How to Read Papers for a Meta-Analysis

In addition to the checklist, I wanted to give some strategies to do targeted reading for a meta-analysis. The idea is that we are trying to extract key pieces of information from each paper. Once you know where to look in a paper, and how to get these pieces of information, you won't have to read through each paper in its entirety!

Most of what you need will come from the 'Methods' section of the paper. The Methods section (aka Methods and Materials) of a paper is usually in one of two places. Many of the papers will have the Methods section right after the Introduction, like Diana's paper and the Schuwerk paper. Some papers will have the Methods section at the end of the paper. Rarely, a paper will not have a heading for the methods section, like the Kircher paper. Within the Methods section, there are often subheadings like Participants (where you can get sample size, age range etc), Task, Design, and/or Procedure sections (where you can find out what they did), and sections that outline their Analyses (where you can find if they did conducted appropriate contrasts).

STEP 1) Is it a social interaction study? The first step is to identify whether the paper meets our criteria for a study of social interaction. I'll have screened most of the papers that I send you, but it's possible that I have made a mistake. So, the first step is to double check that it meets criteria. To find this out, you can go to the Task and/or Procedure subheading of the Methods section.

--

Just a reminder, ***Social Interactions*** *are the reciprocal exchanges between socially engaged individuals. The participant has to at least think that they are interacting with another live person in real time. This definition includes two main parts:*

1) Social engagement *– the participant is socially engaged with a live human partner who is engaging with the participant in real time.*

- *Can simply be through being told that they are engaging with a real person*
- *Person can be outside the scan room or in a distant location connected digitally*
- *The participant has to believe it, and feel as if they are engaged with the partner*

2) Interaction *– the participant has to be involved in a reciprocal exchange with the interaction partner, or at least be ready to reciprocate with the interaction partner.*

- *Exchange can be through eye-gaze, hand gestures, digital chat, game play, etc.*
- *Participant or partner has to at least make one response to something the other does; the response doesn't have to be immediate, but must be contingent on the other's action*

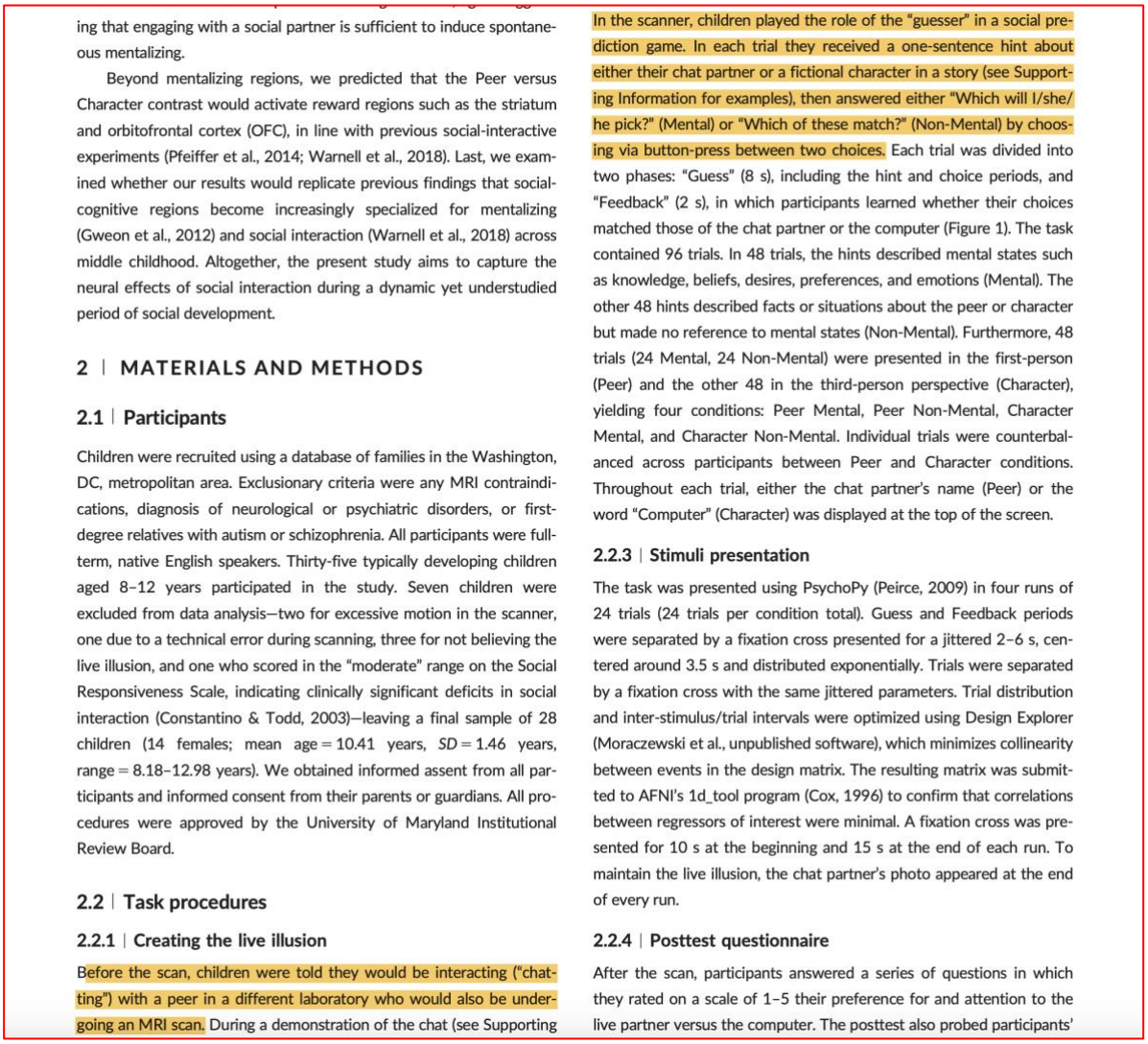
--

First, you'll want to make sure that they convinced participants that they will be engaging with another person in real time. Usually, they are simply told that will be engaging with another person.

Second, you want to make sure that the participant is doing something that involves a reciprocal exchange. Some papers have participants simply observing another person, which

does not count! Reciprocal exchange is fairly broad, so they can be responding to a person's eye-gaze, playing a game, or responding to questions posed by their interaction partner.

Here are a couple examples of methods sections that meet both criteria:



ing that engaging with a social partner is sufficient to induce spontaneous mentalizing.

Beyond mentalizing regions, we predicted that the Peer versus Character contrast would activate reward regions such as the striatum and orbitofrontal cortex (OFC), in line with previous social-interactive experiments (Pfeiffer et al., 2014; Warnell et al., 2018). Last, we examined whether our results would replicate previous findings that social-cognitive regions become increasingly specialized for mentalizing (Gweon et al., 2012) and social interaction (Warnell et al., 2018) across middle childhood. Altogether, the present study aims to capture the neural effects of social interaction during a dynamic yet understudied period of social development.

2 | MATERIALS AND METHODS

2.1 | Participants

Children were recruited using a database of families in the Washington, DC, metropolitan area. Exclusionary criteria were any MRI contraindications, diagnosis of neurological or psychiatric disorders, or first-degree relatives with autism or schizophrenia. All participants were full-term, native English speakers. Thirty-five typically developing children aged 8–12 years participated in the study. Seven children were excluded from data analysis—two for excessive motion in the scanner, one due to a technical error during scanning, three for not believing the live illusion, and one who scored in the "moderate" range on the Social Responsiveness Scale, indicating clinically significant deficits in social interaction (Constantino & Todd, 2003)—leaving a final sample of 28 children (14 females; mean age = 10.41 years, *SD* = 1.46 years, range = 8.18–12.98 years). We obtained informed assent from all participants and informed consent from their parents or guardians. All procedures were approved by the University of Maryland Institutional Review Board.

2.2 | Task procedures

2.2.1 | Creating the live illusion

Before the scan, children were told they would be interacting ("chatting") with a peer in a different laboratory who would also be undergoing an MRI scan. During a demonstration of the chat (see Supporting

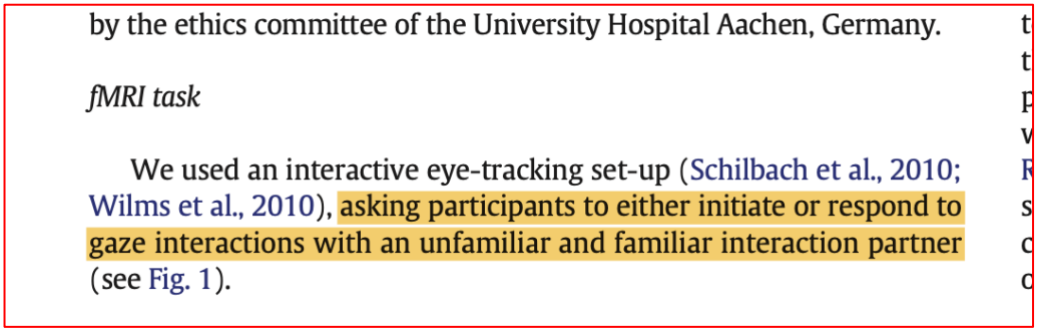
In the scanner, children played the role of the "guesser" in a social prediction game. In each trial they received a one-sentence hint about either their chat partner or a fictional character in a story (see Supporting Information for examples), then answered either "Which will I/she/he pick?" (Mental) or "Which of these match?" (Non-Mental) by choosing via button-press between two choices. Each trial was divided into two phases: "Guess" (8 s), including the hint and choice periods, and "Feedback" (2 s), in which participants learned whether their choices matched those of the chat partner or the computer (Figure 1). The task contained 96 trials. In 48 trials, the hints described mental states such as knowledge, beliefs, desires, preferences, and emotions (Mental). The other 48 hints described facts or situations about the peer or character but made no reference to mental states (Non-Mental). Furthermore, 48 trials (24 Mental, 24 Non-Mental) were presented in the first-person (Peer) and the other 48 in the third-person perspective (Character), yielding four conditions: Peer Mental, Peer Non-Mental, Character Mental, and Character Non-Mental. Individual trials were counterbalanced across participants between Peer and Character conditions. Throughout each trial, either the chat partner's name (Peer) or the word "Computer" (Character) was displayed at the top of the screen.

2.2.3 | Stimuli presentation

The task was presented using PsychoPy (Peirce, 2009) in four runs of 24 trials (24 trials per condition total). Guess and Feedback periods were separated by a fixation cross presented for a jittered 2–6 s, centered around 3.5 s and distributed exponentially. Trials were separated by a fixation cross with the same jittered parameters. Trial distribution and inter-stimulus/trial intervals were optimized using Design Explorer (Moraczewski et al., unpublished software), which minimizes collinearity between events in the design matrix. The resulting matrix was submitted to AFNI's 1d_tool program (Cox, 1996) to confirm that correlations between regressors of interest were minimal. A fixation cross was presented for 10 s at the beginning and 15 s at the end of each run. To maintain the live illusion, the chat partner's photo appeared at the end of every run.

2.2.4 | Posttest questionnaire

After the scan, participants answered a series of questions in which they rated on a scale of 1–5 their preference for and attention to the live partner versus the computer. The posttest also probed participants'



by the ethics committee of the University Hospital Aachen, Germany.

fMRI task

We used an interactive eye-tracking set-up (Schilbach et al., 2010; Wilms et al., 2010), asking participants to either initiate or respond to gaze interactions with an unfamiliar and familiar interaction partner (see Fig. 1).

IF YOU DETERMINE THAT THE STUDY DOES NOT QUALIFY AS A STUDY OF A SOCIAL INTERACTION, MARK THAT IT DOESN'T, AND YOU DON'T HAVE TO DO ANYTHING ELSE!

STEP 2) Is there an appropriate contrast? Once you have determined that the paper meets our criteria for a study of social interaction, you next need to determine if the paper has an appropriate contrast that targets social interaction. Here again, most of the earlier papers I send you will be screened for this criterion, but please check again to make sure that I did not make a mistake. Also, this is not something that will always be apparent, so if you have any questions, please ask.

Just to remind you again, an appropriate contrast is one that isolates the brain activations involved in social interaction by subtracting out all the other processes that might be involved. For example in Diana's paper, she had participants make guesses about an abstract character as a control condition because this also includes decision making, similar visual/verbal information, and the same motor responses. The only thing different between the peer and character conditions is that the peer condition involves interacting with a real person. Therefore, by subtracting the character condition, we can isolate social interaction. Not all papers will use the terminology "social interaction", so you will have to make sure.

Common control conditions for studies of social interaction include:
1) Playing with a person vs playing with a computer
2) Playing with another vs playing alone
3) Thinking of a person vs thinking of an abstract character

There are plenty of other appropriate control conditions too that are clever. I'll do my best to do a first pass of screening for this stuff, but please double check in case I messed up.

Rest (aka baseline) is NOT an appropriate condition, and you might notice that some of the older papers do this! Rest/baseline just means that they looked at brain activation without any contrast, so it is a mix of all the different neurocognitive processes that might be involved in the task.

Also, papers might not have an appropriate control condition even though they meet our criteria for social interaction, in which case we cannot use them. For example, some studies might have participants play two different games with another person, one easy and hard, and do a contrast that is examining the effect of game difficulty. In this case, because both conditions involve social interaction, it does not count!

Here are a couple examples (shared_emo = sharing emotion with another):

How to Read Papers for a Meta-Analysis

Shared_emo > unshared_emo		
VS[a]	R	
VMPFC/OFC[a]	L	11

Human > Baseline	
R	**Middle Frontal Gyrus** Superior Medial Frontal Gyrus
R	**Inferior Parietal Cortex** Temporo-Parietal Junction
L	**Angular Gyrus** Superior Parietal Cortex
Computer > Baseline	
R	**Middle Frontal Gyrus**
L/R	**Superior Frontal Gyrus (medial part)** **Angular Gyrus** Temporo-Parietal Junction
R	**Inferior Frontal Gyrus (orbital part)** Superior Temporal Pole
L	**Middle Frontal Gyrus**
Human > Computer	
	Superior Frontal Gyrus (medial part) Anterior Cingulate Cortex
	Thalamus

STEP 3) Sample information. Once you have determined that the study meets criteria and has an appropriate contrast, you need to write down sample characteristics, including sample size, gender break down, and average age (also age range and standard deviation if they provide it). For this analysis, we are only considering studies using healthy, neurotypical samples, so any study looking at autism, for example, will need to be excluded. All this information will be under the Participants subheading of the Methods section.

1) Sample size, sometimes referred to as 'n' or 'N' (for number), is simply the number of participants included in the final analysis of the study. Pay attention because sometimes papers will first list the number of people that enrolled in the study first, and explain why the excluded certain people before listing the final sample size used in the analysis. Finally, we are interested specifically in the sample size of the scanned participants.
2) Gender breakdown – All papers should indicate the number of this sample that is male/female. You can simply mark down the number of females in the sample.
3) Age – All papers should indicate the mean (average) age of the sample used. Often times, papers will also include the age range of the study (youngest to oldest), and the standard deviation (SD) in age of the sample. If they include age range and standard deviation, please mark these as well. Sometimes, papers will only list the age, age range, and SD of the people enrolled in the study rather than the final sample; in which case you can mark these down, but make a note.

How to Read Papers for a Meta-Analysis

Here are a couple examples from papers:

Materials and methods

Participants

32 right-handed volunteers participated in the study, which was approved by the ethics committee of the Medical Faculty of the University of Cologne. 12 participants were excluded due to excessive movements (n = 4), technical problems with the eye-tracker (n = 5), and disbelief in the cover story (n = 3). 20 participants (9 female/11 male, *M* = 27.75 years, *SD* ± 6.44) were included in the analyses.

METHODS

Participants

Sixty female volunteers (30 pairs of friends) completed the experiment after giving informed consent. Magnetic resonance imaging (MRI) data were obtained from 30 of them. They were recruited from the Charité Universitätsmedizin Berlin and the Free University Berlin, and through social media advertisements. The study was approved by the ethics committee of the Charité Universitätsmedizin Berlin, Campus Mitte. Participants were aged between 20 and 33 years (*M* = 24.76, SEM = 0.364) and fit common inclusion criteria for fMRI research. Each participant received 25 or 30 Euros depending on the duration of the session. A control questionnaire confirmed that the two participants in each pair regarded each other as close friends [on a scale from 0 (not true at all) to 10 (entirely true), the item 'How close is this friend to you?' was rated with *M* = 8.00, s.d. = 1.775; the item 'How important is this friendship for your life?' was rated with *M* = 8.667, s.d. = 1.348]. Two participants were excluded from fMRI analysis owing to technical error in the acquisition of brain images. Two further participants reported developing a headache in the scanner and were therefore excluded from both behavioral and fMRI data analyses. Thus, behavioral analysis included *n* = 58 participants, and fMRI analyses *n* = 26 participants.

One final note about samples. Sometimes, a paper will have more than one sample. This is sometimes the case with papers that have two or more different experiments. In these cases, we need to mark the sample characteristics of each sample. Also, we'll have to make note about which sample corresponds with each contrast to make sure that we are tagging them properly. Here is an example of that:

Methods

Go to:

Participants

Participants for both experiments were recruited from the Boston community. Participants had never met the experimenter prior to the scan date. All participants gave informed written consent as approved by the MIT human subjects committee (COUHES) and were compensated for their participation. Participants were excluded if they self-reported any psychiatric or neurological diagnosis on a medical questionnaire. In addition, all participants completed a Social Communication Questionnaire (Berument et al., 1999) and scored below 15 (i.e. in the normal range). For Experiment 1, data were collected from 23 participants. Seven were excluded for the following reasons: head movement exceeding the criteria detailed below (n=4), falling asleep (2), and equipment malfunction (1). The final sample consisted of 16 adults (7 males, age 18–29 years, mean 22.8 ± 3.1 years). For Experiment 2, data were collected from 14 participants. One was excluded for exceeding the motion criteria, leaving a sample of 13 adults (6 males, age18–29 years, mean 22.6 ± 3.5 years).