



Universiteit  
Leiden  
The Netherlands

## Understanding existential self-understanding : philosophy meets cognitive neuroscience

Stee, M.J. van

### Citation

Stee, M. J. van. (2017, June 21). *Understanding existential self-understanding : philosophy meets cognitive neuroscience*. Retrieved from <https://hdl.handle.net/1887/49803>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/49803>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/49803> holds various files of this Leiden University dissertation.

**Author:** Stee, M.J. van

**Title:** Understanding existential self-understanding : philosophy meets cognitive neuroscience

**Issue Date:** 2017-06-21

# 1

## Cognitive Neuroscience of Love

Cognitive neuroscience (CNS) is a scientific discipline that investigates the neurobiological underpinnings of mental processes. It combines the methods of cognitive psychology with those of neuroscience. In doing so, CNS addresses the question how brain activity gives rise to mental processes such as perception, action, attention, and especially cognitive processes such as memory, language, and reasoning. CNS research has taken a great flight in recent decades, due to the advent of neuroimaging techniques. Before, we could only learn about the brain and its involvement in mental processes through animal studies and through post-mortem autopsies on patients whose deficits were studied while they were still alive. With the advent of neuroimaging techniques it has become possible to study healthy human brains, in action. Although most research focuses on processes such as perception, memory and language, some CNS research investigates topics that pertain more directly to existential self-understanding. CNS of (so-called) free will, for example, has received a lot of attention as potentially telling us something new about ourselves. Other CNS literatures have yet to receive such attention. In this chapter, I investigate CNS of love to determine what contributions to our understanding of existential selfhood it has to offer.

CNS research specifically targeting something called ‘the existential self’ or ‘existential self-understanding’ does not exist. Certain studies use the term ‘existential’ in their title, but they take their cues from crude versions of Heideggerian thought. For example, they investigate what neural activity correlates with people contemplating their own death (Klackl, Jonas, and Kronbichler 2012; Quirin et al. 2011). The term ‘neuroexistentialism’ has also been used, albeit not in actual CNS experiments but in the theorizing surrounding CNS. First, one researcher uses it to label the argument that certain neuroscientific findings make more sense if we think of humans as being-in-the-world rather than as subjects dependent for their mental life on representations in the brain divorced from the outer world (Iacoboni 2007), a view more generally labelled embodied embedded cognition. Second, ‘neuroexistentialism’ sometimes labels the view that we are our brain, a potentially alienating thought that may induce a renewed search for meaning (Churchland 2013; Flanagan 2009). We encounter some of the ideas behind

these views when I turn to the question what future CNS may contribute to human self-understanding in chapter 4. First, we look into empirical research literatures, however: CNS of love in this chapter and CNS of self-reflection in the next.

When it comes to things that are of existential importance to people, love provides key instances. Clemens' identity crisis has come about after he lost the people he loves most. The situations of Bob and Anna also involve their significant others, specifically the dog and the husband that they are supposed to love but currently stand in ambiguous relations to. Deidre may be said to doubt between two loves: her passion for mathematics having grown weary, her love of philosophy growing. Edward experiences crisis exactly when he realizes he has been in denial about his love for Ms. Kenton. Insight into love is bound to give insight into existential aspects of ourselves.

Cognitive neuroscience research into love has only been around for a short while. The study that is generally considered to have been the start of the field was reported in 2000, by Andreas Bartels and Semir Zeki. I conducted a search for CNS of love studies by performing Web of Science and PubMed searches with the keywords 'neuroimaging' and 'love'. I looked through the reference lists of the research and review articles I thus found to find further studies. For a study to be included it had to satisfy the following inclusion criteria: it had to be a neuroimaging study (including electroencephalography (EEG) studies) reporting original data; of love; performed on a healthy population; and reported before 2015. I excluded studies that had 'love' amongst their key words, but looked at the modulating effects of love on e.g. attention or memory processing, as these studies ultimately aim to contribute to our understanding of attention or memory, rather than to our understanding of love. In November 2016, I made an additional search, and found two more recent studies of a similar kind. A total number of 16 + 2 studies were found. Details of these studies can be found in appendix A.

In this chapter, I review CNS of love with an eye to the questions regarding love it addresses and the conceptual assumptions about love that are embedded in choices made in the research process. In effect, this chapter is a conceptual review of CNS of love. In the first part, I walk through the set-up of a CNS experiment, following the order in which it is reported in CNS research articles: introduction, methods, results, discussion. I pay particular attention to interpretative choices researchers have to make, which generally receive little scrutiny. In the second part of the chapter, I tease out and analyze the conceptual implications of these choices. Finally, I draw conclusions regarding the questions CNS of love addresses and the role interpretation plays in its research process.<sup>2</sup>

---

<sup>2</sup> Parts of this chapter are included in a paper on cognitive neuroscience and self-understanding (van Stee forthcoming).

## 1. CNS experiments on love and the use of interpretation

### Introduction: questions CNS of love aims to address

In the introduction of CNS research articles, researchers position their study within the wider literature. They introduce the question they aim to address and argue for its relevance to the wider research enterprise. Both the question itself and the framing within the wider literature provide information regarding the questions researchers aim to address with their study. If, for example, the research question is introduced as ‘what are the neural correlates to early-stage intense romantic love?’ and researchers argue for the relevance of that question by reference to neuroscientific literature studying pair bonding in other animals, then we know that researchers are also interested in the further question how the neural underpinnings of human romantic love compare to those of pair bonding in other animals. At the end of the article, in the discussion section, researchers return to the wider frame and discuss the implications of the results they found for those further questions.

All studies in CNS of love ask for the neural correlates of love, be it romantic love (Acevedo et al. 2012; Aron et al. 2005; Bartels and Zeki 2000; Fisher et al. 2010; Guerra et al. 2011; Kim et al. 2009; Langeslag et al. 2007; Langeslag, van der Veen, and Röder 2014; Ortigue et al. 2007; Tiedt et al. 2014; Xu et al. 2011; Zeki and Romaya 2010), maternal love (Bartels and Zeki 2004; Noriuchi, Kikuchi, and Senoo 2008) or unconditional love (Beauregard et al. 2009). All are interested in one way or the other to determine to what extent these neural correlates are specific to love or shared with other processes. From the choice of contrast condition (about which more in the next section) and the wider theoretical frame, I can deduce researchers’ interest in several sets of questions.

A first set of questions tries to delineate neural activity correlated with love from neural activity correlated with processes closely related to love, so as to determine what neural activity may specifically underlie love. Research questions include:

- How do the neural correlates for love compare to those of friendship? What are the overlap and differences between the neural correlates (Acevedo et al. 2012; Bartels and Zeki 2000, 2004; Kim et al. 2009; Langeslag et al. 2007; Langeslag, van der Veen, and Röder 2014; Ortigue et al. 2007; Tiedt et al. 2014; Zeki and Romaya 2010)?
- How do neural correlates overlap or differ between love and familiarity (Acevedo et al. 2012; Aron et al. 2005; Bartels and Zeki 2004; Fisher et al. 2010; Guerra et al. 2011; Xu et al. 2011)?
- How do neural correlates of romantic love compare to those of sexual attraction (Diamond and Dickenson 2012; Ortigue et al. 2010)?

- How do neural correlates to romantic love overlap with or differ from neural correlates to attraction to beauty? That is, does processing pictures of a romantic partner differ from processing pictures of a beautiful person of the same sex and age as the romantic partner (Langeslag et al. 2007)?
- How does neural activity correlating with romantic love compare to the neural activity correlating with emotional arousal more generally (Guerra et al. 2011; Vico et al. 2010)?

In a second set of questions, researchers try to delineate the neural correlates to different forms of love from each other. Besides the review by Stephanie Ortigue and colleagues (2010), primary research articles address questions like:

- To what extent do neural mechanisms involved in romantic pair bonding overlap with or differ from attachment mechanisms in parent-infant bonding (Acevedo et al. 2012; Bartels and Zeki 2004)?
- How do neural correlates of unconditional love differ or overlap with neural correlates correlated with parental or romantic love (Beauregard et al. 2009)?
- How do neural correlates differ or overlap for romantic love versus a passion for a pursuit (Ortigue et al. 2007)?

A further set of questions seeks to explore love in terms of motivation and/or argue for a characterization of love as a motivational state:

- What are the neural mechanisms of the motivational force of love (Aron et al. 2005; Bartels and Zeki 2004; Beauregard et al. 2009; Ortigue et al. 2007)?
- Is romantic love better characterized as an emotion or as a motivation (Aron et al. 2005; Fisher et al. 2010)?

Now that research is underway, researchers ask questions to uncover more about neural activity involved in love and its proper functional interpretation.

- Can the neural results generally found in CNS of love be replicated through the use of magnetoencephalography (MEG) (Tiedt et al. 2014)?
- Do the neural results align with peripheral measures, such as heart rate, skin conductance (i.e. sweating) and electromyography (EMG) of the zygomatic major muscle (i.e. the muscle mediating the smile) (Guerra et al. 2011; Vico et al. 2010)?
- How should enhanced late positive potentials be interpreted? Do they reflect familiarity or positive emotion or something else?

Researchers may also compare different groups of participants to see whether neural correlates to their love experiences are different or the same. This has so far only been done for romantic love:

- Do neural correlates of romantic love differ for males versus females (Tiedt et al. 2014; Zeki and Romaya 2010)?
- Do they differ for homosexuals versus heterosexuals (Zeki and Romaya 2010)?
- Do neural correlates for romantic love differ in Western (American) vs. Eastern (Chinese) people (Xu et al. 2011)?
- Do neural correlates for romantic love differ when people hold traditional cultural values as compared to modern cultural values (Xu et al. 2011)?
- What are the differences and similarities between neural correlates to early-stage romantic love and neural correlates to (reported) long-term romantic love (Acevedo et al. 2012)?
- Do neural correlates to romantic love differ for people who are in a happy relationship with the person they are in love with versus for people who have recently been rejected but are still in love (Fisher et al. 2010)?

Although no study directly investigates the question of the overlap or difference between neural correlates underlying human love and equivalents in other animals, several studies are framed in those terms:

- How do pair bonding and motivational mechanisms in humans compare to those in other monogamous animals (Acevedo et al. 2012; Aron et al. 2005; Fisher et al. 2010)?
- How do mechanisms involved in maternal love in humans compare to mechanisms involved in maternal behavior in other animals (Bartels and Zeki 2004; Noriuchi, Kikuchi, and Senoo 2008)?

One study presents the beloved in different ways and compares neural activity patterns in response to those changes:

- How does neural activity differ when mothers watch their beloved child be happy or when they watch their beloved child be distressed (Noriuchi, Kikuchi, and Senoo 2008)?

Finally, studies may inquire after (and aim to predict) changes throughout the course of a love relationship:

- Do the neural correlates correlated with romantic love change over the course of 6 months and does this change correlate with a change in the experience of love (Kim et al. 2009)?

- Does brain activation in early-stage romantic love reliably correlate with (and therefore predict) later satisfaction in the relationship (Xu et al. 2011)?

### **Methods: setting up experiments in CNS of love**

To actually address their research questions, cognitive neuroscientists have to operationalize love. They make choices regarding what participants to recruit. They choose what tasks to make participants perform whilst their brains are being scanned, which in turn involves choosing or developing stimuli and instructions, and choosing or developing control tasks. They also make choices regarding what brain imaging technique to use and how to set its parameters, and what analysis protocol to employ on the eventual data. All these choices are reported in the methods section of research articles. In this section, I summarize the choices researchers in CNS of love make that have conceptual consequences: selection of participants; choice of stimuli; wording of instructions; and choice of control task. The questions that are directly addressed in CNS of love are more specific than the ones given in the introduction of research articles. Methodological choices ensure that studies zoom in on an aspect of love. Later, in the second part of this chapter, I draw out the precise conceptual implications of methodological choices made in CNS of love.

### ***Participants***

Researchers have to select participants for their studies, strictly speaking in such a way that their pool of participants forms a representative sample of the population at large about which they aim to discover something. For pragmatic reasons, participants in experiments tend to be undergraduate students of the university at which researchers work: they are easily accessible, tend to be healthy and can be induced to participate through receiving course credit or a small amount of money in return. What is more, their homogeneity as a group increases the likelihood of finding significant results, as noise in the data due to individual differences between participants is likely to be less than between people of different cultural and educational backgrounds. It can be questioned to what extent these WEIRD people (i.e. people from Western, Educated, Industrialized, Rich, and Democratic societies) are representative of the overall population (Henrich, Heine, and Norenzayan 2010). Generally though, researchers write up their research articles as if they are.

CNS of love is interesting for the way in which researchers select participants to their experiments. Whereas for most CNS studies any healthy undergraduate will do, researchers in CNS of romantic love demand more. 'Are you truly, madly, deeply in love?'

is what the poster asked with which Bartels and Zeki (2000) recruited their participants. Reporting ‘yes’ in response to this question is not enough. Researchers follow up with interviews and other checks and they are strict too. Acevedo and colleagues (2012) report excluding approximately 40% of the people who professed to being madly in love. Bartels and Zeki (2000) included only 17 out of 70 participants that signed up. For CNS of maternal love on the other hand, matters are not that strict: being a mother will do. For unconditional love, finally, a procedure was used to select experts. Specifically, directors of living communities for people with severe intellectual disabilities were asked to select those employees that they deemed most capable of unconditional love. Researchers thus make different choices as to how to determine whom to include, depending on the type of love they are investigating. They do not report reasons for choosing participants in the particular way that they do. See table 1 for an overview over the ways in which participants are recruited and the characteristics that participants have.

*Table 1: Participants in Cognitive Neuroscience of Love*

Study	Selection criteria	Female / male	Age range (mean)	Relationship length (mean)
Acevedo et al. (2012)  Long-term romantic love	Flyer, newspaper ads, word of mouth: ‘Are you still madly in love with your long-term partner?’  Phone screening: Relationship length >10 years Monogamous heterosexual relation Feelings of intense romantic love	10 ♀ / 7 ♂  equals ± 60% of original respondents	39-67 years (53 years)	‘being married’ 10-29 years (21 years)
Aron et al. (2005)  Romantic love	Flyer, word of mouth: ‘currently intensely in love’	10 ♀ / 7 ♂	18-26 years (20.6 years)	‘being in love’ 1-17 months (7.4 months)
Bartels & Zeki (2000)  Romantic love	Posters, via the internet: ‘truly, deeply and madly in love’  Written statement describing how much they are in love + interview	11 ♀ / 6 ♂  out of 70 original respondents, 75% of which female (write B&Z)	21-37 years (24.5 years)	(2.4 years)
Bartels & Zeki (2004)  Maternal love	Posters in nurseries: being mother	20 ♀	27-49 years (34 years)	‘age child’ 9 months – 6 years (24.4 months)

Study	Selection criteria	Female / male	Age range (mean)	Relationship length (mean)
Beauregard et al. (2009)  Unconditional love	Directors of l'Arche residential communities where people with intellectual disabilities live together with so-called assistants helped recruit those assistants: 'with a very high capacity for unconditional love'; who understand the meaning of 'unconditional love'; and find work at l'Arche very gratifying.	9 ♀ / 8 ♂	20-63 years (36 years)	
Fisher et al. (2010)  Rejected romantic love	Flyer, word of mouth: 'Have you just been rejected in love but can't let go?'  Interview: mixed emotions obsessively thinking about beloved rejecter	10 ♀ / 5 ♂	18-21 years (19.8 years)	4-48 months (21 months)  end relationship: 1-32 weeks ago (9 weeks ago)
Guerra et al. (2011)  Romantic and filial love	Undergraduate students in a romantic relationship reside close to family and partner have a positive relationship with father	35 ♀	20-29 years (21.7 years)	
Kim et al. (2009)  Romantic love	Advertisements by broadcaster: couples who had fallen in love < 100 days ago. Were fine with appearing in a documentary on love being broadcast on Korean television.	5 ♀ / 5 ♂  5 couples selected out of ± 100 couples who volunteered	18-24 years (21.1 years)	< 100 days
Langeslag et al. (2007)  Romantic love	Posters at Dutch university: in love with someone of opposite sex	9 ♀ / 9 ♂	18-34 years (21.5 years)	(12.1 months)  'love duration': 2.5-36 months (12.6 months)
Langeslag, Van der Veen, Röder (2014)  Romantic love	Dutch students: in love with someone of opposite sex for < 9 months	9 ♀ / 6 ♂	18-25 years (20.8 years)	1-6.5 months (3.9 months)  'love duration': 2.5-8 months (5.1 months)
Noriuchi et al. (2008)  Maternal love	Mothers of infants	13 ♀	(31.1 years)	'age infant': (16.5 months)

Study	Selection criteria	Female / male	Age range (mean)	Relationship length (mean)
Ortigue et al. (2007)  Romantic love	Advertisements to students: 'individuals who are currently intensely in love'  Heterosexual Dating, engaged, or married Have a favorite passion in life (e.g. science, sports, art) Report thinking about passionate hobby 60% of the day	36 ♀	(20.1 years)	'being in love': 1-60 months (15.3 months)
Tiedt et al. (2014)  Romantic love	in a long-term relationship (> 1 year)	14 ♀ / 14 ♂	20-35 years (25.5 years)	12-96 months (40.2 months)
Vico et al. (2010)  Love	Undergraduate students in a romantic relationship reside close to partner and four other loved ones	30 ♀	20-27 years	
Xu et al. (2011)  Romantic love	Flyers to Beijing student email lists: "currently in a relationship and very intensely in love"	10 ♀ / 8 ♂	19-25 years (21.6 years)	1.3-13 months (6.5 months)
Zeki & Romaya (2010)  Romantic love	Advertisements: 'passionately in love'  In sexual relationship with lover Equal number of males and females in heterosexual and homosexual relationships	12 ♀ / 12 ♂	19-47 years (26.3 years)	4 months – 23 years (3.7 years)

*Most studies also used selection criteria that have to do with ability to perform the experiment and participant comparability on the neural level. Examples are: normal or corrected-to-normal vision; no use of (antidepressant) medication; no history of psychiatric, neurological or substance abuse disorders; right-handedness. As these criteria are not particular to CNS of love, I have not included them in the table.*

### **Tasks: stimuli, instructions, control tasks**

To study the neural correlates to love, cognitive neuroscientists have to operationalize love. That is to say, they have to make their participants experience love or otherwise be loving whilst their brains are being scanned. To this end, researchers develop tasks: they employ stimuli and instruct their participants on what to do. Importantly, researchers also develop control tasks. Performing a task generally involves several cognitive

processes, but only one is of interest. The control task also involves the irrelevant cognitive processes, but not the process of interest. Neural activity occurring during the control task is subtracted from neural activity correlated with the experimental task. The net result is data on the neural activity related to the process of interest only. This is called the subtraction method. It depends on several assumptions, for example regarding pure insertion of mental processes. Other methods exist too, each with its own assumptions. The subtraction method is still the most basic one, commonly used in CNS, and the only one used in the studies on the CNS of love. When devising operationalizations, researchers therefore think hard about finding a proper contrast between experimental and control task. The contrast may arise through using different stimuli for the experimental and control task or by changing the task participants are instructed to perform in response to the same stimuli.

The tasks employed in CNS of love are relatively homogenous. Most studies rely on pictures for stimuli. Participants have to bring pictures of their loved one that researchers then include in the experiment. Instructions differ only slightly: 'think about (non-sexual) experiences with this person' instructs a study on romantic love (Acevedo et al. 2012). 'View the pictures, think of the viewed person, and relax' demands another (Bartels and Zeki 2000) and a third, on maternal love, simply states that participants should 'view the pictures and relax' (Bartels and Zeki 2004). Participants thus do not have to press any buttons or otherwise respond physically; researchers hope they will simply induce an experience of love in their participants by showing them pictures of loved ones. In a variation on this general approach, Noriuchi, Kikuchi, and Senoo (2008) do not use pictures, but rather video recordings in their study on maternal love, instructing their participants to simply watch the recordings.

Neural activation during exposure to the loved one is then contrasted with neural activation during exposure to others. If there were no control task, the neural activity that would be measured would include neural activity related to the experience of love, but also to face perception, amongst other things. A control task is needed that requires participants to look at pictures of the face of someone they do not love. If pictures of random strangers are chosen for the control condition, however, net neural activity may reflect processing a familiar face versus an unfamiliar face, rather than a beloved face versus an unloved face. On the other hand, if researchers ask their participants to bring pictures of a friend they know as well as their partner, their feelings for the friend may not differ enough from the feelings for their partner to lead to neural activation patterns that are different enough to reach statistical significance. The choices involved are thus many and they are complex. In practice, researchers in CNS of romantic love use a variety of contrasts, all of which involve pictures of people of the same sex and age as the romantic partner. Some further control for familiarity by using pictures of a highly

familiar acquaintance such as a colleague or classmate (Acevedo et al. 2012; Aron et al. 2005; Fisher et al. 2010; Xu et al. 2011) whereas others (also) control for friendly feelings by using pictures of friends or beloved family members (Acevedo et al. 2012; Bartels and Zeki 2000; Guerra et al. 2011; Kim et al. 2009; Langeslag et al. 2007; Tiedt et al. 2014; Zeki and Romaya 2010) and for perceived beauty, by using pictures of people who were independently rated to be very beautiful (Langeslag et al. 2007).

Mario Beauregard and colleagues (2009) take a rather different approach. They vary the instructions instead of the stimuli between experimental and control condition. Pictures are once again used as stimuli, but all pictures are of unfamiliar children and adults, always with intellectual disabilities. In the experimental condition, participants are instructed to “self-generate a feeling of unconditional love towards the depicted person.” In the control condition they are instructed to passively view similar pictures. Again, neural activity during the control condition is subtracted from neural activity during the experimental condition.

*Table 2: Tasks in Cognitive Neuroscience of Love*

Study	Stimuli	Instructions	Control task
Acevedo et al. (2012) Long-term romantic love	color picture partner	'think about (non-sexual) experiences with person'	Same instructions. Pics of: close friend highly familiar acquaintance low familiar acquaintance (all same sex and $\pm$ same age)
Aron et al. (2005) Romantic love	picture beloved (30 sec)	'think about events with person that were pleasurable, but not sexual'	Same instructions. Pics of: familiar acquaintance (of same sex and age)  countback task
Bartels & Zeki (2000) Romantic love	color picture partner (17.36 sec)	'view the pictures, think of the viewed person and relax'	Same instructions. Pics of: three friends (of same sex, similar age, and known for at least as long as partner)
Bartels & Zeki (2004) Maternal love	picture own child (15 sec)	'view the pictures and relax'	Same instructions. Pics of: other child (known to participant, same age as own child) best friend acquaintance

Study	Stimuli	Instructions	Control task
Beauregard et al. (2009) Unconditional love	pictures of unfamiliar children and adults, all with intellectual disabilities  (9 sec)	'self-generate a feeling of unconditional love towards the individuals depicted'	Similar pictures. Instructions: 'simply look at the individuals depicted'
Fisher et al. (2010) Rejected romantic love	picture of beloved rejecter that stimulates feelings of intense romantic passion  (30 sec)	'think about events that occurred with this person'	Same instructions. Pics of: familiar acquaintance (same sex, similar age and familiarity)  compared (rejecter – neutral) here to (beloved – neutral) in Aron et al (2005)
Guerra et al. (2011) Romantic and filial love	picture of boyfriend picture of father  (4 sec)	'view each picture for the entire time it is on screen'	Pics of: unfamiliar boyfriend of other participant unfamiliar father of other participant unfamiliar baby
Kim et al. (2009) Romantic love	5 pictures lover, taken at studio, different facial expressions.  (30 sec)	[not reported]	Pics of: friend (5, taken at studio) blurred human faces gray screen  Entire experiment repeated after 6 months
Langeslag et al. (2007) Romantic love	picture beloved  (0.25 sec)	'focus on the fixation cross, pay attention to the faces; you'll have to answer questions about them'  at 4 random moments between pictures, a question appeared: 'was the person on the previous picture your beloved / friend/ a stranger?'	Pics of: friend beautiful person (blended pics from beautiful people database) (both same sex as beloved)
Langeslag, Van der Veen, Röder (2014) Romantic love	picture beloved (10% of trials)  (0.250 sec)	'respond to the target stimulus by pressing a button' Accuracy is more important than speed.  [= oddball task]	Pics of: friend (10% of trials) stranger (80% of trials) (same sex as beloved, friend ± as well known)

Study	Stimuli	Instructions	Control task
Noriuchi et al. (2008) Maternal love	Video clips of own infant, either smiling or in distress. No sound.  (32 sec)	[none reported]	Video clips of four unfamiliar infants, also either smiling or in distress. No sound. (wearing same clothes as own infant)
Ortigue et al. (2007) Romantic love	name beloved as prime  (0.026 sec)	'indicate as rapidly and as accurately as possible whether or not an English word was presented'  [= lexical decision task with subliminal priming: 40 positive emotional nouns 40 pronounceable non-words 40 blank]	Control primes noun 'passion' (passionate hobby, e.g. science, sports, art) name 'friend'/acquaintance (same sex as beloved, similar age and duration of friendship)
Tiedt et al. (2014) Romantic love	picture beloved  (6 sec)	'view the photographs and imagine a comfortable, non-erotic situation related to the person presented'	Pics: 2 friends for whom no romantic feelings (yet same sex as beloved and known > 1 year, often for longer than beloved)
Vico et al. (2010) Love	5 pictures of loved ones  (4 sec / 0.5 sec)	'just look at the screen and the pictures appearing in it for as long as they are on' 'try not to blink'	Pics: 5 unfamiliar people (database) 5 unfamiliar people (loved by other participants) 5 famous faces 5 unfamiliar babies
Xu et al. (2011) Romantic love	picture beloved  (30 sec)	'while viewing the pictures, think about pleasurable and rewarding events (that you recalled in the pre-scan interview)'	Pics of acquaintance (same sex as beloved) 'think about neutral events (that you recalled in pre-scan interview)'  Countback task
Zeki & Romaya (2010) Romantic love	6-8 pictures lover  (16 sec)	'subjects were allowed to scan the images freely'  'to ensure consistent attention over time, participants were required to press a key when a circular bulls-eye prompt appeared'	6-8 pictures of friends (same sex as lover) blank screen

*Pictures were often standardized by researchers, to ensure comparability between pictures of beloveds and control pictures. Standardization may involve cropping the picture to include the*

*head only, transforming color pictures into gray scale, etcetera. Countback tasks were used in several experiments between experimental trials, to prevent spillover of feelings from one trial to the next. I have only included them in the table when neural data during countback was included as a contrast for analysis.*

### **Post-scan questionnaires**

After the actual experiment has taken place, participants are often asked for a self-report on their experiences during the experiment. Sometimes this is done in a free-style interview manner (e.g. Acevedo et al. 2012), but often participants are required to rate the extent to which they felt love on a scale (Bartels and Zeki 2000, 2004; Beauregard et al. 2009). Optionally, they also have to rate other feelings. Mostly, this is done to verify that participants actually experienced love and to exclude participants who did not, or whose feelings of love were contaminated by other emotions. For example, Bartels and Zeki excluded one mother who reported being angry at her partner at the time of the experiment, which spilled over into the feelings she experienced towards her child. Scores regarding experienced love can also be used in correlational analyses.

### **Results: neural correlates of love**

Results follow from an application of the experimental set-up as described in the methods section. The way in which they are reported is dictated by the way in which the research questions were introduced. Interpretative choice comes into play in the discussion section, where the implications of results are addressed, but not in the results section itself. Here, I just want to give readers unfamiliar with it an impression of what types of results may be reported in CNS of love.

Generally, the results of CNS studies provide information about the brain and its activity. This information about the brain is interesting in itself to cognitive neuroscientists. It does not speak to issues of self-understanding directly, though. A study in CNS of love, for example, likely reports that neural activity in the caudate nucleus and ventral tegmental area specifically correlates with looking at pictures of beloved people, but significantly less with looking at pictures of friends, acquaintances, or unknown attractive people. Several studies also find activity in the insula and anterior cingulate cortex and deactivation in amygdala (for reviews, see: (Diamond and Dickenson 2012; Ortigue et al. 2010)).

Comparing different types of love, it seems that studies of romantic love more often detect activity in the ventral tegmental area than studies of maternal love and unconditional love, whereas studies of maternal love and unconditional love find activity

in periaqueductal grey matter that studies in romantic love do not report. These results are very tentative however, particularly given the fact that both these areas are very small and it is therefore uncertain whether activity takes place exactly there instead of right next to it. All types of love appear to activate regions in the striatum (Ortigue et al. 2010).

When comparing neural correlates to love in different types of people, several studies do not detect such a difference: neural correlates to looking at pictures of a beloved so far appear the same for women and men, homosexuals and heterosexuals (Zeki and Romaya 2010), Chinese and Americans (Xu et al. 2011). It is not clear how to interpret such an absence of difference on the neural level. Null results, as absences of differences are called, cannot falsify hypotheses.

### **Discussion: from brain activity to ideas about love**

The discussion section of research articles forms a pair with the introduction section, much like the methods and results sections also form a pair. In it, cognitive neuroscientists discuss the implications of their results against the background of the research literatures to which they aim to contribute. Researchers focus on the results that speak most clearly to their question and compare these to results found in similar studies. In this process, they almost always also interpret what the neural activity may stand for in functional terms. That is to say, they form hypotheses regarding what the neural activity may have to say about love.

Up until here I have drawn out the choices researchers make in their research process: the questions they choose to focus on; ways in which they recruit and select participants; the stimuli, instructions and control tasks they employ. I have summarized some results in CNS of love and have now given a very general description of what happens in the discussion section of research articles. Cognitive neuroscientists are bound to recognize all that I say, even though they do not generally scrutinize these choices systematically. I want to devote the rest of this section to an analysis of the ways in which reverse inferences, i.e. inferences from neural activity patterns to ideas about e.g. love, influence issues of self-understanding in CNS. Many cognitive neuroscientists have heard of reverse inferences and their problems. That does not mean they understand the issues fully: I have come across excellent neuroscientists who think it is only an issue for localization research, not for research on networks; or who think it is only a problem in new fields of study, but not well-established ones. It does not mean either, unfortunately, that researchers think of the complications with reverse inferences every time they write up a discussion section to their own research articles. It remains an “unqualified truth” that “[m]any studies have drawn invalid conclusions” based on

unwarranted reverse inferences (Farah 2014, S26). What is more, even those researchers who understand reverse inferences and their complications very well, still appear to be unaware of the full implications of them for CNS' (in)ability to put existing ideas to the test. Let me unpack this.

Reverse inferences are inferences in the opposite (reverse) direction from the usual inference that can be drawn on the basis of neuroimaging research. Studies in cognitive neuroscience generally draw inferences of the type 'if cognitive process X is engaged (as operationalized through task comparison Y), then neural correlate Z is active'. For example, if participants experience love (as operationalized through viewing pictures of their beloved compared to viewing pictures of an acquaintance), then the ventral tegmental area is active. Reverse inferences are drawn in the opposite direction, from neural activity to engagement of a particular cognitive process: 'given that neural correlate Z is active in task comparison Y and that certain previous studies have found Z to be active when cognitive process X was putatively engaged, the activity of Z in the current study indicates that task comparison Y also engages cognitive process X' (Poldrack 2006). For example, given that the ventral tegmental area is active in the comparison between looking at a beloved versus a stranger, and that previous studies have found the ventral tegmental area to be active during reward processing, researchers conclude that the ventral tegmental activity in their study indicates that looking at a beloved versus an acquaintance also involves reward processing.

In this way, reverse inferences give the impression that cognitive neuroscience research speaks to our understanding of human behavior and human experience; in this case we appear to find evidence that experiencing love is rewarding. Strictly speaking, however, reverse inferences are invalid. They are an example of the logical fallacy of affirming the consequent. Many other cognitive processes besides reward processing may lead to activity in the ventral tegmental area. From the finding that reward processing (if p) leads to activity in the ventral tegmental area (then q), it therefore does not follow that when ventral tegmental activity is found (if q) this has to be due to reward processing (\*then p). Reverse inferences could be valid if it were the case that only reward processing (if and only if p) leads to activity in the ventral tegmental area (then q). Bar full specificity, activity in ventral tegmental area increases the probability of reward processing going on, though the selectivity of neural activity for a mental process determines whether this probability increase is tiny or substantial. Determining selectivity is hard, however, and large databases are compiled to this end (about which more in chapter 4). Currently, reverse inferences cannot be drawn with any substantial confidence yet (Poldrack 2006).

Cognitive neuroscientists are often careful not to phrase the main conclusions of their research articles in the form of a reverse inference. That would constitute an

unwarranted reverse inference. Instead, main conclusions are about the brain and not about behavior or experience: about finding out that activity in the ventral tegmental area and striatum correlate with experiencing love, for example. Discussion sections of research articles generally include informal reverse inferences, however (Poldrack 2006). These inferences are made in the same way, but not treated as conclusions, but as hypotheses. Researchers aim to account for all neural activity they found in functional terms, with references to other studies that found similar activity and interpreted it in these functional terms too. Academic journals in which cognitive neuroscientists publish require them to do so. This is not strange: researchers hypothesize what neural activity they expect given the task participants are required to perform. When results come in, they account for the neural activity that was actually found, both the activity that aligns with their hypothesis and the activity they had not expected beforehand. These reverse inferences are informal in that they are intended as hypotheses to be tested further.

Nonetheless, informal reverse inferences are tricky as they can easily morph into unwarranted reverse inferences. As researchers try to back up their functional interpretations of neural activity by referring to others who have interpreted the activity along similar lines, a particular interpretation that is strictly speaking informal can become so prevalent over time within the research community that it seems like fact. For example, although the presence of a dopamine-regulated network involving many midbrain areas is rather well-established, the functional interpretation of this network as ‘reward network’ is a matter of convention. Based on reviews of results (Berridge and Kringelbach 2008; Lammel, Lim, and Malenka 2014), I would say ‘pleasure network’, ‘motivation network’ or also ‘aversion network’ may be equally (in)apt functional labels. What is more, discussion sections of research articles are not only read by cognitive neuroscientists, but also by academics from other disciplines interested in neuroscience, as well as by science journalists. Often, these readers do not realize the hypothetical nature of informal reverse inferences and understand them to be research findings. Unfortunately, cognitive neuroscientists themselves do not always properly appreciate the tentative nature of informal reverse inferences either and their research communication fails to flag the informality. Particularly when communicating with audiences who may not be as interested in the ventral tegmental area as cognitive neuroscientists themselves are (e.g. the general public, or funding bodies), unwarranted reverse inferences occur frequently (Farah 2014).

Then there is the question of the basis on which informal reverse inferences are drawn in a new field of study. To take an example from CNS of love: Bartels and Zeki (2004) found reduced activity in a wide variety of regions. They divided these regions into two sets: on the one hand, areas involved in cognition and (negative) emotions

(“middle prefrontal, inferior parietal and middle temporal cortices mainly in the right hemisphere, as well as the posterior cingulate cortex”) and on the other hand, areas “associated to negative emotions and to social, moral and ‘theory of mind’ tasks” (“amygdala, temporal poles, parietotemporal junction and mesial prefrontal cortex”). They then explain the deactivations in these regions as “suppression in the neural machineries associated with the critical social assessment of other people and with negative emotions” and suggest that “these findings therefore bring us closer to explaining in neurological terms why ‘love makes blind’.”

This example illustrates how researchers can draw on common sense / stereotypical ideas about the phenomenon they are studying (‘love makes blind’) for ideas regarding how to interpret unexpected neural activity. They can also draw on psychological theories and steer interpretation towards the theory they prefer. The Aron lab, for example, uses its studies to argue that love should be conceived of as a motivational state, rather than as an emotion (Aron et al. 2005; Fisher et al. 2010). Intricately, ideas about the type of contribution to self-understanding to be expected from CNS can also guide researchers in drawing informal reverse inferences, e.g. following the schema of natural science debunking lofty images of humanity. Researchers’ personal temperament may sometimes play a role too. Helen Fisher considers love to be closely related to addiction (Fisher et al. 2010) and spends a section of her discussion focusing on CNS studies of addictions to various drugs that found neural activity in regions where her study also found neural activity. Although it may seem to an outsider (or also to a cognitive neuroscientist who is not paying enough critical attention) that these studies provide neural evidence for love making blind, for loving being a motivational state, for love being addictive, it is in fact the other way around: researchers’ ideas about love steer their interpretations of the data.

A certain amount of freedom in interpreting results functionally, i.e. in drawing informal reverse inferences, is thus present in cognitive neuroscience. To some extent, researchers can choose what activity to focus on, choose what studies to compare it with and thus how to frame and label their results functionally. Reviewers have to accept these choices for a study to get published. An overall criterion is whether researchers can refer to other studies in which the same choices were made. Even though others may come up with different theories about love, after which experiments may be designed pitting the theories against each other, the practice of having to substantiate interpretations with references to others means that there is a strong drive towards standardization of interpretations. Especially where CNS research is detached from psychological theorizing, as is the case with much of CNS of love, interpretations made in early studies become corner stones that subsequent studies repeat. The risk is that it may seem as if a body of research is compiling supporting a particular idea about love,

whereas in fact what is compiling is a body of data interpreted along the same lines by researchers, even though other interpretations could be just as valid. It may seem less and less necessary to put this idea about love (or any other process) to the test. Via this mechanism, cognitive neuroscience may unwarrantedly fortify whatever understanding researchers already held, rather than adding to our understanding.

Such pernicious effects can occur in any subfield of CNS. Socio-cultural biases come into play most strongly for topics about which researchers hold convictions as human beings too. Love is a good example of a topic that features prominently in people's self-understanding. Pernicious effects are magnified when research is disseminated to the wider audience at an early stage. Again, this is most likely to occur for topics that feature prominently in people's self-understanding. It is easy to imagine a science journalist who read Bartels and Zeki's study think up a headline running 'A brain in love makes blind' with an article below it stating that cognitive neuroscientists have now found the brain mechanisms by which love makes people blind. Thus, the research that seems most relevant to our self-understanding is ironically also most prone to be biased when disseminated to the wider public.

## **2. Conceptual implications: a passive peak experience**

Interpretation and choice play a role at various stages in CNS' research process. The previous section has shown us that they come into play in *selecting participants*; in developing *stimuli* to be used in the tasks participants have to perform; in wording *instructions* to describe those tasks; in developing *control tasks* to contrast activity during the experimental task with; and in the discussion of results, through *labelling* of results, choosing a frame of reference within which the results' implications are discussed and through both of these, by choosing what *informal reverse inference* to make about the functional implications of the results. That interpretation and choice play a role in the research process is inevitable and not a problem in itself. Each moment of interpretation has implications for the actual concept that is being studied, however. It is crucial to critically examine these conceptual implications, both to ensure that conclusions drawn on the basis of research are justified, as to be able to assess the reach of a field of CNS research and develop ideas about how to extend this reach further. Such systematic review of operationalizations and their conceptual implications happens rarely within the CNS community. As of yet, it has not happened at all for CNS of love. Let us look at some of the conceptual implications of the interpretative choices that researchers in CNS of love make.

### **Selecting participants: love as peak experience, love as easy, love as difficult**

In CNS of love, selection of participants is a moment in the research process with important implications for the concept of love under study. Studies focusing on different types of love recruit their participants in starkly different ways. Cognitive neuroscientists interested in *romantic love* solicit participants with flyers asking ‘are you truly, madly, deeply in love?’. In several studies, those potential participants are included who report the most intense feelings of being in love. The conceptual consequence of selecting participants in this way is that romantic love is treated as *a peak experience* in these studies. Even though happily married couples may also describe their experience as ‘(romantic) love’, as we can imagine Clemens to have done before the accident, they are unlikely to be allowed to participate in this research. Research on romantic love performed so far thus becomes narrowed down to research on infatuation. It is research into the neural correlates of being truly, madly, deeply in love (and limited to those people who are inclined to embrace such descriptions of their feelings).

Selection processes in CNS of *maternal love*, on the other hand, are very different. Potential participants are required to be mothers and that is it. This only makes sense under the assumption that maternal love is *automatic and inescapable*. Participants are merely required to be female and to have a young child; they are treated as if they love their child fully and consistently, i.e. all of the time. In an interview after the experiment, however, one mother reported not experiencing love towards her child during the experiment, and her data therefore was excluded from analysis (Bartels and Zeki 2004). The existence of post-partum depression and the prevalence of ambiguities in human loving more generally both call into question the assumption that all mothers cannot help but experience love towards their child, all of the time. Implicit assumptions about maternal love embedded in current participant selection procedures may stem from, or at least are in line with, socio-cultural stereotypes about pink clouds accompanying early motherhood.

The selection procedure for participants to the study on the neural correlates of *unconditional love* is intricate. Participants are recruited amongst caretakers of mentally impaired people and only those who are most capable of unconditional love according to directors of these caretaking communities, are included. Embedded in the selection procedure, therefore, is the assumption that unconditional love is *difficult* and therefore something which *requires a special aptitude and/or (professional) expertise*. This contrasts starkly with conceptual assumptions regarding maternal love, even though maternal love is also often thought of as unconditional in the sense that it does not depend on receiving anything in return from the beloved.

A further pattern in participant selection in CNS of love is that *love* is largely operationalized as *female*. In the sixteen studies reviewed here, 243 females participated compared to 84 males. An earlier review by Ortigue and colleagues (2010) lists six studies with a total of 99 female and 21 male participants. Part of this discrepancy is inexplicable. Why did some researchers recruit female participants only (Guerra et al. 2011; Ortigue et al. 2010; Vico et al. 2010)? And why did reviewers accept this? Officially, researchers should strive for an equal number of male and female participants. The informal rule of thumb in CNS is that you aim for a group of 20 participants, with a 7 vs. 13 gender distribution as the limit of acceptability. When researchers recruit their participants among psychology undergraduates, they often find it easier to find female participants, as psychology students are more often female than male. Yet in these studies, researchers do not even seem to have tried to recruit males. What is more, the discrepancy in CNS of love as a whole is remarkable, especially considering that several studies recruited their participants outside of the university.

It seems some self-selection on the side of participants has taken place. Recruitment ads and flyers asking: ‘are you madly in love?’ apparently make more women than men respond. Socio-cultural biases may be at play. Generally, it is more socially acceptable for females than for males to embrace a statement like ‘I’m truly, madly, deeply in love’ and its associations with being swept away by one’s feelings, being dependent on the beloved for one’s happiness, etcetera. Indeed, the study with the largest gender discrepancy was the one in which participants had to identify with having recently been rejected in love (dumped, basically) and being unable to move on; all of which flies in the face of the self-reliance that is to a larger extent expected of men than of women. Researchers do not appear to reflect on these issues. Wording recruitment flyers in ways that are less emotionally dramatic than ‘are you truly, madly, deeply in love?’ may help reduce the gender discrepancy for participants in CNS of love.

Finally, part of the gender discrepancy is due to a research interest in maternal love, but not paternal love. A theoretical reason for focusing on women appears to be that researchers want to compare their research on human parental love to hormone research performed on parental behavior in other species, such as rats (Bartels and Zeki 2004). There, only females have been studied and thus a comparison to human females makes most sense. Nevertheless, human male parents are clearly able to display parenting behavior and they too experience love for their offspring. Researchers have yet to include them as participants to their experiments. Socio-cultural stereotypes regarding maternal love being special in ways that paternal love is supposedly not may be partially responsible. What is more, these stereotypes may be reinforced by CNS research. Experiments may result in neural activity patterns that are specific to mothers looking at their own child versus an unknown child. Researchers may discuss this result

as potentially being the neural underpinning of what is so special about maternal love. Yet, they have not put this idea to the test. Including males in CNS of parental love is the only way to figure out to what extent neural mechanisms subserving parental love overlap or differ for males versus females. Current research runs the risk of unwarrantedly reinforce existing ideas about mother love instead of adding to our understanding.

### **Tasks: love as a passive experience**

CNS of love so far employs visual stimuli only, mostly pictures of the face of the beloved. In experiments, love is thus treated as an experience in *response to seeing* (a visual representation of) *the beloved*. Naturally, this is only one way in which we encounter our beloveds in real life. It makes pragmatic sense that CNS researchers start with visual stimuli: it is easier to show pictures to participants lying in a scanner than to present them with the smell of their beloved, for example. Further research could extend the range of stimuli, such as through confronting participants with the voice of their beloved.

In all studies in CNS of love except one, participants are instructed to relax and look at a visual representation of their beloved. Regardless of what cognitive neuroscientists themselves think of love, their choice for this operationalization has conceptual consequences. In it, the idea is embedded that love is *an experience, undergone passively*. Love is treated as an emotional experience outside of people's control: they cannot help whether they experience love or not. Let me unpack this.

First, love is assumed to be an experience, and more specifically an emotional experience, a feeling. Participants are not required to engage in any behavior towards their loved one, nor are their attitudes toward their loved one part of the operationalization. Second, love is operationalized as a passive experience: it happens to people or does not, in a way that is beyond their control. It only makes sense to instruct participants to just relax and look at pictures of their loved ones under the assumption that they will be unable to help that they experience love when confronted with the pictures and the accompanying thoughts of their romantic partner. Participants are also supposed to not be able to control what they feel in response to the pictures of others. This should not be love.

The operationalization by Beauregard and colleagues (2009) is interesting because it differs in this respect. They ask participants to self-generate feelings of unconditional love in response to pictures of unfamiliar, (visibly) mentally handicapped people and contrast this to a condition of passively viewing similar pictures. Clearly, if love were nothing but a feeling that happens to you, passive viewing should also induce

unconditional love, without someone being able to stop the experience. In fact, all participants report feeling love in both conditions, but they report more love in the self-generate condition. This tells us that, at least for these people (selected for their capacity to love unconditionally) in these circumstances (much of the time actively trying to self-generate feelings of love), they cannot help but love. At the same time however, the significant difference between the two conditions suggests that love can also be an activity rather than a passivity and that to some extent it can be under someone's own control.

The point here is not that current operationalizations are wrong or that cognitive neuroscientists are simplistic to think of love in these terms. I do not think for a second that cognitive neuroscientists believe they have captured all there is to say about love with their operationalization. Unlike Raymond Tallis, by the way, who looks at one study in CNS of romantic love and then critiques and dismisses a strawman version of it on the grounds that “as anyone knows who has been in love – indeed anyone who is not a Martian – love is not like a response to a simple stimulus such as a picture. [...] It encompasses many things” (Tallis 2011, 77). Indeed it does, but experimental research cannot proceed by manipulating ‘many things’ at once. This operationalization captures an aspect of love and researchers’ choice for it is probably guided by pragmatic concerns as much as conceptual ones. The point is that the limitations of this operationalization also limit the conclusions we can draw on the basis of the research performed with the help of it, and that cognitive neuroscientists and readers of cognitive neuroscience need to be fully aware of that. They are not conclusions about love in general. Given participant selection and operationalization practices in CNS of romantic love, for example, what is investigated in these studies is the neural activity that correlates with the love one cannot help but feel towards the person one is infatuated with, upon seeing a picture of their face.

It is inevitable that operationalizing a mental process means zooming in on an aspect of that process. This happens necessarily, throughout all of cognitive neuroscience, whether researchers investigate love or moral judgment or language learning or visual perception. It is not a problem in itself, *as long as* researchers and everybody reading the research remain aware of the limits to the conclusions that can be drawn on the basis of research. Yet what happens is that researchers write ‘romantic love’ instead of something more precise such as ‘infatuation’ or ‘the passive experience of infatuation’ in their titles, abstracts, introduction and discussion sections. To others reading their research it may thus seem as if their conclusions pertain to romantic love generally.

Second, this has crucial consequences for setting up informative review articles and databases, both of which are powerful resources in CNS. In review articles and

databases many studies are taken together. By pooling together all data from these studies, overall neural activity patterns can be drawn out that exist in spite of different operationalization details. Accurate labelling is important to ensure that similar studies are grouped together and different studies are kept apart in reviewing efforts.

Yet what happens is that in CNS of love, for example, ‘maternal love’ studies are included in a review article (Ortigue et al. 2010). Other studies exist that employ highly similar operationalizations, i.e. mothers looking at pictures of their infants, but they do not include ‘love’ amongst their keywords or in their title (Leibenluft et al. 2004; Nitschke et al. 2004). These studies thus investigate the same process, but label it differently, which means that they are not included in review articles and not pooled together in databases. Vice versa, if databases group together studies on the neural correlates of ‘love’, this includes studies on romantic love, maternal love and unconditional love, just like the review by Ortigue and colleagues (2010) does, even though these types of love may be markedly different from each other, involving different psychological processes and experiences. It is vital to be aware of these differences, for they may be accountable for differences in neural activation patterns.

Third, it is important to be aware of the conceptual restrictions resulting from operationalization to keep in mind that not all of ‘romantic love’ is already being investigated by cognitive neuroscientists, and to develop ideas of how research could be extended. Researchers could recruit participants who report ‘romantic love’ but not ‘infatuation’, for example. Or they could ask of their participants to control their feelings of love to some extent, by suppressing them under one condition and giving them free reign under another. Studies could also not focus on the experience of love so much as on love behavior or love attitudes. Yet what happens in current CNS of love (as in other subfields of CNS) is that researchers repeat an early operationalization. The first study in CNS of love, by Bartels and Zeki (2000) aimed to address love as a positive emotion and embedded its research question and findings in scientific literature on the neuroscience of emotions more generally. They devised this particular operationalization with their aims in mind. All the others that came after them just copied the operationalization, even though some are primarily interested in the motivational aspect to love (e.g. Aron et al. 2005).

All in all then, interpretation and choice play a role in CNS’ research process. This is inevitable and not a problem in itself. It is essential to be reflexively aware of the choices and their conceptual consequences, however. Currently, cognitive neuroscientists by and large are not. I return to these issues in chapter 4, when I present my ideas on how philosophy may help CNS progress.

### 3. Conclusions CNS of love

So, what types of understanding does current CNS of love contribute? What is its research process like and what role does interpretation play in it? And what consequences does this have for its contributions to our understanding of the existential aspect of the human condition?

The first main finding from this chapter is that interpretation plays an important role in the research process of cognitive neuroscience. Conceptual assumptions about love seep in through participant selection processes, operationalization of the research topic into a task with stimuli and instructions and contrasted to a control task, but also during discussion of results through labeling of results and choices made regarding what informal reverse inferences to draw. In CNS of love, interpretative choices have the consequence that love is generally conceived of as a passive, emotional experience in response to seeing a beloved. Operationalizations of different types of love embody further conceptual assumptions: romantic love is operationalized as the peak experience of infatuation; maternal love is treated as easy and automatic; unconditional love is treated as hard and requiring expertise.

To repeat, interpretation is an inevitable part of the research process and not a problem in itself. However, when people claim that the scientific method is the only way “to move a topic along on sure footing” (Gazzaniga, Ivry, and Mangun 2002, 2) or when they proclaim an opposition between “speculative philosophy” versus “well-grounded experimental disciplines” (Churchland 2008, 409), their claims have to be toned down somewhat. Any experiment’s ‘footing’ or ‘ground’ is only as good as the conceptual assumptions on which it rests. What is crucial is critical awareness of interpretative choices and their conceptual implications and appropriate communication of research results in light of them. As said, current lack of critical awareness imperils the validity of operationalizations, the validity of conclusions drawn based on experiments, and the comparability of studies to each other. With the conceptual review I have developed in this chapter, I aim to further CNS of love on these points.

Second, the main contribution to our understanding of existential selfhood that we may expect from CNS of love is insight into the neural correlates of loving. Current CNS research addresses questions about the ways in which neural activity correlates with experiencing love. Studies first and foremost aim to delineate neural activity that correlates specifically with (a type or component of) love, be it a region, a network, a hormone or any other type of neural correlate. The verdict is still out on those questions. At the same time it has to be noted that we know a lot more than we knew a mere twenty years ago. In the future, CNS of love may contribute to existential self-understanding in

those circumstances in which we are interested in our brains and in ourselves as ‘embrained’ beings (more about this in chapter 4).

A subset of questions in CNS of love revolves around comparisons between two groups of participants, to see whether they show different neural activation patterns in response to seeing their beloved. It is unclear how we should interpret the findings that no clear neural activation differences are found between, e.g., homosexuals and heterosexuals, or between people who have recently fallen in love and have been in love for a long time, or between Chinese and Americans. The absence of evidence that they are different in terms of their neural activity patterns does not constitute evidence that they are the same in terms of neural activity patterns (hypotheses cannot be falsified on the basis of null results), let alone that it would constitute evidence regarding their experiences being the same (Henson 2006).

Anna, Bob, Clemens, Deidre and Edward are not pondering their brains, nor themselves as embrained beings. They are reflecting on who they are in terms of what they love. Current CNS does not have insight to offer into love directly. For that, we would have to be able to draw conclusions about behavior and experience from brain data, that is, we would have to be able to draw inferences in the reverse direction from ordinary ones in CNS. This requires knowledge about what neural activity correlates specifically with (components of) love and only (those components of) love. Current CNS of love has not found such specific neural correlates; in fact, such a result seems far off.

Occasionally, I meet people who are incredulous that CNS of love is even possible. Surely, they say, a natural science cannot investigate something as personal and subjective and meaningful as love? Well, it can, CNS of love is on its way. One has to see it for what it is, however. Experiments in CNS of love are set up to determine what neural activity correlates with (aspects of) loving. That is, researchers in CNS of love mainly aim to determine something about the brain-in-action. Experiments are not intended to replace people’s personal stories about their love relationships, nor do researchers aim to replace philosophical analyses with their particular experiments. (Researchers often hold personal stories about their own love relationships and they are mostly blissfully unaware of philosophical analyses.) What they first and foremost aim to do is to complement existing insight with insight into the neural activity enabling loving (more about this in chapter 4). Dismissals of CNS of love, whether by philosophers or by members of the general public, are often based on wrong ideas about what it is. In that respect, they are not so different from inflated statements regarding CNS’ contributions to human self-understanding.

This brings us to the third main finding from this chapter. To many, it may seem as if CNS contributes more to our understanding of the human condition than just

insight into the neural underpinnings of behavior and experience. Several factors play a role in producing this misleading impression. The informal reverse inferences researchers draw when they discuss their findings are often taken to be research findings themselves, particularly by readers unfamiliar with CNS' reporting conventions, but sometimes also by uncritical cognitive neuroscientists. Especially when speaking to audiences not composed of people who think the ventral tegmental area is inherently interesting, cognitive neuroscientists may cross the line and unwarrantedly present hypothetical reverse inferences as research results.

What is more, this mechanism is exacerbated by incentives in reporting and reviewing practices towards repeating what others have done, both during operationalization and for data interpretation. A particular interpretation of data can become so widespread that it may seem to be a fact. The so-called reward network is an example of this. Also, pre-existing understandings of love impact research strategies and interpretations of research data in ways that make it seem as if they are corroborated by CNS research. We have seen examples of this for the idea that mother love is biologically special and for the idea that love makes blind. In chapter 3 we will see some ways in which philosophers try to deal with interpretation in their research process. In chapter 4, I develop proposals on how interpretation in CNS could be improved with the help of those philosophical strategies. For now, I would like to conclude that particularly when cognitive neuroscience understands itself to be an interpretation-free research endeavor, it runs the risk of overlooking biases and other limitations in its operationalizations and interpretations of data.

The existential crises, questions and dissatisfaction that Anna, Bob, Clemens, Deidre and Edward experience do not only rely on their ability to love, however imperfectly, but also on their ability to self-reflect, however imperfectly. Self-reflection is a topic that cognitive neuroscientists investigate as well. In the next chapter, I review CNS of self-reflection along similar lines as I have reviewed CNS of love here. CNS of self-reflection is further developed than CNS of love: more studies have been performed and more critical attention has been paid to its operationalizations. Let us see what interpretative choices have been made there, what conceptual implications these have, and what potential contributions CNS of self-reflection has to make to our understanding of existential self-understanding.

