H.
Lachnit, H.
Leibrecht, O.
Losekant
Lüders
Mayer
Mayer-Königsbüscher, J.
Mylius
Nimsky
Oertel
Pagenstecher
Pinquart, M.
Quaschner
Ries
Rohrmann, E.
Rosenow
Rüdinger, S.
Sassenhagen, J.
Schäfer, L.
Schallert, O.
Schmidt, E.
Schubö, A.
Sniehotta
Straube, B.
Strik
Tackenberg
Teepker
Theisen
Tillmann, C.
Tune, S.
Vill-Debney, D.
Walter, P.
Westermann, R.
Wiese, R.
Wright
Zingel, C.

in Nijmegen:
Levinson, S.
Senft, G.

in Zürich:
Clematide
Dellwo, V.
Volk

Autorenbeteiligungen an den Studien

Study I: Riedl, L., Nagels, A., Sammer, G., & Straube, B. (2020). A multimodal speech-gesture training intervention for patients with schizophrenia and its neural underpinnings–the study protocol of a randomized controlled pilot trial. *Frontiers in psychiatry*, *11*, 110. DOI: 10.3389/fpsyt.2020.00110.

Projektadministration	**LR**
Konzeptionalisierung	BS, AN, GS, **LR**
Manuskriptentwurf	**LR**
Manuskriptüberarbeitung	**LR**, BS, AN, GS
Forschungsmittelakquisition	BS, AN, GS
Betreuung	BS

Study II: Riedl, L., Nagels, A., Sammer, G., Choudhury, M., Nonnenmann, A., Sütterlin, A., Feise, C., Haslach, M., Bitsch, F. & Straube, B. (2021). A Novel Multimodal Speech-Gesture Training and Its Impact on Quality of Life and Neural Processing in Patients With Schizophrenia Spectrum Disorder. A Pilot Randomized Controlled Trial. PsyArXiv. DOI: 10.31234/osf.io/a8wn4

Projektadministration	**LR**
Konzeptionalisierung	BS, AN, GS, **LR**
Rekrutierung	**LR**, AN, MC, AS, CF, MH, FB
Datenerhebung	**LR**, AN, MC, AS, CF, MH, FB
Datenanalyse	**LR**
Dateninterpretation	**LR**, BS
Manuskriptentwurf	**LR**
Manuskriptüberarbeitung	**LR**, BS, AN, GS, MC, FB
Forschungsmittelakquisition	BS, AN, GS
Betreuung	BS

Weitere Beteiligungen an zitierten Artikeln

Choudhury, M., Steines, M., Nagels, A., **Riedl, L.**, Kircher, T., & Straube, B. (2021). Neural Basis of Speech-Gesture Mismatch Detection in Schizophrenia Spectrum Disorders. *Schizophrenia Bulletin*, *sbab059*. https://doi.org/10.1093/schbul/sbab059

Gau, R., Noble, S., … **Riedl, L.**, ... (2021). Brainhack: Developing a culture of open, inclusive, community-driven neuroscience. *Neuron*. https://doi.org/10.1016/j.neuron.2021.04.001

Danksagung

Die Seiten 97-98 (Danksagung) enthalten persönliche Daten. Sie sind deshalb nicht Bestandteil der Online-Veröffentlichung.

Ehrenwörtliche Erklärung

Die Seite 99 (Ehrenwörtliche Erklärung) enthält persönliche Daten. Sie ist deshalb nicht Bestandteil der Online-Veröffentlichung.

References

Allen, D. J., Antonitis, J. J., & Magaro, P. A. (1978). Reinforcing effects of prerecorded words and delayed speech feedback on the verbal behavior of a neologistic schizophrenic. *Perceptual and Motor Skills*, *46*(2), 343–346. https://doi.org/10.2466/pms.1978.46.2.343

Andreasen, N. C. (1983). The scale for the assessment of negative symptoms (SANS). University of Iowa Press.

Andreasen, N. C. (1984). The scale for the assessment of positive symptoms (SAPS). University of Iowa Press.

Andric, M., & Small, S. L. (2012). Gesture's Neural Language. *Frontiers in Psychology*, *99*(2), https://doi.org/1–12. 10.3389/fpsyg.2012.00099

Annen, S., Roser, P., & Brüne, M. (2012). Nonverbal Behavior During Clinical Interviews: Similarities and Dissimilarities Among Schizophrenia, Mania, and Depression. *The Journal of Nervous and Mental Disease*, *200*(1), 26–32. https://doi.org/10.1097/NMD.0b013e31823e653b

Ardila, A., Bernal, B., & Rosselli, M. (2016). How Localized are Language Brain Areas? A Review of Brodmann Areas Involvement in Oral Language. *Archives of Clinical Neuropsychology*, *31*(1), 112–122. https://doi.org/10.1093/arclin/acv081

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*(1), 26–41. https://doi.org/10.1016/j.media.2007.06.004

Baker, R. (1971). The use of operant conditioning to reinstate speech in mute schizophrenics. *Behaviour Research and Therapy*, *9*(4), 329–336. https://doi.org/10.1016/0005-7967(71)90045-3

Bambini, V., Arcara, G., Bechi, M., Buonocore, M., Cavallaro, R., & Bosia, M. (2016). The communicative impairment as a core feature of schizophrenia:

Frequency of pragmatic deficit, cognitive substrates, and relation with quality of life. *Comprehensive Psychiatry*, *71*, 106–120. https://doi.org/10.1016/j.comppsych.2016.08.012

Bavelas, J. B., Gerwing, G., Allison, M., & Sutton, C. (2011). Dyadic evidence for grounding with abstract deictic gestures. In G. Stam & M. Ishino (Eds.), *Integrating gestures: The interdisciplinary nature of gesture* (pp. 49–60). John Benjamins Publishing Company.

Beaty, R. E., Silvia, P. J., & Benedek, M. (2017). Brain networks underlying novel metaphor production. *Brain and Cognition*, *111*, 163–170. https://doi.org/10.1016/j.bandc.2016.12.004

Beeman, M. (1998). Coarse semantic coding and discourse comprehension. In C. Chiarello & M. Beeman (Eds.), *Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience* (pp. 255–284). Lawrence Erlbaum Associates, Inc., Publishers.

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, *37*(1), 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042

Bell, V. (2017). Open science in mental health research. *The Lancet Psychiatry*, *4*(7), 525–526. https://doi.org/10.1016/S2215-0366(17)30244-4

Benedek, M., Beaty, R., Jauk, E., Koschutnig, K., Fink, A., Silvia, P. J., Dunst, B., & Neubauer, A. C. (2014). Creating metaphors: The neural basis of figurative language production. *NeuroImage*, *90*, 99–106. https://doi.org/10.1016/j.neuroimage.2013.12.046

Bergemann, N., Parzer, P., Jaggy, S., Auler, B., Mundt, C., & Maier-Braunleder, S. (2008). Estrogen and Comprehension of Metaphoric Speech in Women Suffering From Schizophrenia: Results of a Double-Blind, Placebo-Controlled Trial. *Schizophrenia Bulletin*, *34*(6), 1172–1181. https://doi.org/10.1093/schbul/sbm138

Berndl, K., Cranach, M. von, & Grüsser, O.-J. (1986). Impairment of perception and recognition of faces, mimic expression and gestures in schizophrenic patients. *European Archives of Psychiatry and Neurological Sciences*, *235*(5), 282–291. https://doi.org/10.1007/BF00515915

Berndl, K., Grüsser, O.-J., Martini, M., & Remschmidt, H. (1986). Comparative studies on recognition of faces, mimic and gestures in adolescent and middle-aged schizophrenic patients. *European Archives of Psychiatry and Neurological Sciences*, *236*(2), 123–130. https://doi.org/10.1007/BF00454022

Binder, J. R. (2017). Current Controversies on Wernicke's Area and its Role in Language. *Current Neurology and Neuroscience Reports*, *17*(8), 58. https://doi.org/10.1007/s11910-017-0764-8

Blairy, S., Neumann, A., Nutthals, F., Pierret, L., Collet, D., & Philippot, P. (2008). Improvements in Autobiographical Memory in Schizophrenia Patients after a Cognitive Intervention. *Psychopathology*, *41*(6), 388–396. https://doi.org/10.1159/000155217

Bor, J., Brunelin, J., d'Amato, T., Costes, N., Suaud-Chagny, M.-F., Saoud, M., & Poulet, E. (2011). How can cognitive remediation therapy modulate brain activations in schizophrenia? An fMRI study. *Psychiatry Research*, *192*(3), 160–166. https://doi.org/10.1016/j.pscychresns.2010.12.004

Bosco, F. M., Gabbatore, I., Gastaldo, L., & Sacco, K. (2016). Communicative-Pragmatic Treatment in Schizophrenia: A Pilot Study. *Frontiers in Psychology*, *0*. https://doi.org/10.3389/fpsyg.2016.00166

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The Word Frequency Effect. *Experimental Psychology*, *58*(5), 412–424. https://doi.org/10.1027/1618-3169/a000123

Bucci, S., Startup, M., Wynn, P., Baker, A., & Lewin, T. J. (2008). Referential delusions of communication and interpretations of gestures. *Psychiatry Research*, *158*(1), 27–34. https://doi.org/10.1016/j.psychres.2007.07.004

Buchain, P. C., Vizzotto, A. D. B., Henna Neto, J., & Elkis, H. (2003). Randomized controlled trial of occupational therapy in patients with treatment-resistant schizophrenia. *Revista Brasileira de Psiquiatria*, *25*(1), 26–30. https://doi.org/10.1590/S1516-44462003000100006

Castillejos, M. C., Martín-Pérez, C., & Moreno-Küstner, B. (2019). Incidence of psychotic disorders and its association with methodological issues. A systematic review and meta-analyses. *Schizophrenia Research*, *204*, 458–459. https://doi.org/10.1016/j.schres.2018.07.031

Cavelti, M., Kircher, T., Nagels, A., Strik, W., & Homan, P. (2018). Is formal thought disorder in schizophrenia related to structural and functional aberrations in the language network? A systematic review of neuroimaging findings. *Schizophrenia Research*, *199*, 2–16. https://doi.org/10.1016/j.schres.2018.02.051

Charlson, F. J., Ferrari, A. J., Santomauro, D. F., Diminic, S., Stockings, E., Scott, J. G., McGrath, J. J., & Whiteford, H. A. (2018). Global Epidemiology and Burden of Schizophrenia: Findings From the Global Burden of Disease Study 2016. *Schizophrenia Bulletin*, *44*(6), 1195–1203. https://doi.org/10.1093/schbul/sby058

Choudhury, M., Steines, M., Nagels, A., Riedl, L., Kircher, T., & Straube, B. (2021). Neural Basis of Speech-Gesture Mismatch Detection in Schizophrenia Spectrum Disorders. *Schizophrenia Bulletin*, *sbab059*. https://doi.org/10.1093/schbul/sbab059

Cienki, A. (2008). Why study metaphor and gesture? In A. Cienki & C. Müller (Eds.), *Metaphor and Gesture* (3rd ed.). John Benjamins Publishing Company.

Cirillo, M. A., & Seidman, L. J. (2003). Verbal Declarative Memory Dysfunction in Schizophrenia: From Clinical Assessment to Genetics and Brain Mechanisms. *Neuropsychology Review*, *13*(2), 43–77. https://doi.org/10.1023/A:1023870821631

Clegg, J., Brumfitt, S., Parks, R. W., & Woodruff, P. W. R. (2007). Speech and language therapy intervention in schizophrenia: A case study. *International Journal of Language & Communication Disorders*, *42*(S1), 81–101. https://doi.org/10.1080/13682820601171472

Covington, M. A., He, C., Brown, C., Naçi, L., McClain, J. T., Fjordbak, B. S., Semple, J., & Brown, J. (2005). Schizophrenia and the structure of language: The linguist's view. *Schizophrenia Research*, *77*(1), 85–98. https://doi.org/10.1016/j.schres.2005.01.016

Cox, R. W., & Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine*, *10*(4–5), 171–178. https://doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L

Crawford, M. J., & Patterson, S. (2007). Arts therapies for people with schizophrenia: An emerging evidence base. *Evidence-Based Mental Health*, *10*(3), 69. https://doi.org/10.1136/ebmh.10.3.69

Cuevas, P., He, Y., Steines, M., & Straube, B. (2021). The processing of semantic complexity and co-speech gestures in schizophrenia: A naturalistic, multimodal fMRI study. *BioRxiv*, 2021.05.18.444612. https://doi.org/10.1101/2021.05.18.444612

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

de Bonis, M., Epelbaum, C., Deffez, V., & Féline, A. (1997). The comprehension of metaphors in schizophrenia. *Psychopathology*, *30*(3), 149–154. https://doi.org/10.1159/000285041

DeLisi, L. E. (2001). Speech Disorder in Schizophrenia: Review of the Literature and Exploration of Its Relation to the Uniquely Human Capacity for Language. *Schizophrenia Bulletin*, *27*(3), 481–496. https://doi.org/10.1093/oxfordjournals.schbul.a006889

DeWitt, I., & Rauschecker, J. P. (2013). Wernicke's area revisited: Parallel streams and word processing. *Brain and Language*, *127*(2), 181–191. https://doi.org/10.1016/j.bandl.2013.09.014

Diagnosedaten der Patientinnen und Patienten in Krankenhäusern. (2008). [Krankenhausstatistik]. Statistisches Bundesamt.

Dick, A. S., Goldin-Meadow, S., Hasson, U., Skipper, J. I., & Small, S. L. (2009). Co-speech gestures influence neural activity in brain regions associated with processing semantic information. *Human Brain Mapping*, *30*(11), 3509–3526. https://doi.org/10.1002/hbm.20774

Dick, A. S., Goldin-Meadow, S., Solodkin, A., & Small, S. L. (2012). Gesture in the developing brain. *Developmental Science*, *15*(2), 165–180. https://doi.org/10.1111/j.1467-7687.2011.01100.x

Dick, A. S., Mok, E. H., Beharelle, A. R., Goldin-Meadow, S., & Small, S. L. (2014). Frontal and temporal contributions to understanding the iconic co-speech gestures that accompany speech. *Human Brain Mapping*, *35*(3), 900–917.

DlexDB. (n.d.). Retrieved September 27, 2021, from http://www.dlexdb.de/

Dollfus, S., & Petit, M. (1995). Negative symptoms in schizophrenia: Their evolution during an acute phase. *Schizophrenia Research*, *17*(2), 187–194. https://doi.org/10.1016/0920-9964(94)00087-O

Driskell, J. E., & Radtke, P. H. (2003). The Effect of Gesture on Speech Production and Comprehension. *Human Factors*, *45*(3), 445–454. https://doi.org/10.1518/hfes.45.3.445.27258

Eack, S. M., & Newhill, C. E. (2007). Psychiatric Symptoms and Quality of Life in Schizophrenia: A Meta-Analysis. *Schizophrenia Bulletin*, *33*(5), 1225–1237. https://doi.org/10.1093/schbul/sbl071

Edwards, B. G., Barch, D. M., & Braver, T. S. (2010). Improving prefrontal cortex function in schizophrenia through focused training of cognitive control. *Frontiers in Human Neuroscience*, *4*. https://doi.org/10.3389/fnhum.2010.00032

Eigsti, I.-M. (2013). A Review of Embodiment in Autism Spectrum Disorders. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00224

Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J. (2017). MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PLOS ONE*, *12*(9), e0184661. https://doi.org/10.1371/journal.pone.0184661

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2018). fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*, *16*(1), 111–116. https://doi.org/10.1038/s41592-018-0235-4

Esteban, O., Markiewicz, C. J., Burns, C., Goncalves, M., Jarecka, D., Ziegler, E., Berleant, S., Ellis, D. G., Pinsard, B., Madison, C., Waskom, M., Notter, M. P., Clark, D., Manhães-Savio, A., Clark, D., Jordan, K., Dayan, M., Halchenko, Y. O., Loney, F., … Ghosh, S. (2020). *nipy/nipype: 1.5.1*. Zenodo. https://doi.org/10.5281/zenodo.4035081

Esteban, O., Markiewicz, C. J., Goncalves, M., DuPre, E., Kent, J. D., Salo, T., Ciric, R., Pinsard, B., Blair, R. W., Poldrack, R. A., & Gorgolewski, K. J. (2020). *fMRIPrep: A robust preprocessing pipeline for functional MRI*. Zenodo. https://doi.org/10.5281/zenodo.4252786

Evans, A. C., Janke, A. L., Collins, D. L., & Baillet, S. (2012). Brain templates and atlases. *NeuroImage*, *62*(2), 911–922. https://doi.org/10.1016/j.neuroimage.2012.01.024

Falkai, P. (Ed.). (2016). *Praxishandbuch Schizophrenie – Diagnostik Therapie Versorgungsstrukturen* (1st ed., Vol. 84). Thieme.

Falkai, P., Malchow, B., & Schmitt, A. (2017). Aerobic exercise and its effects on cognition in schizophrenia. *Current Opinion in Psychiatry*, *30*(3), 171–175. https://doi.org/10.1097/YCO.0000000000000326

Fiquer, J. T., Moreno, R. A., Brunoni, A. R., Barros, V. B., Fernandes, F., & Gorenstein, C. (2018). What is the nonverbal communication of depression? Assessing expressive differences between depressive patients and healthy volunteers during clinical interviews. *Journal of Affective Disorders*, *238*, 636–644. https://doi.org/10.1016/j.jad.2018.05.071

Fonov, V., Evans, A., McKinstry, R., Almli, C., & Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, *47*, S102. https://doi.org/10.1016/S1053-8119(09)70884-5

Fourie, E., Palser, E. R., Pokorny, J. J., Neff, M., & Rivera, S. M. (2020). Neural Processing and Production of Gesture in Children and Adolescents With Autism Spectrum Disorder. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.03045

Foxx, R. M., McMorrow, M. J., Davis, L. A., & Bittle, R. G. (1988). Replacing a chronic schizophrenic man’s delusional speech with stimulus appropriate responses. *Journal of Behavior Therapy and Experimental Psychiatry*, *19*(1), 43–50. https://doi.org/10.1016/0005-7916(88)90009-2

Frisk, V., & Milner, B. (1990). The role of the left hippocampal region in the acquisition and retention of story content. *Neuropsychologia*, *28*(4), 349–359. https://doi.org/10.1016/0028-3932(90)90061-R

Friston, K. J., Holmes, A. P., Poline, J.-B., Grasby, P. J., Williams, S. C. R., Frackowiak, R. S. J., & Turner, R. (1995). Analysis of fMRI Time-Series Revisited. *NeuroImage*, *2*(1), 45–53. https://doi.org/10.1006/nimg.1995.1007

Gaebel, W., & Wölwer, W. (2010). Gesundheitsberichterstattung des Bundes Heft 50. *Berlin: Rober Koch Institut*.

Galuppi, A., Turola, M. C., Nanni, M. G., Mazzoni, P., & Grassi, L. (2010). Schizophrenia and quality of life: How important are symptoms and functioning? *International Journal of Mental Health Systems*, *4*(1), 31. https://doi.org/10.1186/1752-4458-4-31

Gau, R., Noble, S., Heuer, K., Bottenhorn, K. L., Bilgin, I. P., Yang, Y.-F., Huntenburg, J. M., Bayer, J. M. M., Bethlehem, R. A. I., Rhoads, S. A., Vogelbacher, C., Borghesani, V., Levitis, E., Wang, H.-T., Van Den Bossche, S., Kobeleva, X., Legarreta, J. H., Guay, S., Atay, S. M., … Zuo, X.-N. (2021). Brainhack: Developing a culture of open, inclusive, community-driven neuroscience. *Neuron*. https://doi.org/10.1016/j.neuron.2021.04.001

Gernsbacher, M. A., Tallent, K. A., & Bolliger, C. M. (1999). Disordered Discourse in Schizophrenia Described by the Structure Building Framework. *Discourse Studies*, *1*(3), 355–372. https://doi.org/10.1177/1461445699001003004

Geschwind, N. (1970). The Organization of Language and the Brain. *Science*, *170*(3961), 940–944.

Glaesmer, H., Grande, G., Braehler, E., & Roth, M. (2011). The German Version of the Satisfaction With Life Scale (SWLS). *European Journal of Psychological Assessment*, *27*(2), 127–132. https://doi.org/10.1027/1015-5759/a000058

Glauer, M., Häusig, S., Krüger, M., Betsch, T., Renkewitz, F., Sedlmeier, P., & Winkler, I. (2007). Typizitätsnormen für vertreter von 30 kategorien. *Neurolinguistik*, *21*, 21–31.

Gold, C., Solli, H. P., Krüger, V., & Lie, S. A. (2009). Dose–response relationship in music therapy for people with serious mental disorders: Systematic review and meta-analysis. *Clinical Psychology Review*, *29*(3), 193–207. https://doi.org/10.1016/j.cpr.2009.01.001

Goldin-Meadow, S. (2005). *Hearing gesture: How our hands help us think*. Harvard University Press.

Goldin-Meadow, S. (2017). Using our hands to change our minds. *WIREs Cognitive Science*, *8*(1–2), e1368. https://doi.org/10.1002/wcs.1368

Goldin-Meadow, S., & Alibali, M. W. (2013). Gesture's Role in Speaking, Learning, and Creating Language. *Annual Review of Psychology*, *64*(1), 257–283. https://doi.org/10.1146/annurev-psych-113011-143802

Gordon, A., Davis, P. J., Patterson, S., Pepping, C. A., Scott, J. G., Salter, K., & Connell, M. (2018). A randomized waitlist control community study of Social

Cognition and Interaction Training for people with schizophrenia. *British Journal of Clinical Psychology*, *57*(1), 116–130. https://doi.org/10.1111/bjc.12161

Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Frontiers in Neuroinformatics*, *5*. https://doi.org/10.3389/fninf.2011.00013

Gorgolewski, K. J., Alfaro-Almagro, F., Auer, T., Bellec, P., Capotă, M., Chakravarty, M. M., Churchill, N. W., Cohen, A. L., Craddock, R. C., Devenyi, G. A., Eklund, A., Esteban, O., Flandin, G., Ghosh, S. S., Guntupalli, J. S., Jenkinson, M., Keshavan, A., Kiar, G., Liem, F., … Poldrack, R. A. (2017). BIDS apps: Improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. *PLOS Computational Biology*, *13*(3), e1005209. https://doi.org/10.1371/journal.pcbi.1005209

Goss, J. (2011). *Speech-gesture mismatches and dynamic coordination in schizophrenia* [Ph.D., The University of Chicago]. http://search.proquest.com/docview/914436105

Green, A., Straube, B., Weis, S., Jansen, A., Willmes, K., Konrad, K., & Kircher, T. (2009). Neural integration of iconic and unrelated coverbal gestures: A functional MRI study. *Human Brain Mapping*, *30*(10), 3309–3324. https://doi.org/10.1002/hbm.20753

Green, A., Straube, B., Weis, S., & Kircher, T. (2008). Perception of Speech and Iconic Gestures in Patients with Schizophrenia. *Schizophrenia Research*, *102*(1–3), 85–86. https://doi.org/10.1016/S0920-9964(08)70260-3

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, *48*(1), 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060

Grice, H. P. (1975). Logic and conversation. In *Speech acts* (pp. 41–58). Brill.

Gupta, T., Osborne, K. J., & Mittal, V. A. (2021). Abnormal Gesture Perception and Clinical High-Risk for Psychosis. *Schizophrenia Bulletin*, *47*(4), 938–947. https://doi.org/10.1093/schbul/sbab056

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences*, *9*(9), 416–423. https://doi.org/10.1016/j.tics.2005.07.004

Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Frontiers in Psychology*, *4*, 416. https://doi.org/10.3389/fpsyg.2013.00416

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (4th ed., pp. 819–836). MIT Press.

Hanlon, F. M., Houck, J. M., Pyeatt, C. J., Lundy, S. L., Euler, M. J., Weisend, M. P., Thoma, R. J., Bustillo, J. R., Miller, G. A., & Tesche, C. D. (2011). Bilateral hippocampal dysfunction in schizophrenia. *NeuroImage*, *58*(4), 1158–1168. https://doi.org/10.1016/j.neuroimage.2011.06.091

Harrison, P. J. (2004). The hippocampus in schizophrenia: A review of the neuropathological evidence and its pathophysiological implications. *Psychopharmacology*, *174*(1), 151–162. https://doi.org/10.1007/s00213-003-1761-y

Haut, K. M., Lim, K. O., & MacDonald Iii, A. (2010). Prefrontal Cortical Changes Following Cognitive Training in Patients with Chronic Schizophrenia: Effects of Practice, Generalization, and Specificity. *Neuropsychopharmacology*, *35*(9), 1850–1859. https://doi.org/10.1038/npp.2010.52

He, Y., Gebhardt, H., Steines, M., Sammer, G., Kircher, T., Nagels, A., & Straube, B. (2015). The EEG and fMRI signatures of neural integration: An investigation of meaningful gestures and corresponding speech. *Neuropsychologia*, *72*, 27–42. https://doi.org/10.1016/j.neuropsychologia.2015.04.018

He, Y., Steines, M., Sammer, G., Nagels, A., Kircher, T., & Straube, B. (2021). Modality-specific dysfunctional neural processing of social-abstract and non-social-concrete information in schizophrenia. *NeuroImage: Clinical*, *29*, 102568. https://doi.org/10.1016/j.nicl.2021.102568

He, Y., Steines, M., Sommer, J., Gebhardt, H., Nagels, A., Sammer, G., Kircher, T., & Straube, B. (2018). Spatial–temporal dynamics of gesture–speech integration: A simultaneous EEG-fMRI study. *Brain Structure and Function*, *223*, 3073–3089. https://doi.org/10.1007/s00429-018-1674-5

Hegarty, J. D., Baldessarini, R. J., Tohen, M., Waternaux, C., & Oepen, G. (1994). One hundred years of schizophrenia: A meta-analysis of the outcome literature. *American Journal of Psychiatry*, *151*(10), 1409–1416.

Heim, S. (2020). Kognitive Kommunikationsstörungen bei Schizophrenie und ihre Relevanz für die Logopädie/Sprachtherapie. *Sprache · Stimme · Gehör*. https://doi.org/10.1055/a-1043-7838

Heim, S., Dehmer, M., & Berger-Tunkel, M. (2019). Beeinträchtigungen von Sprache und Kommunikation bei Schizophrenie. *Der Nervenarzt*, *90*(5), 485–489. https://doi.org/10.1007/s00115-018-0647-5

Herholz, P. (n.d.). *BIDSonym*. https://peerherholz.github.io/BIDSonym/

Herholz, P., Ludwig, R. M., & Poline, J.-B. (2021). *BIDSonym - A BIDSapp for the pseudo-anonymization of neuroimaging datasets*. PsyArXiv. https://doi.org/10.31234/osf.io/3aknq

Hinzen, W., & Rosselló, J. (2015). The linguistics of schizophrenia: Thought disturbance as language pathology across positive symptoms. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00971

Holle, H., & Gunter, T. C. (2007). The Role of Iconic Gestures in Speech Disambiguation: ERP Evidence. *Journal of Cognitive Neuroscience*, *19*(7), 1175–1192. https://doi.org/10.1162/jocn.2007.19.7.1175

Holler, J., & Levinson, S. C. (2019). Multimodal Language Processing in Human Communication. *Trends in Cognitive Sciences*, *23*(8), 639–652. https://doi.org/10.1016/j.tics.2019.05.006

Hooker, C. I., Bruce, L., Fisher, M., Verosky, S. C., Miyakawa, A., & Vinogradov, S. (2012). Neural activity during emotion recognition after combined cognitive plus social cognitive training in schizophrenia. *Schizophrenia Research*, *139*(1), 53–59. https://doi.org/10.1016/j.schres.2012.05.009

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, *137*(2), 297–315. https://doi.org/10.1037/a0022128

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–effect relationship in psychotherapy. *American Psychologist*, *41*(2), 159.

Hubbard, A. L., McNealy, K., Zeeland, A. A. S.-V., Callan, D. E., Bookheimer, S. Y., & Dapretto, M. (2012). Altered integration of speech and gesture in children with autism spectrum disorders. *Brain and Behavior*, *2*(5), 606–619. https://doi.org/10.1002/brb3.81

Huppert, J. D., Weiss, K. A., Lim, R., Pratt, S., & Smith, T. E. (2001). Quality of life in schizophrenia: Contributions of anxiety and depression. *Schizophrenia Research*, *51*(2), 171–180. https://doi.org/10.1016/S0920-9964(99)00151-6

Iakimova, G., Passerieux, C., Denhière, G., Laurent, J.-P., Vistoli, D., Vilain, J., & Hardy-Baylé, M.-C. (2010). The influence of idiomatic salience during the comprehension of ambiguous idioms by patients with schizophrenia. *Psychiatry Research*, *177*(1–2), 46–54. https://doi.org/10.1016/j.psychres.2010.02.005

Jääskeläinen, E., Juola, P., Hirvonen, N., McGrath, J. J., Saha, S., Isohanni, M., Veijola, J., & Miettunen, J. (2013). A Systematic Review and Meta-Analysis of Recovery in Schizophrenia. *Schizophrenia Bulletin*, *39*(6), 1296–1306. https://doi.org/10.1093/schbul/sbs130

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage*, *17*(2), 825–841. https://doi.org/10.1006/nimg.2002.1132

Jeong, B., Wible, C. G., Hashimoto, R.-I., & Kubicki, M. (2009). Functional and anatomical connectivity abnormalities in left inferior frontal gyrus in schizophrenia. *Human Brain Mapping*, *30*(12), 4138–4151.

Jones, P., Murray, R., Jones, P., Rodgers, B., & Marmot, M. (1994). Child developmental risk factors for adult schizophrenia in the British 1946 birth cohort. *The Lancet*, *344*(8934), 1398–1402. https://doi.org/10.1016/S0140-6736(94)90569-X

Joue, G., Boven, L., Willmes, K., Evola, V., Demenescu, L. R., Hassemer, J., Mittelberg, I., Mathiak, K., Schneider, F., & Habel, U. (2020). Metaphor processing is supramodal semantic processing: The role of the bilateral lateral temporal regions in multimodal communication. *Brain and Language*, *205*, 104772. https://doi.org/10.1016/j.bandl.2020.104772

Joyal, M., Bonneau, A., & Fecteau, S. (2016). Speech and language therapies to improve pragmatics and discourse skills in patients with schizophrenia. *Psychiatry Research*, *240*, 88–95. https://doi.org/10.1016/j.psychres.2016.04.010

Karakuła, H., Krukow, P., & Kalinowska, J. (2013). Disintegration of social cognitive processes in schizophrenia. *Psychiatria Polska*, *47*(6), 1023–1039.

Karow, A., Moritz, S., Lambert, M., Schoder, S., & Krausz, M. (2005). PANSS Syndromes and Quality of Life in Schizophrenia. *Psychopathology*, *38*(6), 320–326. https://doi.org/10.1159/000088921

Kawakubo, Y., Kamio, S., Nose, T., Iwanami, A., Nakagome, K., Fukuda, M., Kato, N., Rogers, M. A., & Kasai, K. (2007). Phonetic mismatch negativity predicts social skills acquisition in schizophrenia. *Psychiatry Research*, *152*(2–3), 261–265. https://doi.org/10.1016/j.psychres.2006.02.010

Kelly, A. M. C., & Garavan, H. (2005). Human Functional Neuroimaging of Brain Changes Associated with Practice. *Cerebral Cortex*, *15*(8), 1089–1102. https://doi.org/10.1093/cercor/bhi005

Kelly, S. D., Barr, D. J., Church, R. B., & Lynch, K. (1999). Offering a Hand to Pragmatic Understanding: The Role of Speech and Gesture in Comprehension and Memory. *Journal of Memory and Language*, *40*(4), 577–592.

Kelly, S. D., Özyürek, A., & Maris, E. (2010). Two Sides of the Same Coin: Speech and Gesture Mutually Interact to Enhance Comprehension. *Psychological Science*, *21*(2), 260–267. https://doi.org/10.1177/0956797609357327

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363. https://doi.org/10.1037/0033-295X.85.5.363

Kircher, T., & Gauggel, S. (2008). *Neuropsychologie der Schizophrenie: Symptome, Kognition, Gehirn* (T. Kircher, Ed.). Springer-Verlag Berlin Heidelberg.

Kircher, T., Leube, D., Erb, M., Grodd, W., & Rapp, A. (2007). Neural correlates of metaphor processing in schizophrenia. *NeuroImage*, *34*(1), 281–289. https://doi.org/10.1016/j.neuroimage.2006.08.044

Kircher, T., Straube, B., Leube, D., Weis, S., Sachs, O., Willmes, K., Konrad, K., & Green, A. (2009). Neural interaction of speech and gesture: Differential activations of metaphoric co-verbal gestures. *Neuropsychologia*, *47*(1), 169–179. https://doi.org/10.1016/j.neuropsychologia.2008.08.009

Kissling, W., Höffler, J., Seemann, U., Müller, P., Rüther, E., Trenckmann, U., Über, A., Schulenburg, J.-M. G. v d, Glaser, P., Glaser, T., Mast, O., & Schmidt, D. (1999). Die direkten und indirekten Kosten der Schizophrenie. *Fortschritte der Neurologie · Psychiatrie*, *67*(1), 29–36. https://doi.org/10.1055/s-2007-993735

Klein, A., Ghosh, S. S., Bao, F. S., Giard, J., Häme, Y., Stavsky, E., Lee, N., Rossa, B., Reuter, M., Neto, E. C., & Keshavan, A. (2017). Mindboggling morphometry of human brains. *PLOS Computational Biology*, *13*(2), e1005350. https://doi.org/10.1371/journal.pcbi.1005350

Kondel, T., Hirsch, S., & Laws, K. (2006). Name relearning in elderly patients with schizophrenia: Episodic and temporary, not semantic and permanent. *Cognitive Neuropsychiatry*, *11*(1), 1–12. https://doi.org/10.1080/13546800444000137

Kostova, M., Passerieux, C., Laurent, J. P., Saint-Georges, C., & Hardy-Baylé, M. C. (2003). Functional analysis of the deficit in semantic context processes in schizophrenic patients: An event-related potentials study. *Neurophysiologie Clinique/Clinical Neurophysiology*, *33*(1), 11–22. https://doi.org/10.1016/S0987-7053(03)00006-6

Krach, S., Blümel, I., Marjoram, D., Lataster, T., Krabbendam, L., Weber, J., van Os, J., & Kircher, T. (2009). Are women better mindreaders? Sex differences in neural correlates of mentalizing detected with functional MRI. *BMC Neuroscience*, *10*(1), 9. https://doi.org/10.1186/1471-2202-10-9

Kramer, S., Bryan, K., & Frith, C. D. (2001). Mental Illness and Communication. *International Journal of Language & Communication Disorders*, *36*(S1), 132–137. https://doi.org/10.3109/13682820109177872

Kuperberg, G., & Heckers, S. (2000). Schizophrenia and cognitive function. *Current Opinion in Neurobiology*, *10*(2), 205–210. https://doi.org/10.1016/S0959-4388(00)00068-4

Lakoff, G., & Johnson, M. (2008). *Metaphors We Live By.* University of Chicago Press.

Laliberte-Rudman, D., Yu, B., Scott, E., & Pajouhandeh, P. (2000). Exploration of the Perspectives of Persons With Schizophrenia Regarding Quality of Life. *American Journal of Occupational Therapy*, *54*(2), 137–147. https://doi.org/10.5014/ajot.54.2.137

Lanczos, C. (1964). Evaluation of Noisy Data. *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, *1*(1), 76–85. https://doi.org/10.1137/0701007

Landre, N. A., Taylor, M. A., & Kearns, K. P. (1992). Language Functioning in Schizophrenic and Aphasic Patients. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology*, *5*(1), 7–14.

Langdon, R., Coltheart, M., Ward, P. B., & Catts, S. V. (2002). Disturbed communication in schizophrenia: The role of poor pragmatics and poor mind-reading. *Psychological Medicine*, *32*(7), 1273–1284. https://doi.org/10.1017/S0033291702006396

Lavelle, M., Healey, P. G. T., & McCabe, R. (2013). Is Nonverbal Communication Disrupted in Interactions Involving Patients With Schizophrenia? *Schizophrenia Bulletin*, *39*(5), 1150–1158. https://doi.org/10.1093/schbul/sbs091

Lavelle, M., Healey, P. G. T., & McCabe, R. (2014). Nonverbal Behavior During Face-to-face Social Interaction in Schizophrenia: A Review. *The Journal of Nervous and Mental Disease*, *202*(1), 47–54. https://doi.org/10.1097/NMD.0000000000000031

Liberman, R. P., Wallace, C. J., Blackwell, G., Kopelowicz, A., Vaccaro, J. V., & Mintz, J. (1998). Skills Training Versus Psychosocial Occupational Therapy for Persons With Persistent Schizophrenia. *American Journal of Psychiatry*, *155*(8), 1087–1091. https://doi.org/10.1176/ajp.155.8.1087

MacNeil, L. K., & Mostofsky, S. H. (2012). Specificity of dyspraxia in children with autism. *Neuropsychology*, *26*(2), 165.

Man, D. W., Law, K. M., & Chung, R. C. (2012). Cognitive training for Hong Kong Chinese with schizophrenia in vocational rehabilitation. *Hong Kong Medical Journal = Xianggang Yi Xue Za Zhi*, *18*(S6), 18–22.

Maricchiolo, F., Gnisci, A., Bonaiuto, M., & Ficca, G. (2009). Effects of different types of hand gestures in persuasive speech on receivers’ evaluations. *Language and Cognitive Processes*, *24*(2), 239–266. https://doi.org/10.1080/01690960802159929

Marini, A., Spoletini, I., Rubino, I. A., Ciuffa, M., Bria, P., Martinotti, G., Banfi, G., Boccascino, R., Strom, P., Siracusano, A., Caltagirone, C., & Spalletta, G. (2008). The language of schizophrenia: An analysis of micro and macrolinguistic abilities and their neuropsychological correlates. *Schizophrenia Research*, *105*(1), 144–155. https://doi.org/10.1016/j.schres.2008.07.011

Martin, L. A. L., Koch, S. C., Hirjak, D., & Fuchs, T. (2016). Overcoming Disembodiment: The Effect of Movement Therapy on Negative Symptoms in Schizophrenia—A Multicenter Randomized Controlled Trial. *Frontiers in Psychology*, *7*, 483. https://doi.org/10.3389/fpsyg.2016.00483

Martin, P., Tewesmeier, M., Albers, M., Schmid, G., & Scharfetter, C. (1994). Investigation of gestural and pantomime performance in chronic schizophrenic inpatients. *European Archives of Psychiatry and Clinical Neuroscience*, *244*(2), 59–64. https://doi.org/10.1007/BF02193520

MATLAB (Version R2017b). (2017). [Computer software]. The MathWorks, Inc. https://de.mathworks.com/

Matthews, N., Gold, B. J., Sekuler, R., & Park, S. (2013). Gesture imitation in schizophrenia. *Schizophrenia Bulletin*, *39*(1), 94–101. https://doi.org/10.1093/schbul/sbr062

McCabe, R., Healey, P. G. T., Priebe, S., Lavelle, M., Dodwell, D., Laugharne, R., Snell, A., & Bremner, S. (2013). Shared understanding in psychiatrist–patient communication: Association with treatment adherence in schizophrenia. *Patient Education and Counseling*, *93*(1), 73–79. https://doi.org/10.1016/j.pec.2013.05.015

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press.

Mittal, V. A., Tessner, K. D., McMillan, A. L., Delawalla, Z., Trotman, H. D., & Walker, E. F. (2006). Gesture behavior in unmedicated schizotypal adolescents. *Journal of Abnormal Psychology*, *115*(2), 351–358. https://doi.org/10.1037%2F0021-843X.115.2.351

Mondada, L., & Oloff, F. (2011). Gesture in overlap: The situated establishment of speakership. In G. Stam & M. Ishino (Eds.), *Integrating Gestures: The interdisciplinary nature of gesture* (Vol. 4, p. 321).

Morimoto, T., Matsuda, Y., Matsuoka, K., Yasuno, F., Ikebuchi, E., Kameda, H., Taoka, T., Miyasaka, T., Kichikawa, K., & Kishimoto, T. (2018). Computer-assisted cognitive remediation therapy increases hippocampal volume in patients with schizophrenia: A randomized controlled trial. *BMC Psychiatry*, *18*(1), 83. https://doi.org/10.1186/s12888-018-1667-1

Morsella, E., & Krauss, R. M. (2004). The Role of Gestures in Spatial Working Memory and Speech. *The American Journal of Psychology*, *117*(3), 411–424.

Nagels, A., Kircher, T., Grosvald, M., Steines, M., & Straube, B. (2019). Evidence for gesture-speech mismatch detection impairments in schizophrenia. *Psychiatry Research*, *273*, 15–21. https://doi.org/10.1016/j.psychres.2018.12.107

Narvaez, J. M., Twamley, E. W., McKibbin, C. L., Heaton, R. K., & Patterson, T. L. (2008). Subjective and objective quality of life in schizophrenia. *Schizophrenia Research*, *98*(1), 201–208. https://doi.org/10.1016/j.schres.2007.09.001

Nathan, M. J., & Alibali, M. W. (2011). How gesture use enables intersubjectivity in the classroom. In G. Stam & M. Ishino (Eds.), *Integrating gestures: The interdisciplinary nature of gesture* (pp. 257–266).

Ojeda, N., Peña, J., Sánchez, P., Bengoetxea, E., Elizagárate, E., Ezcurra, J., & Gutiérrez Fraile, M. (2012). Efficiency of cognitive rehabilitation with REHACOP in chronic treatment resistant Hispanic patients. *NeuroRehabilitation*, *30*(1), 65–74. https://doi.org/10.3233/NRE-2011-0728

Özyürek, A., Willems, R. M., Kita, S., & Hagoort, P. (2007). On-line Integration of Semantic Information from Speech and Gesture: Insights from Event-related Brain Potentials. *Journal of Cognitive Neuroscience*, *19*(4), 605–616. https://doi.org/10.1162/jocn.2007.19.4.605

Park, S., Matthews, N., & Gibson, C. (2008). Imitation, Simulation, and Schizophrenia. *Schizophrenia Bulletin*, *34*(4), 698–707. https://doi.org/10.1093/schbul/sbn048

Patil, I. (2021). Visualizations with statistical details: The “ggstatsplot” approach. *Journal of Open Source Software*, *6*(61), 3167. https://doi.org/10.21105/joss.03167

Pavot, W., & Diener, E. (2009). Review of the Satisfaction With Life Scale. In E. Diener (Ed.), *Assessing Well-Being: The Collected Works of Ed Diener* (pp. 101–117). Springer Netherlands. https://doi.org/10.1007/978-90-481-2354-4_5

Pecher, D., Boot, I., & Van Dantzig, S. (2011). Chapter seven - Abstract Concepts: Sensory-Motor Grounding, Metaphors, and Beyond. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 54, pp. 217–248). Academic Press. https://doi.org/10.1016/B978-0-12-385527-5.00007-3

Penadés, R., González-Rodríguez, A., Catalán, R., Segura, B., Bernardo, M., & Junqué, C. (2017). Neuroimaging studies of cognitive remediation in schizophrenia: A systematic and critical review. *World Journal of Psychiatry*, *7*(1), 34–43. https://doi.org/10.5498/wjp.v7.i1.34

Penadés, R., Pujol, N., Catalán, R., Massana, G., Rametti, G., García-Rizo, C., Bargalló, N., Gastó, C., Bernardo, M., & Junqué, C. (2013). Brain Effects of Cognitive Remediation Therapy in Schizophrenia: A Structural and Functional Neuroimaging Study. *Biological Psychiatry*, *73*(10), 1015–1023. https://doi.org/10.1016/j.biopsych.2013.01.017

Penadés, R., Segura, B., Inguanzo, A., García-Rizo, C., Catalán, R., Masana, G., Bernardo, M., & Junqué, C. (2020). Cognitive remediation and brain connectivity: A resting-state fMRI study in patients with schizophrenia. *Psychiatry Research: Neuroimaging*, *303*, 111140. https://doi.org/10.1016/j.pscychresns.2020.111140

Peng, S.-M., Koo, M., & Kuo, J.-C. (2010). Effect of Group Music Activity as an Adjunctive Therapy on Psychotic Symptoms in Patients With Acute Schizophrenia. *Archives of Psychiatric Nursing*, *24*(6), 429–434. https://doi.org/10.1016/j.apnu.2010.04.001

Pereira, G. M. (2015). *The interplay of gesture, gaze, and speech: Analyzing task-based interaction* [Ph.D., Universität des Saarlandes]. https://doi.org/doi:10.22028/D291-23650

Pine, K. J., Reeves, L., Howlett, N., & Fletcher, B. (C). (2013). Giving cognition a helping hand: The effect of congruent gestures on object name retrieval. *British Journal of Psychology*, *104*(1), 57–68.

Poldrack, R. A. (2000). Imaging Brain Plasticity: Conceptual and Methodological Issues— A Theoretical Review. *NeuroImage*, *12*(1), 1–13. https://doi.org/10.1006/nimg.2000.0596

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, *84*, 320–341. https://doi.org/10.1016/j.neuroimage.2013.08.048

Presentation (Software) Documentation. (n.d.). Retrieved March 29, 2017, from https://www.neurobs.com/presentation/docs

R Core Team. (2013). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. http://www.R-project.org/

Rapp, A. M., Mutschler, D. E., & Erb, M. (2012). Where in the brain is nonliteral language? A coordinate-based meta-analysis of functional magnetic resonance imaging studies. *NeuroImage*, *63*(1), 600–610. https://doi.org/10.1016/j.neuroimage.2012.06.022

Rapp, A., & Schmierer, P. (2010). Proverbs and nonliteral language in schizophrenia. A systematic methodological review of all studies published 1931-2010. *Schizophrenia Research*, *117*(2), 422. https://doi.org/10.1016/j.schres.2010.02.775

Redcay, E. (2008). The superior temporal sulcus performs a common function for social and speech perception: Implications for the emergence of autism.

Neuroscience & Biobehavioral Reviews, *32*(1), 123–142. https://doi.org/10.1016/j.neubiorev.2007.06.004

Reichenberg, A., Weiser, M., Rabinowitz, J., Caspi, A., Schmeidler, J., Mark, M., Kaplan, Z., & Davidson, M. (2002). A Population-Based Cohort Study of Premorbid Intellectual, Language, and Behavioral Functioning in Patients With Schizophrenia, Schizoaffective Disorder, and Nonpsychotic Bipolar Disorder. *American Journal of Psychiatry*, *159*(12), 2027–2035.

Riedl, L. (2021). *Fmriprep2spm* [MATLAB]. https://github.com/LydiaRiedl/fmriprep2spm

Riedl, L., Nagels, A., Sammer, G., & Straube, B. (2020). A Multimodal Speech-Gesture Training Intervention for Patients With Schizophrenia and Its Neural Underpinnings – the Study Protocol of a Randomized Controlled Pilot Trial. *Frontiers in Psychiatry*, *11*, 110. https://doi.org/10.3389/fpsyt.2020.00110

Rolls, E. T., Joliot, M., & Tzourio-Mazoyer, N. (2015). Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. *NeuroImage*, *122*, 1–5. https://doi.org/10.1016/j.neuroimage.2015.07.075

Rosenkranz, A., Kircher, T., & Nagels, A. (2019). Neuropsychological correlates underlying verbal fluency deficits in schizophrenia: The role of attention and executive function. *Interdisciplinary Linguistics and Psychiatric Research on Language Disorders*, 45–54.

Rossetti, I., Brambilla, P., & Papagno, C. (2018). Metaphor Comprehension in Schizophrenic Patients. *Frontiers in Psychology*, *9*, 670. https://doi.org/10.3389/fpsyg.2018.00670

RStudio, P. (2021). *RStudio: Integrated Development Environment for R* (1.4.1106.5) [Computer software]. http://www.rstudio.com/

Rudner, M. (2018). Working Memory for Linguistic and Non-linguistic Manual Gestures: Evidence, Theory, and Application. *Frontiers in Psychology*, *9*, 679. https://doi.org/10.3389/fpsyg.2018.00679

Saha, S., Chant, D., Welham, J., & McGrath, J. (2005). A Systematic Review of the Prevalence of Schizophrenia. *PLOS Medicine*, *2*(5), e141. https://doi.org/10.1371/journal.pmed.0020141

Santana, E., & De Vega, M. (2011). Metaphors are Embodied, and so are Their Literal Counterparts. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00090

Santos, A. E. dos, Pedrão, L. J., Zamberlan-Amorim, N. E., Furlan, J. D. L., & Carvalho, A. M. P. (2021). Effectiveness of the speech language intervention on the communicative behavior in groups of individuals diagnosed with schizophrenia. *CoDAS*, *33*(4). https://doi.org/10.1590/2317-1782/20202020088

Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughead, J., Calkins, M. E., Eickhoff, S. B., Hakonarson, H., Gur, R. C., Gur, R. E., & Wolf, D. H. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *NeuroImage*, *64*, 240–256. https://doi.org/10.1016/j.neuroimage.2012.08.052

Saykin, A. J., Gur, R. C., Gur, R. E., Mozley, P. D., Mozley, L. H., Resnick, S. M., Kester, D. B., & Stafiniak, P. (1991). Neuropsychological Function in Schizophrenia: Selective Impairment in Memory and Learning. *Archives of General Psychiatry*, *48*(7), 618–624. https://doi.org/10.1001/archpsyc.1991.01810310036007

Saykin, A. J., Shtasel, D. L., Gur, R. E., Kester, D. B., Mozley, L. H., Stafiniak, P., & Gur, R. C. (1994). Neuropsychological Deficits in Neuroleptic Naive Patients With First-Episode Schizophrenia. *Archives of General Psychiatry*, *51*(2), 124–131. https://doi.org/10.1001/archpsyc.1994.03950020048005

Schülke, R., & Straube, B. (2019). Transcranial Direct Current Stimulation Improves Semantic Speech–Gesture Matching in Patients With Schizophrenia Spectrum Disorder. *Schizophrenia Bulletin*, *45*(3), 522–530. https://doi.org/10.1093/schbul/sby144

Segrin, C. (1990). A meta-analytic review of social skill deficits in depression. *Communication Monographs*, *57*(4), 292–308. https://doi.org/10.1080/03637759009376204

Shenton, M. E., Dickey, C. C., Frumin, M., & McCarley, R. W. (2001). A review of MRI findings in schizophrenia. *Schizophrenia Research*, *49*(1), 1–52. https://doi.org/10.1016/S0920-9964(01)00163-3

Silverman, L. B., Bennetto, L., Campana, E., & Tanenhaus, M. K. (2010). Speech-and-gesture integration in high functioning autism. *Cognition*, *115*(3), 380–393. https://doi.org/10.1016/j.cognition.2010.01.002

Skipper, J. I., Goldin-Meadow, S., Nusbaum, H. C., & Small, S. L. (2007). Speech-associated gestures, Broca's area, and the human mirror system. *Brain and Language*, *101*(3), 260–277.

Slotnick, S. D. (2017). Cluster success: FMRI inferences for spatial extent have acceptable false-positive rates. *Cognitive Neuroscience*, *8*(3), 150–155. https://doi.org/10.1080/17588928.2017.1319350

Slotnick, S. D., Moo, L. R., Segal, J. B., & Hart, J. (2003). Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Cognitive Brain Research*, *17*(1), 75–82. https://doi.org/10.1016/S0926-6410(03)00082-X

Solanki, R. K., Singh, P., Midha, A., & Chugh, K. (2008). Schizophrenia: Impact on quality of life. *Indian Journal of Psychiatry*, *50*(3), 181–186. https://doi.org/10.4103/0019-5545.43632

Statistical Parametric Mapping (Version 12). (2018). [Computer software]. http://www.fil.ion.ucl.ac.uk/spm/

Steines, M., Nagels, A., Kircher, T., & Straube, B. (2021). The role of the left and right inferior frontal gyrus in processing metaphoric and unrelated co-speech gestures. *NeuroImage*, *237*, 118182. https://doi.org/10.1016/j.neuroimage.2021.118182

Stevenson, R. A., VanDerKlok, R. M., Pisoni, D. B., & James, T. W. (2011). Discrete neural substrates underlie complementary audiovisual speech integration processes. *NeuroImage*, *55*(3), 1339–1345. https://doi.org/10.1016/j.neuroimage.2010.12.063

Stieglitz Ham, H., Corley, M., Rajendran, G., Carletta, J., & Swanson, S. (2008). Brief Report: Imitation of Meaningless Gestures in Individuals with Asperger Syndrome and High-functioning Autism. *Journal of Autism and Developmental Disorders*, *38*(3), 569–573. https://doi.org/10.1007/s10803-007-0417-x

Straube, B., Green, A., Bromberger, B., & Kircher, T. (2011a). The differentiation of iconic and metaphoric gestures: Common and unique integration processes. *Human Brain Mapping*, *32*(4), 520–533. https://doi.org/10.1002/hbm.21041

Straube, B., Green, A., Bromberger, B., & Kircher, T. (2011b). The differentiation of iconic and metaphoric gestures: Common and unique integration processes. *Human Brain Mapping*, *32*(4), 520–533. https://doi.org/10.1002/hbm.21041

Straube, B., Green, A., Chatterjee, A., & Kircher, T. (2011). Encoding Social Interactions: The Neural Correlates of True and False Memories. *Journal of Cognitive Neuroscience*, *23*(2), 306–324. https://doi.org/10.1162/jocn.2010.21505

Straube, B., Green, A., Jansen, A., Chatterjee, A., & Kircher, T. (2010). Social cues, mentalizing and the neural processing of speech accompanied by gestures. *Neuropsychologia*, *48*(2), 382–393. https://doi.org/10.1016/j.neuropsychologia.2009.09.025

Straube, B., Green, A., Sass, K., & Kircher, T. (2014). Superior Temporal Sulcus Disconnectivity During Processing of Metaphoric Gestures in Schizophrenia. *Schizophrenia Bulletin*, *40*(4), 936–944. https://doi.org/10.1093/schbul/sbt110

Straube, B., Green, A., Sass, K., Kirner-Veselinovic, A., & Kircher, T. (2013). Neural integration of speech and gesture in schizophrenia: Evidence for differential processing of metaphoric gestures. *Human Brain Mapping*, *34*(7), 1696–1712. https://doi.org/10.1002/hbm.22015

Straube, B., Green, A., Weis, S., Chatterjee, A., & Kircher, T. (2009). Memory Effects of Speech and Gesture Binding: Cortical and Hippocampal Activation in Relation to Subsequent Memory Performance. *Journal of Cognitive Neuroscience*, *21*(4), 821–836. https://doi.org/10.1162/jocn.2009.21053

Straube, B., He, Y., Steines, M., Gebhardt, H., Kircher, T., Sammer, G., & Nagels, A. (2013). Supramodal neural processing of abstract information conveyed by speech and gesture. *Frontiers in Behavioral Neuroscience*, *7*, 51–64. https://doi.org/10.3389/fnbeh.2013.00120

Subramaniam, K., Luks, T. L., Garrett, C., Chung, C., Fisher, M., Nagarajan, S., & Vinogradov, S. (2014). Intensive cognitive training in schizophrenia enhances working memory and associated prefrontal cortical efficiency in a manner that drives long-term functional gains. *NeuroImage*, *99*, 281–292. https://doi.org/10.1016/j.neuroimage.2014.05.057

Suffel, A., Nagels, A., Steines, M., Kircher, T., & Straube, B. (2020). Feeling addressed! The neural processing of social communicative cues in patients with major depression. *Human Brain Mapping*, *41*(13), 3541–3554. https://doi.org/10.1002/hbm.25027

Tamminga, C. A., Stan, A. D., & Wagner, A. D. (2010). The Hippocampal Formation in Schizophrenia. *American Journal of Psychiatry*, *167*(10), 1178–1193. https://doi.org/10.1176/appi.ajp.2010.09081187

Tang, W., Yao, X., & Zheng, Z. (1994). Rehabilitative Effect of Music Therapy for Residual Schizophrenia: A One-Month Randomised Controlled Trial in Shanghai. *The British Journal of Psychiatry*, *165*(S24), 38–44. https://doi.org/10.1192/S0007125000292969

The Lisbon Council. (2019). *Study on open science: Monitoring trends and drivers* (Open Science Monitor D.2.4 Final Report). European Commission.

Tomasello, R., Kim, C., Dreyer, F. R., Grisoni, L., & Pulvermüller, F. (2019). Neurophysiological evidence for rapid processing of verbal and gestural information in understanding communicative actions. *Scientific Reports*, *9*(1), 1–17. https://doi.org/10.1038/s41598-019-52158-w

Troisi, A., Spalletta, G., & Pasini, A. (1998). Non-verbal behaviour deficits in schizophrenia: An ethological study of drug-free patients. *Acta Psychiatrica Scandinavica*, *97*(2), 109–115.

Tschacher, W., Giersch, A., & Friston, K. (2017). Embodiment and Schizophrenia: A Review of Implications and Applications. *Schizophrenia Bulletin*, *43*(4), 745–753. https://doi.org/10.1093/schbul/sbw220

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 Bias Correction. *IEEE Transactions on Medical Imaging*, *29*(6), 1310–1320. https://doi.org/10.1109/TMI.2010.2046908

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage*, *15*(1), 273–289. https://doi.org/10.1006/nimg.2001.0978

Uchiyama, H., Seki, A., Kageyama, H., Saito, D. N., Koeda, T., Ohno, K., & Sadato, N. (2006). Neural substrates of sarcasm: A functional magnetic-resonance imaging study. *Brain Research*, *1124*(1), 100–110. https://doi.org/10.1016/j.brainres.2006.09.088

Uchiyama, H. T., Saito, D. N., Tanabe, H. C., Harada, T., Seki, A., Ohno, K., Koeda, T., & Sadato, N. (2012). Distinction between the literal and intended meanings of sentences: A functional magnetic resonance imaging study of metaphor and sarcasm. *Cortex*, *48*(5), 563–583.

Vancampfort, D., Probst, M., Helvik Skjaerven, L., Catalán-Matamoros, D., Lundvik-Gyllensten, A., Gómez-Conesa, A., Ijntema, R., & De Hert, M. (2012). Systematic Review of the Benefits of Physical Therapy Within a Multidisciplinary Care Approach for People With Schizophrenia. *Physical Therapy*, *92*(1), 11–23. https://doi.org/10.2522/ptj.20110218

Vianin, P., Urben, S., Magistretti, P., Marquet, P., Fornari, E., & Jaugey, L. (2014). Increased activation in Broca's area after cognitive remediation in schizophrenia. *Psychiatry Research: Neuroimaging*, *221*(3), 204–209. https://doi.org/10.1016/j.pscychresns.2014.01.004

Viher, P. V., Abdulkadir, A., Savadijev, P., Stegmayer, K., Kubicki, M., Makris, N., Karmacharya, S., Federspiel, A., Bohlhalter, S., Vanbellingen, T., Müri, R., Wiest, R., Strik, W., & Walther, S. (2020). Structural organization of the praxis network predicts gesture production: Evidence from healthy subjects and patients with schizophrenia. *Cortex*, *132*, 322–333. https://doi.org/10.1016/j.cortex.2020.05.023

Walther, S., Eisenhardt, S., Bohlhalter, S., Vanbellingen, T., Müri, R., Strik, W., & Stegmayer, K. (2016). Gesture Performance in Schizophrenia Predicts Functional Outcome After 6 Months. *Schizophrenia Bulletin*, *42*(6), 1326–1333. https://doi.org/10.1093/schbul/sbw124

Walther, S., Kunz, M., Müller, M., Zürcher, C., Vladimirova, I., Bachofner, H., Scherer, K. A., Nadesalingam, N., Stegmayer, K., Bohlhalter, S., & Viher, P. V. (2020). Single Session Transcranial Magnetic Stimulation Ameliorates Hand Gesture Deficits in Schizophrenia. *Schizophrenia Bulletin*, *46*(2), 286–293. https://doi.org/10.1093/schbul/sbz078

Walther, S., & Mittal, V. A. (2016). Why We Should Take a Closer Look at Gestures. *Schizophrenia Bulletin*, *42*(2), 259–261. https://doi.org/10.1093/schbul/sbv229

Walther, S., Mittal, V. A., Stegmayer, K., & Bohlhalter, S. (2020). Gesture deficits and apraxia in schizophrenia. *Cortex*, *133*, 65–75. https://doi.org/10.1016/j.cortex.2020.09.017

Walther, S., Stegmayer, K., Sulzbacher, J., Vanbellingen, T., Müri, R., Strik, W., & Bohlhalter, S. (2015). Nonverbal social communication and gesture control in schizophrenia. *Schizophrenia Bulletin*, *41*(2), 338–345. https://doi.org/10.1093/schbul/sbu222

Walther, S., Vanbellingen, T., Müri, R., Strik, W., & Bohlhalter, S. (2013a). Impaired gesture performance in schizophrenia: Particular vulnerability of meaningless pantomimes. *Neuropsychologia*, *51*(13), 2674–2678. https://doi.org/10.1016/j.neuropsychologia.2013.08.017

Walther, S., Vanbellingen, T., Müri, R., Strik, W., & Bohlhalter, S. (2013b). Impaired pantomime in schizophrenia: Association with frontal lobe function. *Cortex*, *49*(2), 520–527. https://doi.org/10.1016/j.cortex.2011.12.008

Wende, K., Straube, B., Stratmann, M., Sommer, J., Kircher, T., & Nagels, A. (2012). Neural correlates of continuous causal word generation. *NeuroImage*, *62*(3), 1399–1407. https://doi.org/10.1016/j.neuroimage.2012.06.003

Wexler, B. E., Anderson, M., Fulbright, R. K., & Gore, J. C. (2000). Preliminary Evidence of Improved Verbal Working Memory Performance and Normalization of Task-Related Frontal Lobe Activation in Schizophrenia Following Cognitive Exercises. *American Journal of Psychiatry*, *157*(10), 1694–1697. https://doi.org/10.1176/appi.ajp.157.10.1694

White, T. P., Borgan, F., Ralley, O., & Shergill, S. S. (2016). You looking at me?: Interpreting social cues in schizophrenia. *Psychological Medicine*, *46*(1), 149–160. https://doi.org/10.1017/S0033291715001622

Wible, C. G. (2012). Hippocampal temporal-parietal junction interaction in the production of psychotic symptoms: A framework for understanding the schizophrenic syndrome. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00180

Willems, R. M., & Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: A review. *Brain and Language*, *101*(3), 278–289. https://doi.org/10.1016/j.bandl.2007.03.004

Willems, R. M., Özyürek, A., & Hagoort, P. (2007). When Language Meets Action: The Neural Integration of Gesture and Speech. *Cerebral Cortex*, *17*(10), 2322–2333. https://doi.org/10.1093/cercor/bhl141

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. World Health Organization. https://apps.who.int/iris/handle/10665/37958

Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI Time-Series Revisited—Again. *NeuroImage*, *2*(3), 173–181. https://doi.org/10.1006/nimg.1995.1023

Wroblewski, A., He, Y., & Straube, B. (2020). Dynamic Causal Modelling suggests impaired effective connectivity in patients with schizophrenia spectrum disorders during gesture-speech integration. *Schizophrenia Research*, *216*, 175–183. https://doi.org/10.1016/j.schres.2019.12.005

Wu, Y. C., & Coulson, S. (2010). Gestures modulate speech processing early in utterances. *Neuroreport*, *21*(7), 522–526. https://doi.org/10.1097/WNR.0b013e32833904bb

Wüthrich, F., Pavlidou, A., Stegmayer, K., Eisenhardt, S., Moor, J., Schäppi, L., Vanbellingen, T., Bohlhalter, S., & Walther, S. (2020). Nonverbal communication remains untouched: No beneficial effect of symptomatic improvement on poor gesture performance in schizophrenia. *Schizophrenia Research*, *223*, 258–264. https://doi.org/10.1016/j.schres.2020.08.013

Wykes, T. (1998). What are we changing with neurocognitive rehabilitation?: Illustrations from two single cases of changes in neuropsychological performance and brain systems as measured by SPECT. *Schizophrenia Research*, *34*(1), 77–86. https://doi.org/10.1016/S0920-9964(98)00076-0

Wykes, T., Brammer, M., Mellers, J., Bray, P., Reeder, C., Williams, C., & Corner, J. (2002). Effects on the brain of a psychological treatment: Cognitive remediation therapy: Functional magnetic resonance imaging in schizophrenia. *The British Journal of Psychiatry*, *181*(2), 144–152. https://doi.org/10.1192/bjp.181.2.144

Yap, D.-F., So, W.-C., Melvin Yap, J.-M., Tan, Y.-Q., & Teoh, R.-L. S. (2011). Iconic Gestures Prime Words. *Cognitive Science*, *35*(1), 171–183.

Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, *20*(1), 45–57. https://doi.org/10.1109/42.906424