Good practice in food-related neuroimaging

Smeets, P. A. M., Dagher, A., Hare, T. A., Kullmann, S., van der Laan, L. N., Poldrack, R. A., ... Veldhuizen, M. G.

Smeets, P. A. M., Dagher, A., Hare, T. A., Kullmann, S., van der Laan, L. N., Poldrack, R. A., ... Veldhuizen, M. G. (2019). Good practice in food-related neuroimaging. The American journal of clinical nutrition, 109(3), 491-503. DOI: 10.1093/ajcn/nqy344

Good practice in food-related neuroimaging

Paul A.M. Smeets, Alain Dagher, Todd A. Hare, Stephanie Kullmann, Laura N. van der Laan, Russell A. Poldrack, Hubert Preissl, Dana Small, Eric Stice, Maria G. Veldhuizen

Author affiliations:

University Medical Center Utrecht, Brain Center Rudolf Magnus, Utrecht, The Netherlands (PAMS)

Montreal Neurological Institute, McGill University, Montreal, Canada (AD)

Zurich Center for Neureconomics, Department of Economics, University of Zurich, Zurich, Switzerland (TAH)

Institute for Diabetes Research and Metabolic Diseases of the Helmholtz Center Munich at the University of Tübingen; German Center for Diabetes Research (DZD e.V.), Tübingen, Germany (SK, HP)

Amsterdam School of Communication Research, University of Amsterdam, Amsterdam, The Netherlands (LNV)

Department of Psychology, Stanford University, Stanford, California, USA (RAP)

Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA (DS, MGV)

Oregon Research Institute, Eugene, Oregon, USA (ES)

Division of Human Nutrition and Health, Wageningen University & Research, Wageningen, The Netherlands (PAMS)

Author's last names for PubMed indexing: Smeets, Dagher, Hare, Kullmann, Van der Laan, Poldrack, Preissl, Small, Stice, Veldhuizen.

Changed affiliations: L.N. van der Laan's current affiliation is Tilburg Center for Cognition and Communication, Tilburg University, Tilburg, The Netherlands.

Corresponding author:
Paul A.M. Smeets
University Medical Center Utrecht
Heidelberglaan 100, Rm Q2.445
3584 CX Utrecht, The Netherlands
Tel. +318875553172
E-mail**:** p.smeets@umcutrecht.nl

Sources of support: This work is supported by the European Union Seventh Framework Programme under Grant Agreement 607310 (Nudge-it) (PAMS, TAH, SK, LNV, HP).

Running head: Good practice in food-related neuroimaging

Abbreviations: BMI, body mass index; BOLD, blood-oxygen level dependent; COBIDAS, Committee on Best Practice in Data Analysis and Sharing; DEXA, dual-energy x-ray absorptiometry; fMRI, functional magnetic resonance imaging; MRI, magnetic resonance imaging; PET, positron-emission tomography; ROI, region of interest; vmPFC, ventromedial prefrontal cortex.

Abstract

The use of neuroimaging tools, especially functional magnetic resonance imaging (fMRI), in nutritional research has increased substantially over the past two decades. Neuroimaging is a research tool with great potential impact on the field of nutrition, but to achieve that potential appropriate use of techniques and interpretation of neuroimaging results is necessary. In this paper, we present guidelines for good methodological practice in fMRI studies and flag specific limitations in the hope of helping researchers to make the most of neuroimaging tools and avoid potential pitfalls. We highlight specific considerations for food-related studies such as how to statistically adjust for common confounders such as hunger state, menstrual phase, and body mass index as well as how to optimally match different types of food stimuli. Finally, we summarize current research needs and future directions such as the use of prospective designs and more realistic paradigms for studying eating behavior.

Keywords: functional magnetic resonance imaging, neuroimaging, good practice, data sharing, food viewing, food choice, taste, aroma, satiation

Outline

1. Current state of the field of nutritional neuroimaging

1.1 Introduction

The brain plays a central role in the regulation of food intake. It integrates many different state and trait-related neural and hormonal signals that affect eating behavior. Understanding how normal and maladaptive eating behaviors emerge and are maintained is crucial for developing effective eating interventions or treatments, such as weight loss or maintenance programs. Thus, studying the brain structures and processes underlying eating behavior has great potential significance, especially when combined with information on other aspects of physiology and psychology. Since the late 1990's functional neuroimaging techniques have been increasingly used to study food-related brain activity in humans. Among the first studies were taste/flavor positron-emission tomography (PET) studies (1) and functional magnetic resonance imaging (fMRI) (2) and PET studies on the effects of extreme hunger in healthy (3) and obese (4) individuals. Since then, fMRI in particular has become a widely used neuroimaging technique that is often employed to study food-related neural correlates in health and disease. We focus here on task-based fMRI, but many of the issues addressed apply similarly to resting state fMRI, PET and perfusion fMRI as well as structural MRI studies.

We present a set of guidelines for good practice in the use of neuroimaging with the hope of helping researchers make the most of these powerful, but readily misinterpreted or even misused techniques. We view the establishment of a widely accepted set of guidelines as critical at this point in the development of the field, in part because, although simple visual and motor tasks yield large, robust, and readily replicable brain responses in primary visual and motor cortex, higher order tasks

often produce smaller, more variable responses that are harder to replicate. For example, the most commonly used type of fMRI task in the food domain is the presentation of food images. Meta-analyses have shown that even the brain regions most consistently shown to differentially respond to food vs non-food images are significantly active in less than 40% of studies (5). Although brain responses to visual food cues in fasted overweight/obese participants have been found to have relatively good mean-level reproducibility, they had poor within-subject test-retest reliability (6). Another example are fMRI studies that examined the functional significance of the fat mass and obesity-associated gene FTO. Individuals with the “high-risk” AA FTO variant have been found to show less responsivity to high-calorie food images in a fasted state compared to “low-risk” TT individuals reward-related brain regions (7). Also, adults with versus without the AA genotype showed less food cue activation in the prefrontal cortex 30 min after ingesting 75 g of glucose, but no differences in a fasted state (8). In contrast, individuals with the AA or AT genotypes showed greater responsivity to food- (9) and high-calorie food images (10, 11) in reward-related brain areas than “low-risk” TT individuals.

This variability in findings is also due, in part, to divergent characteristics of the individual study designs, highlighting the current scarcity and strong need for direct replication studies. Studies of food stimulus responses and eating behavior differ in many important ways including the structure, timing and stimuli of the fMRI task; software, strategy and parameter settings used for processing and statistical analysis of the data; and individual characteristics like age, gender and eating-related traits and state variables like current hunger level and weight status. In addition, the effect size of food-related brain activation is often modest and isolating specific effects of interest can be challenging because there are many confounders and interacting

factors. For example, in a food viewing task caloric content may well covary with palatability and therefore responses to high versus low calorie foods cannot be attributed to caloric content *per se*. Further, there are clear individual differences in food preferences and familiarity that introduce additional variance (12). Thus, there is a need for better standardization of the food stimuli and fMRI task designs used and the additional data that is collected on participant's state (hunger, mood) and personal characteristics that may be used to control for confounding effects in the analyses.

In addition to the variability between studies and infrequent replication attempts, a lack of sufficient power and rigor in individual experiments is a key factor. Just as in other fields investigating higher cognitive processes, many of the earlier fMRI studies on eating behavior are underpowered (13, 14). Although there is a clear trend towards larger sample sizes in fMRI over the past decade, only recently have tools for better power calculation become available (15, 16). The need for informed study planning is further highlighted by recent empirical demonstrations stressing the importance of appropriate, validated statistical thresholding approaches (17).

Despite previous shortcomings, there is reason to be optimistic that this situation will improve in the near term. This optimism stems from the ongoing development of neuroimaging hardware and analysis software, and especially the adoption of higher quality standards in the field. We believe that replication studies and open data sharing will play a central role in the ongoing efforts to advance the utility and reliability of food-related neuroimaging findings. The current lack of replication efforts means that it remains unknown how robust many of the original findings in the field are, and although meta-analyses can give some initial indications, the accuracy of meta-analytic studies is limited by the number and quality of the primary studies they

aggregate over and is reduced by publication bias and lack of access to primary data (14). The aim of this paper is to foster good practice in food-related neuroimaging by presenting guidelines for good methodological practice, outlining potential pitfalls and providing recommendations for food-related fMRI task implementation.

1.2 What can we learn from fMRI?

FMRI usually refers to blood-oxygen level dependent (BOLD) fMRI. This popular form of fMRI exploits the fact that at a site of increased neuronal firing (brain activation), increased local blood flow leads to a decreased concentration of deoxygenated hemoglobin in the capillaries. This reduces the local distortion of the magnetic field by the para-magnetic deoxy-hemoglobin, which leads to a small increase in the fMRI signal (~0.5 – 4 %). Thus, BOLD fMRI provides an indirect vascular measure of (changes in) neuronal activity. Most fMRI studies use cognitive or sensory tasks in which different task conditions are contrasted to assess neural activation differences of interest (e.g. viewing food images versus viewing non-food images or tasting chocolate milkshake versus a control solution). This provides information on which brain regions become more or less active during a certain task (functional localization) and whether this differs between study conditions such as hunger and satiety or different groups of participants.

In recent years, there is increasing focus on (differences in) functional connectivity, that is, the degree to which task-related brain activation in a specific brain region co-varies with activation in other brain regions (functional interactions) (18). Also, 'resting-state' fMRI, which examines the spatio-temporal networks of correlated activity in the absence of a specific task (lying still with eyes closed, or mere visual

fixation) has become a popular and promising means of assessing individual differences in neurobiology (19, 20).

Brain findings per se can be useful, but often their combination with other measures creates synergy and aids the interpretation of fMRI findings; fMRI results become more meaningful when associations with physiological signals and subjective ratings or individual characteristics can be established and when they are linked to relevant outcomes such as food intake (21, 22) and weight change (23-27). Because the brain is so central in the regulation of food intake and body weight, fMRI is well-suited for connecting different levels of understanding.

Many brain imaging studies of neural response to food stimuli seek to make inferences regarding the role of neural responsivity in the development of adverse physical or mental health problems such as obesity or eating disorders. For instance, it had originally been suggested, based on the evidence that obese versus lean individuals have lower D2 receptor binding as measured by positron emission tomography, that low responsivity of reward circuitry increases the risk for overeating and consequential obesity (28, 29). However, this is an example of the complexity involved in drawing inferences from cross-sectional studies because they are unable to differentiate neural vulnerability factors from neural consequences of these physical and mental health problems.

Prospective studies that can show that the putative neural vulnerability factor predates and predicts future emergence of the adverse public health outcome permit stronger inferences than cross sectional studies. However, they do not rule out the possibility that some omitted third variable explains both the neural response and the emergence of the public health outcome. Indeed, a larger study spanning the full adult age range concluded that there was no relation between D2 receptor levels and

BMI in young adults, and a positive relationship in older individuals (30), casting doubt on the reward deficiency interpretation. Furthermore, a recent meta-analysis failed to find support for the reward deficiency interpretation as well (31). Together this work highlights the importance of prospective studies, meta-analysis, and replication in establishing reliable links between brain structure or function and eating behavior or health outcomes.

Prospective neuroimaging studies in the domain of eating behavior can vary in their breadth and duration. The most basic prospective design is to assess neural responses to experimentally manipulated stimuli or measures of brain morphometry at baseline and then test whether individual differences in these variables predict future increases in, or onset of, the health issue of interest, e.g. future weight gain or onset of obesity among initially non-obese participants. Prospective designs that include repeated-measurements of neural responses at multiple time-points provide information on biological and behavioral trajectories that can capture behavioral and neural plasticity that occurs in response to weight gain or weight loss over time (or vice versa with behavioral or neural interventions). Prospective repeated-measures neuroimaging studies of food-related behavior and health are, thus, useful for studying the mechanisms of action for prevention and treatment interventions.

Overall, neuroimaging has exciting potential to contribute to our understanding of the causes of obesity. The significant increase in the incidence of obesity over the past 50 years has been attributed to an interaction of individual vulnerability and an obesogenic environment replete with inexpensive high-calorie foods (32). Considerable evidence suggests that substantial individual vulnerability to this obesogenic environment resides in the brain. As in the mental health literature (33), the search for endophenotypes, that is, neural, cognitive or personality measures that

correlate with weight gain and BMI, has the potential to: (1) provide intermediate measures for gene discovery (2) provide explanatory mechanisms for the neural computations that lead to over-eating, and thus potentially inform the development of therapies. Moreover, the combination of endophenotype research and genetics, performed in different age-groups may allow us to disentangle the two-way relationship between body mass composition and the brain, as it is known that visceral obesity itself also causes brain changes (34), which may favor further weight gain. However, as with any measurement technique, the ultimate utility of MRI and other neuroimaging methods depends directly on the experimental designs and analysis strategies it is combined with. In the subsequent sections we highlight the importance of and aim to provide initial guidance on good practice and minimal standards in neuroimaging research with a particular focus on its application to questions surrounding dietary behavior, nutrition, and obesity.

2. Methodological aspects – good practice & minimal standards

2.1 Good practice guidelines

A carefully compiled and commonly agreed upon set of good practice guidelines is essential for maximizing the utility of the complex and ever-growing set of neuroimaging techniques available to researchers. Such guidelines facilitate the design, execution, and interpretation of original research studies, and moreover, allow for testing reproducibility, accurate replication (13, 35) and better meta-analyses. In light of the need for such guidelines, the Organization for Human Brain Mapping (OHBM) initiated the Committee on Best Practice in Data Analysis and Sharing (COBIDAS) which set out to define best practices for data analysis and results reporting as well as algorithm and data sharing to promote transparency,

reliability and collaboration. This resulted in a position paper (36) and the COBIDAS report (http://biorxiv.org/content/early/2016/05/20/054262) which provides details for proper reporting and specific good practices.

Two of the most important issues for any fMRI study are: 1) Power in terms of both the number of participants included as well as the task design (e.g. number of trials per condition), and 2) The threshold used for assessing statistical significance and how that was determined, appropriately controlling for multiple comparisons. These comparisons include the testing of multiple voxels and/or regions of interest, but also extend to tests of neuroimaging measures against multiple measures of individual differences in cognition or health status. The following sections will cover multiple aspects of how these general guidelines can be applied to neuroimaging studies of dietary behavior, nutrition, and obesity. After briefly summarizing general good practice guidelines for neuroimaging, we discuss specific experimental design and analysis features for studies using visual, olfactory, or physical foods/liquids as stimuli. We would like to note that the AJCN is committed to the COBIDAS standard and encourages authors to follow the recommendations of that report. Upon submission, authors will be asked to complete a checklist based on Appendix D of the COBIDAS report. All items flagged as mandatory need to be satisfied as a minimal standard. This checklist is available as Supplemental Checklist S1.

2.2 Power calculation and study planning

The prevalence of underpowered studies in neuroimaging, as well as many other scientific disciplines, is one of the biggest, but also most concretely addressable issues we face (14, 37, 38). Power analysis is important not simply to avoid performing a futile study, but also to ensure that any positive findings are likely to be

249 true positives; as noted by (37), low power increases the likelihood that any positive
250 findings are false positives and thus reduces the likelihood that findings from
251 underpowered studies are replicable. To date, sample size calculations based on
252 realistic power analyses have been made only rarely during the planning stages of
253 fMRI studies. At least, such calculations are rarely reported in literature. This
254 shortcoming is by no means specific to the use of fMRI for nutrition research, but is
255 nonetheless a serious limitation and often results in inconclusive, non-replicable, or
256 even misleading findings. We now know that common rules of thumb about statistical
257 power for fMRI studies (e.g. 20-30 participants per group) do not hold in many cases
258 and often result in underpowered studies, particularly when the goal is to examine
259 individual differences (39). Underpowered studies are most often a waste of funding
260 as well as the time and effort of both researchers and study participants (38, 40).
261 Making realistic power calculations requires careful thought and effort, but the
262 necessary tools for doing so are available. Most statistical software packages include
263 dedicated functions for power analyses. Moreover, in recent years, more accessible
264 and fMRI-specific tools, e.g. (15, 41), have been developed to help researchers make
265 appropriate power calculations that incorporate both within and between subjects
266 factors. It is important to remember that power is a function of the number of
267 participants, but also the heterogeneity of the study population and the amount and
268 quality of data collected per participant. In conjunction with sample size calculations,
269 it is important to optimize the design of fMRI tasks in terms of the number, temporal
270 distribution, and duration of different trial types (for general guidelines see
271 http://imaging.mrc-cbu.cam.ac.uk/imaging/DesignEfficiency; for an example tool for
272 testing efficiency of an fMRI task design see http://www.neuropowertools.org/).

The ever-growing number of studies in the literature and the move toward open data sharing means that in many cases, data are readily available for use in making estimates of power and requisite sample sizes for new studies. However, it should be noted that effect sizes based only on published studies are likely to be inflated due to publication bias. Therefore, the use of existing data should generally be complemented by piloting the exact experimental procedures. In many cases, researchers and funding agencies will still need to invest significant time and resources into collecting more specific pilot data to make realistic power calculations. However, the returns on such initial investments are worthwhile, and the cost of not conducting appropriate study planning is far greater.

Lastly, we note that collecting more data (trials or subjects) is not the only way to improve statistical power in fMRI research. The traditional method for analyzing fMRI data (i.e., the mass univariate approach) involves the repeated testing of a regression model in tens or hundreds of thousands of individual voxels. These multiple tests require corrections for multiple comparisons that reduce statistical power. These corrections are necessary for valid inference and cannot be avoided for mass univariate analyses. However, mass univariate analyses are only one means of analyzing fMRI data (42). Multivariate analyses (43) and data reduction or aggregation techniques such as independent or principle components analyses, or predefined regions of interest (ROIs) reduce the number of comparisons conducted and thus the degree of correction required (e.g. p/10 rather that p/50,000). Beyond simply increasing power, there is ample reason to believe that multivariate and network-level analyses (44, 45) provide additional insight into brain function and the application of such techniques to the domain of food choice and nutrition represents an important, and as yet, relatively under-exploited opportunity.

298

299 *2.3. Proper experimental and task design*

300 Eating behavior and nutritional decisions are determined by a plethora of factors. In
301 order to draw strong conclusions from neuroimaging results, we have to know
302 precisely which factors were controlled and which were manipulated. The nature of
303 the scientific question will determine exactly which geno- and phenotypic information
304 is most appropriate to measure or manipulate and report. It is now standard to report
305 body mass index (BMI) as an anthropometric measure, age of the participants and
306 sex. However, for many specific questions a deeper phenotyping may be necessary.
307 For example, it is clear that BMI does not provide enough information concerning
308 body composition (46). Better methods to describe the body composition are bio-
309 impedance measures, DEXA, MRI, or BOD POD® assessment of body composition.
310 However, the method used for a given study should be appropriate for the aims of the
311 study and justified in terms of costs and benefits to both researchers and participants.
312 Ideally, however, there should be overlap in the measures used to allow better
313 accumulation of evidence. Accordingly, a set of high-priority measures, including
314 MRI, has been proposed to achieve common usage and thereby increase the
315 breadth and impact of obesity research (47).

316

317 2.3.1 Hunger state and related factors

318 An important factor to control in nutritional studies is hunger state and caloric
319 deprivation because they affect food wanting and food-related brain responses (3,
320 48-51). In addition, the quantification of food intake is especially important for
321 intervention studies, because nutritional composition can also affect neuronal
322 processes. For example, fasting state studies generally require a 12-hour fast and try

to control for the subjective hunger state using visual analog scale measures of appetite. However, it has been established that macronutrient composition of even a single meal can affect hormonal responses extending beyond 12 hours (52). Thus, there is added value in the assessment, and inclusion as covariates in analyses, of major hormonal factors related to nutrition. E.g. glucose, insulin, leptin and ghrelin could be included for nutritional studies of neural responses in specifically induced feeding states such as hunger versus satiety. This would allow researchers to disentangle physiological and subjective factors related to eating processes.

Another issue is that nutritional preferences are culturally and individually determined, and therefore, the creation and use of standardized food stimuli can be difficult. Moreover, these evaluations are time of day, season and (hunger) state-dependent. For example, a heavy breakfast with savory components is very uncommon in many parts of the world and if studies are performed during the morning hours this has to be taken into account. Thus, acquiring individual evaluations of the experimental stimuli is another standard operating procedure that should be incorporated into neuroimaging studies of nutrition-related behavioral or physiological responses. In addition, it is advisable to use a standardized meal e.g. on the evening before the measurement or at least to request participants in a repeated measurements design to consume the same meal preceding all measurements.

Finally, an important challenge in all nutritional studies, including those using neuroimaging, is that the assessment of nutritional intake is difficult to quantify in normal daily life. Currently, most studies use diaries for nutritional intake. However, such self-reports are unreliable (53). There are several ongoing efforts to measure nutritional intake using smartphone applications. However, an assessment of the

validity and degree of advantage or disadvantage of smartphone-based methods relative to traditional diary methods and the doubly-labeled water method for assessing habitual caloric intake will require further study.

2.3.2 Personal characteristics

In addition to physiological factors, care must be taken to account and, whenever possible, control for psychological factors in studies of the neurobiology of eating behavior. Personality or cognitive traits may modulate food-related brain responses (12).

Most studies test for eating disorders, to exclude clinically relevant diseases. However, it would be advisable to statistically control for subclinical scores on eating disorder scales.

2.3.3 Choosing and matching food-related stimuli

Eating engages all of our senses. The extra-oral sensations of vision and olfaction provide information about food availability to guide food acquisition. The oral sensations of somatosensation (e.g. texture and temperature), chemesthesis (e.g. astringency, spiciness) gustation (sweet, sour, salty, bitter, umami and possibly fat and starch taste) and retronasal olfaction provide information to guide consumption once the food is acquired and in the mouth. For example, one uses oral somatosensation to localize a bone in a bite of fish that needs to be extracted before swallowing while the taste of sweetness produces a metabolic cascade to facilitate glucose metabolism (54). The choice of stimulus will depend upon the particular goals of the study. An in-depth discussion of relevant factors to consider for visual,

372 olfactory and oral food-related stimulation is provided as Online Supporting Material
373 S2.
374
375 *2.4 fMRI data analysis*
376 2.4.1 Statistical thresholding for whole-brain and region of interest (ROI) fMRI
377 analyses can be performed at several levels. When using common mass univariate
378 approaches that take all voxels in the brain into account, appropriate corrections for
379 multiple comparisons must be implemented. This has been noted early on (55) but
380 was highlighted several years ago by a conference paper reporting on scans of a
381 dead salmon who was instructed to perform an emotion recognition task (56). When
382 appropriate correction techniques were not applied, there appeared to be task-
383 related brain activation in the salmon. Naturally, these false-positive activations were
384 no longer seen when appropriate corrections for multiple testing were used.
385 This infamous “case study” is a salient reminder of the importance of employing
386 appropriate statistical methodology in the analysis of neuroimaging data. In many
387 subfields of neuroimaging, it has been commonplace to use rule-of-thumb corrections
388 for multiple comparisons (e.g. a voxel-level threshold of $p < 0.001$ uncorrected
389 combined a cluster-extent size of 10 voxels). However, it is now clear from creative
390 examples like the salmon study and more rigorous and extensive investigations that
391 such rules are inadequate in controlling false-positive rates. Recent comparisons of
392 correction methods for multiple testing in fMRI data indicate that permutation-based
393 procedures are the best choice and that cluster-based methods should be used
394 correctly (17, 57). Specifically, when Gaussian random field theory is used for cluster-
395 based inference, the cluster-forming threshold should be $P = 0.001$ to avoid inflated
396 false-positive rates (17). More stringent cluster forming thresholds also help to avoid

problems in interpreting the very large activation clusters that often result from low cluster forming thresholds (57). Note that cluster-based corrected findings indicate that there is likely to be significant activity somewhere within the cluster rather than indicating that all voxels within the cluster are significant. Thus, if we only show that somewhere within a very large cluster there is probably a significant difference between conditions or groups, then we cannot infer or conclude much at all.

In addition to whole-brain analyses, the current literature on the neurobiology of nutrition is substantial enough to justify region of interest (ROI) analysis for certain brain regions or connections between regions. However, in order to be valid, ROI analyses must be planned a priori, ideally preregistered, and the hypotheses about the region must be clearly stated. To avoid biased results, both anatomical and functional ROIs should be defined based on *independent* datasets or functional localizer tasks. Note that multiple comparison corrections must be applied across the ROIs when multiple ROIs are tested for a given hypothesis. Furthermore, the assumptions underlying cluster-based correction methods are rarely satisfied in small volume analyses and their use in this case should be avoided (58).

2.4.2 Minimizing the influence of movement

FMRI data are prone to movement-related artefacts because movement causes displacement and distortions in the data. In particular oral stimulation can be accompanied by significant movement. Movement from swallowing and other activities like breathing may be larger because of greater body mass. Additionally, there is evidence that head motion and BMI share genetic influences, suggesting that movement is a neurobehavioral trait that is greater in obesity (e.g. (59)). These

movements can be counteracted in real time or modeled post-hoc during data analysis.

a) Real time. Movement can be minimized physically by the use of cushions around the head, a personalized head case from a two-part foam mold or a bite bar. Movement can also be minimized through behavioral training or feedback. One way is to provide the participant with a stationary reference, which has been done by using a cloth strap or tape across the forehead that attaches to the head coil. When the participant moves they can clearly feel this by the friction on their forehead. This feedback works well, and leads to substantial improvement because movement from swallowing mostly results in small movement in the z-plane which is hard to feel in most head coils. Again, training is important to improve comfort and ability to lie still. Training will also allow participants to learn to swallow with minimal movement of the head, by isolating movement to the jaw and tongue during swallowing. The use of real time feedback with a head motion tracker in a mock scanner may be most efficient (available for example at Psychology Software Tools https://pstnet.com/). Scanners with newer software may include real time monitoring of movement and allow experimenters to immediately redo runs that invoked too much movement. Another solution is to remove the need to swallow altogether by suctioning out liquids (60) or instructing participants to hold the liquid in their mouth until they receive a cue to swallow (61). The downside of these methods, as elaborated in Supplemental Material S2, is that a large area of stimulation is overlooked, that aromas in flavors cannot be perceived and that an important part of the process of ingestion is omitted.

b) Post-hoc analysis. Correction for head motion via image registration is performed as a standard part of the fMRI preprocessing pipeline, but it is clear that this is not sufficient to remove the residual effects of head motion on image

intensities (62); for this reason, motion parameters and their derivatives (which quantify change from time point to time point) are often included as nuisance regressors in the statistical model. However, these too may be insufficient to address large amounts of motion, and it is common to reject data from individual participants, runs, or time points based on motion estimates. The state of the art techniques for motion detection and cleaning have been developed in the context of resting state fMRI, where head motion is a critical problem (63). In addition to use of motion estimates and their derivatives as nuisance regressors, it is common to compute a measure of “frame wise displacement”, which measures the overall displacement of the images between each pair of subsequent time points, and a measure called DVARS which quantifies the mean change in image intensity between time points. These measures may be used to “scrub” time points with motion that exceeds a particular threshold (varying from 0.2 to 0.5 mm frame wise displacement (64)) along with surrounding time points; in the context of task-based fMRI analysis, this scrubbing can be performed as part of the statistical model by including single time point regressors for each excluded time point in the model (65). Individual runs or subjects exceeding a threshold level of scrubbed volumes may be dropped; the use of faster imaging with multi slice acquisition can improve the handling of motion by reducing the relative amount of data that needs to be removed.

An estimate of vigor of swallowing and exact timing of swallowing may be obtained with expanding bellows and a spirometer (66, 67), which will allow using swallowing as either the onset of an event-of-interest or, alternatively, as a nuisance regressor to be covaried out. Similarly, movement from breathing can be estimated with most standard scanner equipment and incorporated into the single-subjects analysis. These variables can also be included as regressors in group analyses to address

their confounding effects. Finally, independent component analysis can be used to remove the effects of motion artifacts and physiological noise from breathing and heart beating (68-70).

2.4.3 Analysis of prospective designs

Although significant advances have occurred in analytic approaches for longitudinal data that better account for auto-correlation of data from the same participant over time, missing data, and nested data (71), these advances are not supported by commonly used fMRI analytic packages. The most basic approach if the data are only collected at two time points is to use change scores for the outcome (e.g., T2 BMI – T1 BMI) and simply regress the change scores on BOLD response from the contrast of interest (e.g., (72)). However, it is critical to covary for baseline BMI because change in an outcome over time is typically negatively related to baseline values of the outcome (73). Ideally, we recommend using random effects growth mixture models, or other types of hierarchical linear models that use full information maximum likelihood to confirm that we model change in behavioral outcomes optimally. This is particularly important when data are collected at 3 or more time points, as there is the potential for non-linear change over time (e.g., quadratic growth). The slopes and intercepts (coded to reflect baseline values) can then be exported to any of the standard fMRI analytic statistical packages, and the slopes regressed against the BOLD response, controlling for the intercept (e.g., (27)). For repeated-measures studies, which can include natural history observational studies (e.g., (74)) or intervention trials (e.g., (75)), one can simply use repeated-measures ANOVA models to test for differential change in BOLD response in contrasts of interest over time across two or more groups. Although one might be tempted to

directly contrast BOLD response to the event of interest (e.g., taste of milkshake) from multiple assessment points, we do not recommend this approach because a number of factors can contribute to variation in BOLD signal over time (e.g., variability in physiological variables, instability of MRI hardware), which may introduce bias. Instead, the contrast of the event of interest against an appropriate control event (e.g., tasting tasteless control solution) should be used. An alternative approach is to read out parameter estimates from the contrast of interest at each assessment and use standard data analytic packages, such as SAS or R to conduct regression models or repeated measures analyses, but this requires a ROI approach, which does not make use of all the data collected and may miss important peaks that were not anticipated *a priori*.

2.4.4 Predictive modelling

One of the potential uses of MRI is the prediction of future outcomes, such as eating behavior, weight change or treatment responses. A mounting number of studies suggests neural food cue reactivity can predict outcomes like energy intake outside the lab (76), weight gain (27, 77, 78), weight variability (79) and weight loss success (23, 80).

However, care must be used during model fit in order to achieve predictive accuracy on new samples. When model fit and goodness of fit estimates are obtained from the same data, the estimated goodness of fit is inflated because the data have in a sense been used twice (81). One approach to address this is to use cross-validation to assess out-of-sample predictive accuracy; in this method, the model is fit iteratively to subsets of the data and tested on the remaining data that were held out during training (https://web.stanford.edu/~hastie/ElemStatLearn/). This method provides

more accurate estimates of how well the model can predict outcomes in new samples, however, predictive accuracies can be highly variable with small samples (82), and accuracies can be inflated if many different parameter sets are tested without proper control (83). For this reason, testing a model (e.g. regression, support vector machine, etc.) fit to one dataset against an entirely separate and independent dataset remains the gold standard for quantification of predictive accuracy.

2.5 Preregistration and data sharing

The importance of transparency for reproducible research is increasingly realized. Studies can be registered at accredited public trial registries like clinicaltrials.gov, but that does not preclude exploration of the data beyond the testing of the primary hypotheses, although study plans including planned analyses can also be pre-registered e.g. at the Open Science Framework (osf.io). To counter publication bias there is an increasing number of journals that accepts registered reports; the study plan is peer-reviewed and if accepted, the journal will publish the results of the planned analyses regardless of their nature (see https://cos.io/rr/)

Transparency and reproducibility is further aided by the sharing of research materials such as task scripts and analysis code as well as the data. There is a spectrum of data sharing, which involves a tradeoff between the ease of sharing and the utility of the data (84). On the one hand, meta-analysis has largely relied upon activation coordinates from published papers (85, 86) which are easy to obtain but limited in comparison to meta-analysis based on full statistical images (87). For this reason, it is now recommended to share the unthresholded statistical images from neuroimaging studies using a database such as Neurovault (88). At the other end of the spectrum is the sharing of complete raw datasets, via resources such as

OpenNeuro, INDI/FCP, and NITRC. The sharing of raw datasets requires substantially more time and effort than sharing of coordinates or statistical results, but provides greater utility of the data, such as allowing different analyses to be applied to the same data, or allowing raw data to be combined across studies in a "mega-analysis." Recent projects such as the Human Connectome Project (89) and ENIGMA Consortium (90) have demonstrated the substantial utility of sharing large samples of raw MRI data.

3. Appropriate interpretation

3.1 What can be concluded from fMRI findings (and what not)?

Although research on the exact meaning of changes in the BOLD fMRI signal is still ongoing, most researchers assume that differences in BOLD signal reflect differences in neuronal activity 'averaged' over the piece of brain tissue that was sampled (voxel). One could argue that as long as we can detect apparently meaningful differences between conditions or groups that BOLD fMRI is of use, regardless of the exact underlying neuronal and physiological correlates of these signal differences. Nevertheless, underlying processes such as coupling between neuronal and vascular response may differ between subjects, and may be affected by disease states. Notably, obesity is associated with increased cerebrovascular disease risk and this may affect neurovascular coupling (91). Studies examining cerebrovascular reactivity can be used to assess whether this might be a problem in specific study populations.

A particular point of attention for clinical and intervention studies is that baseline or 'resting state' brain activity may differ between patients and controls or may change

571 due to the study treatment (e.g. meal ingestion or a diet intervention). This may
572 explain observed differences in task-related brain activation, which is usually the
573 main outcome parameter. In addition, because fMRI results usually rely on a
574 comparison between two task conditions or groups the direction of the underlying
575 BOLD signal changes should be examined by extracting cluster parameter estimates
576 to aid interpretation. This allows one to distinguish less deactivation from greater
577 activation, for example. Group x task condition interactions should be reported only
578 where there is a main effect of the task in one of the groups. For example, when
579 there is no clear activation in a region for "food versus non-food" great caution should
580 be exercised in reporting and interpreting a group x stimulus type interaction in this
581 area.
582 It can be challenging to design an fMRI task such that a specific cognitive process is
583 subtracted out by contrasting a task of interest with a control condition. First, in the
584 food domain in particular it is inherently harder to match stimuli due to their sensory
585 complexity and possible cognitive associations and we can only approximate control
586 conditions by matching on as many characteristics as we can. Second, the observed
587 differences in regional brain activation may be driven by associated but not
588 necessarily food-specific processes like arousal, attention, emotion or motivation.
589 This is not necessarily a drawback, but it is important to be aware of this. Third, fMRI
590 is sensitive such that task instructions and mind set or attentional focus can alter the
591 pattern of brain activation observed (see e.g. (92-95)). Thus, when interpreting
592 findings and comparing with the literature it is important to take seemingly minor
593 differences in task design and instruction into account.

As alluded to before, conclusions can be strengthened by showing that differences in BOLD signal changes correlate with relevant parameters like stimulus or personal characteristics.

3.2 Reverse inference

A common practice in the interpretation of neuroimaging results is the use of reverse inference (96). This refers to interpreting activation of a particular brain region as evidence for the engagement of a particular cognitive process. Although they can provide some information, such inferences are not deductively valid and need further substantiation. In particular when activation of a brain region cannot be pinpointed to a specific process or when evidence for selective engagement of that region during a specific neural process is weak, reverse inference should be done with caution. E.g. areas that are often found to be activated in many studies, also outside the food domain, are the insula, cerebellum and prefrontal cortex (97, 98). For such large and heterogeneous regions special care should be taken to consider the exact subregion found in combination with the process of interest. In conclusion, reverse inference should be used with caution and involve as much specificity as possible.

3.3 Comparability of findings in 'the same' brain region

In general, the discussion of fMRI findings often lacks accuracy. Often it is unclear whether the area being discussed is really in the same part of the larger structure, say within a 10-mm radius, and located in the same hemisphere. This may be particularly true for large areas such as the insula and long gyri, e.g. the inferior frontal gyrus. It is advisable to be as specific as possible e.g. by distinguishing between anterior, middle and posterior insula. Likewise, indicative labels such as

619 ‘dorsolateral prefrontal cortex’ or ‘ventromedial prefrontal cortex’ may be used to refer
620 to very different locations. Thus, in all cases comparison of findings between studies
621 should not be done without checking the exact location to allow appropriate wording
622 of the degree of similarity. In addition, it is important to be clear on the paradigm or
623 other relevant aspects of the study such as the sample size or population used,
624 which can significantly affect comparability of findings and thus the strength of the
625 inferences made. We see the open sharing of un-thresholded group-level statistical
626 maps, e.g. through Neurovault.org, as the most promising way to resolve such
627 ambiguities. If these data are available for all published studies then comparing the
628 spatial locations of new and existing findings becomes as simple as overlaying two or
629 more maps.
630 A useful approach to overcome regional/functional imprecision is to use meta-
631 analytical results to pinpoint functional areas. Online repositories of meta-analyses
632 such as the ANIMA database (99) or Neurosynth (www.neurosynth.org) (86) can be
633 queried to identify specific functional locations e.g. the vmPFC area that encodes
634 stimulus value or the insular subregion that responds to taste stimuli.
635

636 ## 4. Research needs and future directions

637

638 *4.1 Fostering comparability, data pooling and meta-analysis*
639 Scientific progress can be promoted by better comparability of research findings,
640 allowing better data pooling and more accurate meta-analyses. This requires better
641 standardization of (neuroimaging) methods and associated measures, along with the
642 application of advanced analysis and modelling techniques to nutritional
643 neuroscience data (100). This would be aided by minimal standards in the field as to

644 which descriptive data must be reported, in addition to common descriptives such as
645 age and gender. This might include as a minimum handedness, BMI and a measure
646 of hunger state but could be expanded for many studies by additional measures such
647 as information on diet, body composition (body fat %), hormonal status (menstrual
648 cycle phase, appetite-related hormones) and personal as well as personality
649 characteristics (dietary restraint, food attitudes, reward sensitivity, impulsivity).
650 Task-related fMRI studies would do well to use established paradigms with
651 standardized stimuli adjusted for the population under study and also evaluated by
652 the study participants to confirm e.g. familiarity. This is aided by sharing of the stimuli
653 used in online databases (see Supplemental Table S1 in Supplemental Material S2)
654 and sharing of the associated task paradigms and code, preferably at established
655 repositories like the Open Science Framework (https://osf.io/) and GitHub.
656 An excellent way to make more of existing data or achieve greater yield from studies
657 is to employ the same paradigm and analysis pipeline across many centers. This is
658 particularly useful when it concerns specific (clinical) populations that may be hard to
659 recruit in sufficient numbers by a single center. An example of this are the ENIGMA
660 (90) working groups that assess cortical thickness for different disorders by pooling
661 results obtained from the analysis of anatomical MRI scans from many centers
662 (http://enigma.ini.usc.edu/). While mainly focused on brain disorders so far, there is
663 an eating disorder group as well (http://enigma.ini.usc.edu/ongoing/enigma-
664 anorexia/).
665 Another noteworthy initiative is the use of standardized analysis pipelines for
666 neuroimaging data analysis (101) as provided at the OpenNeuro platform
667 https://www.openneuro.org/. This may help to reduce variation in study results and
668 allows researchers to see how robust their outcomes are, when assessed with

different software packages. As a minimum, (neuroimaging) analysis scripts should be shared alongside data to better allow replication by others.

4.2 Toward predicting future outcomes

The vast majority of nutritional neuro-imaging studies are cross-sectional. As alluded to above, to learn more about the causality of obesity and eating disorders it is crucial to promote long-term follow-up studies, e.g. by adding MRI measures to adequately powered cohort studies. Adding to existing or newly formed cohorts would also ensure detailed phenotyping. Individual differences in fMRI task responses or structural data at baseline can then be used to predict future changes in relevant outcomes such as onset of a disease state or growth in symptoms (see e.g. (27)). Ideally, phenotyping including neural measures would be done repeatedly to be able to examined neural plasticity that may occur in response to (nutritional) interventions or disease conditions (e.g., onset of an eating disorder or obesity).

4.3 Technological advances

4.3.1 More realistic food cue exposure and choice context – potential of virtual reality

Another direction for future work is the development of more realistic fMRI paradigms, which better reflect the reality of food cue exposure and choice. A supine-positioned, immobile participant lying in a narrow, noisy MRI tube, located in a hospital, might reasonably be expected to behave differently than one walking around a supermarket or sitting at the dinner table. There is ample evidence that situational factors influence momentary goals and preferences and thereby food choice (102, 103). For example, in-store communication and cues at the consumption

site can trigger hedonic- or health-related goals and thereby steer choices towards goal-congruent alternatives (104-108). The abovementioned contextual cues, which are normally present at the point of purchase, are lacking in most fMRI studies. However, possibly more problematic, situational factors in the fMRI research setting, like seeing medical equipment, might activate associated information (i.e., thoughts about disease, medical treatments) and influence current goals (e.g., prevention of disease) itself and thereby influence behavior. It is unknown how the presence of medical equipment influences food choice and underlying cognition and this is a relevant topic for further study. Further, given the strong effects of situational factors on choice and potentially on the neural processes leading to that choice, it is important that authors describe the complete study setting with a high level of detail. For example, it should minimally be mentioned whether the experiment was carried out at a hospital or at a research-dedicated MRI scanner in a non-medical facility.

Aside from these situational factors, functional MRI food choice tasks are generally highly simplified, showing (cut-out) images on a plain background, and are thus very different from the real-life food choice environment (109-113). Situational and task-related factors combined might result in very different choices in fMRI research than in real life. If choice behavior differs between fMRI tasks and real life how can we be confident that the cognitive process we measure during choice is the one we actually aim to measure? So far, to our knowledge only a few studies have related choices made in the scanner to a 'real-life' measure of eating behavior, namely intake at a subsequent *ad libitum* lab buffet meal (114) and intake at a buffet lunch the next day (115). In the former study, however, in-scanner choices were not related to intake at the buffet. To assess how representative food choice behavior in fMRI tasks is for real-life food choices future studies should incorporate real-life measures of eating

719 behavior and relate these to in-scanner behaviors. This will allow us to establish the
720 need for more realistic fMRI food choice paradigms.

721 One approach to develop more realistic fMRI paradigms is by using virtual reality
722 (VR). VR provides the ultimate level of immersion, creating a sense of physical
723 presence in the 3D virtual environment. VR has been successfully applied in a wide
724 range of fields including psychiatry and medicine (116, 117). Moreover, in the past
725 years, several virtual supermarkets have been developed (118-120), which enables
726 collection of purchase data in a very controlled yet realistic environment. VR has a
727 major potential for use in neuroimaging food choice research because individuals
728 quickly feel 'embedded' in VR environments, such that the actual situation (lying in an
729 MRI scanner) is suppressed in favor of the virtual situation (walking in the
730 supermarket) (121). Several studies have shown that purchasing behavior in virtual
731 supermarkets is relatively similar to actual purchase behavior (122-125). However,
732 increased realism might come at the cost of increased noise and excessive visual
733 stimulation which might decrease sensitivity to detect signals of interest. To our
734 knowledge, to date only one virtual supermarket paradigm that can be used in fMRI
735 research has been developed
736 (http://nutritionalneuroscience.eu/index.php/resources/neuroshop-virtual-
737 supermarket). In this paradigm, participants can first freely navigate through the
738 virtual supermarket with a joystick. This serves to embed the participant in the virtual
739 supermarket and foster involvement in the task of grocery shopping. Subsequently,
740 participants perform a more standardized fMRI choice task in which shelves with the
741 same design are shown and choice blocks are interspersed with movies of walking
742 around from shelf to shelf, in order to maintain embedding. This provides a first step

743 towards exploiting the potential of virtual reality to produce more ecologically valid
744 measures of food choice and underlying neural processes.

745 4.3.2 More realistic feeding paradigms

746 To better mimic ingestive behavior there is need to move beyond stimulation with
747 passive reception of small boluses of liquid. The major hurdle here has been the
748 sensitivity of fMRI to movement. However, recent advances in hardware and software
749 offer hope that sequences can be compiled that will be more robust and perhaps
750 even allow us to measure responses to active sipping, swallowing, and even chewing
751 solid foods. For example, multi-echo fMRI increases the signal to noise ratio by a
752 factor of 4 (126), while multiband acquisition provides enhanced speed to increase
753 the temporal resolution allowing greater ability to deconvolve the BOLD response in
754 the context of movement. Also in development is echo planar imaging with the
755 “keyhole technique”, which increases the signal readout even further allowing 25-
756 30% increases in either spatial or temporal resolution. These improvements in data
757 acquisition can then be coupled to new technology enabling delivery of solid foods to
758 participants lying in the scanner bore. Although there is some way to go and chewing
759 poses additional risk for movement artefacts as well as aliasing of activity from the
760 temporalis muscles, such technologies are on the horizon (127).

761

762 **5. Conclusions**

763 The potential of functional neuroimaging for leveraging our understanding of the
764 drivers of eating behavior is substantial because it can elucidate the underlying
765 neural processes and how these are affected by the diverse determinants of eating
766 behavior. However, to maximize the yield of neuroimaging methods it is of paramount

767 importance to adhere to high standards in terms of experimental and task design and
768 subsequent data analysis to ensure sufficient detection power, specificity and
769 interpretability. To accommodate the complexity of nutrition research and to be able
770 to distinguish noise from meaningful variability, the use of standardized methods,
771 proper phenotyping and reporting of sufficient methodological detail are necessary to
772 enhance data pooling and meta-analyses of nutritional imaging data. Moreover, there
773 is a need for more prospective and repeated measures studies to elucidate etiology
774 and establish neural markers so as to provide novel and specific targets for
775 intervention.
776

777 **Acknowledgments**

778 None of the authors has a potential conflict of interest. All authors designed and
779 wrote the paper. PAMS had primary responsibility for final content. All authors read
780 and approved the final manuscript.

References

1. Small DM, Jones-Gotman M, Zatorre RJ, Petrides M, Evans AC. Flavor processing: more than the sum of its parts. Neuroreport 1997;8(18):3913-7.
2. Cerf B, Lebihan D, Van de Moortele PF, Mac LP, Faurion A. Functional lateralization of human gustatory cortex related to handedness disclosed by fMRI study. AnnNYAcadSci 1998;855:575-8.
3. Tataranni PA, Gautier JF, Chen K, Uecker A, Bandy D, Salbe AD, Pratley RE, Lawson M, Reiman EM, Ravussin E. Neuroanatomical correlates of hunger and satiation in humans using positron emission tomography. ProcNatlAcadSciUSA 1999;96(8):4569-74.
4. Gautier JF, Chen K, Salbe AD, Bandy D, Pratley RE, Heiman M, Ravussin E, Reiman EM, Tataranni PA. Differential brain responses to satiation in obese and lean men. Diabetes 2000;49(5):838-46.
5. van der Laan LN, de Ridder DT, Viergever MA, Smeets PA. The first taste is always with the eyes: a meta-analysis on the neural correlates of processing visual food cues. Neuroimage 2011;55(1):296-303.
6. Drew Sayer R, Tamer GG, Jr., Chen N, Tregellas JR, Cornier MA, Kareken DA, Talavage TM, McCrory MA, Campbell WW. Reproducibility assessment of brain responses to visual food stimuli in adults with overweight and obesity. Obesity (Silver Spring, Md) 2016. doi: 10.1002/oby.21603.
7. Karra E, O'Daly OG, Choudhury AI, Yousseif A, Millership S, Neary MT, Scott WR, Chandarana K, Manning S, Hess ME, et al. A link between FTO, ghrelin, and impaired brain food-cue responsivity. The Journal of clinical investigation 2013;123(8):3539-51. doi: 10.1172/jci44403.
8. Heni M, Kullmann S, Veit R, Ketterer C, Frank S, Machicao F, Staiger H, Haring HU, Preissl H, Fritsche A. Variation in the obesity risk gene FTO determines the postprandial cerebral processing of food stimuli in the prefrontal cortex. Mol Metab 2014;3(2):109-13. doi: 10.1016/j.molmet.2013.11.009.
9. Kuhn AB, Feis DL, Schilbach L, Kracht L, Hess ME, Mauer J, Bruning JC, Tittgemeyer M. FTO gene variant modulates the neural correlates of visual food perception. Neuroimage 2016;128:21-31. doi: 10.1016/j.neuroimage.2015.12.049.
10. Rapuano KM, Zieselman AL, Kelley WM, Sargent JD, Heatherton TF, Gilbert-Diamond D. Genetic risk for obesity predicts nucleus accumbens size and responsivity to real-world food cues. Proceedings of the National Academy of Sciences of the United States of America 2017;114(1):160-5. doi: 10.1073/pnas.1605548113.
11. Wiemerslage L, Nilsson EK, Solstrand Dahlberg L, Ence-Eriksson F, Castillo S, Larsen AL, Bylund SB, Hogenkamp PS, Olivo G, Bandstein M, et al. An obesity-associated risk allele within the FTO gene affects human brain activity for areas important for emotion, impulse control and reward in response to food images. The European journal of neuroscience 2016;43(9):1173-80. doi: 10.1111/ejn.13177.
12. van der Laan LN, Smeets PAM. You are what you eat: a neuroscience perspective on consumers' personality characteristics as determinants of eating behavior. Current Opinion in Food Science 2015;3(0):11-8. doi: http://dx.doi.org/10.1016/j.cofs.2014.11.001.
13. Carp J. The secret lives of experiments: methods reporting in the fMRI literature. Neuroimage 2012;63(1):289-300. doi: 10.1016/j.neuroimage.2012.07.004.
14. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, Munafo MR. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 2013;14(5):365-76. doi: 10.1038/nrn3475.
15. Mumford JA. A power calculation guide for fMRI studies. Social cognitive and affective neuroscience 2012;7(6):738-42. doi: 10.1093/scan/nss059.

16. Durnez J, Degryse J, Moerkerke B, Seurinck R, Sochat V, Poldrack R, Nichols T. Power and sample size calculations for fMRI studies based on the prevalence of active peaks. bioRxiv 2016. doi: 10.1101/049429.
17. Eklund A, Nichols TE, Knutsson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proceedings of the National Academy of Sciences of the United States of America 2016;113(28):7900-5. doi: 10.1073/pnas.1602413113.
18. O'Reilly JX, Woolrich MW, Behrens TEJ, Smith SM, Johansen-Berg H. Tools of the trade: psychophysiological interactions and functional connectivity. Social cognitive and affective neuroscience 2012;7(5):604-9. doi: 10.1093/scan/nss055.
19. Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. A default mode of brain function. ProcNatlAcadSciUSA 2001;98(2):676-82.
20. Buckner RL, Andrews-Hanna JR, Schacter DL. The brain's default network: anatomy, function, and relevance to disease. Annals of the New York Academy of Sciences 2008;1124:1-38. doi: 10.1196/annals.1440.011.
21. Spetter MS, de Graaf C, Viergever MA, Smeets PA. Anterior cingulate taste activation predicts ad libitum intake of sweet and savory drinks in healthy, normal-weight men. J Nutr 2012;142(4):795-802.
22. Mehta S, Melhorn SJ, Smeraglio A, Tyagi V, Grabowski T, Schwartz MW, Schur EA. Regional brain response to visual food cues is a marker of satiety that predicts food choice. AmJClinNutr 2012.
23. Murdaugh DL, Cox JE, Cook EW, III, Weller RE. fMRI reactivity to high-calorie food pictures predicts short- and long-term outcome in a weight-loss program. Neuroimage 2012;59(3):2709-21.
24. Demos KE, Heatherton TF, Kelley WM. Individual differences in nucleus accumbens activity to food and sexual images predict weight gain and sexual behavior. J Neurosci 2012;32(16):5549-52. doi: 10.1523/jneurosci.5958-11.2012.
25. Yokum S, Ng J, Stice E. Attentional bias to food images associated with elevated weight and future weight gain: an FMRI study. Obesity (SilverSpring) 2011;19(9):1775-83.
26. Yokum S, Gearhardt AN, Harris JL, Brownell KD, Stice E. Individual differences in striatum activity to food commercials predict weight gain in adolescents. Obesity (Silver Spring, Md) 2014. doi: 10.1002/oby.20882.
27. Stice E, Burger KS, Yokum S. Reward Region Responsivity Predicts Future Weight Gain and Moderating Effects of the TaqIA Allele. J Neurosci 2015;35(28):10316-24. doi: 10.1523/jneurosci.3607-14.2015.
28. Volkow ND, Wang GJ, Telang F, Fowler JS, Thanos PK, Logan J, Alexoff D, Ding YS, Wong C, Ma Y, et al. Low dopamine striatal D2 receptors are associated with prefrontal metabolism in obese subjects: Possible contributing factors. Neuroimage 2008.
29. Wang GJ, Volkow ND, Logan J, Pappas NR, Wong CT, Zhu W, Netusil N, Fowler JS. Brain dopamine and obesity. Lancet 2001;357(9253):354-7.
30. Dang LC, Samanez-Larkin GR, Castrellon JJ, Perkins SF, Cowan RL, Zald DH. Associations between dopamine D2 receptor availability and BMI depend on age. Neuroimage 2016;138:176-83. doi: 10.1016/j.neuroimage.2016.05.044.
31. Benton D, Young HA. A meta-analysis of the relationship between brain dopamine receptors and obesity: a matter of changes in behavior rather than food addiction? Int J Obes (Lond) 2016;40 Suppl 1:S12-21. doi: 10.1038/ijo.2016.9.
32. Schwartz MW, Seeley RJ, Zeltser LM, Drewnowski A, Ravussin E, Redman LM, Leibel RL. Obesity Pathogenesis: An Endocrine Society Scientific Statement. Endocrine reviews 2017;38(4):267-96. doi: 10.1210/er.2017-00111.
33. Braff DL. The importance of endophenotypes in schizophrenia research. Schizophrenia research 2015;163(1-3):1-8. doi: 10.1016/j.schres.2015.02.007.

34. Willette AA, Kapogiannis D. Does the brain shrink as the waist expands? Ageing research reviews 2015;20:86-97. doi: 10.1016/j.arr.2014.03.007.
35. Poldrack RA, Fletcher PC, Henson RN, Worsley KJ, Brett M, Nichols TE. Guidelines for reporting an fMRI study. Neuroimage 2008;40(2):409-14.
36. Nichols TE, Das S, Eickhoff SB, Evans AC, Glatard T, Hanke M, Kriegeskorte N, Milham MP, Poldrack RA, Poline JB, et al. Best practices in data analysis and sharing in neuroimaging using MRI. Nat Neurosci 2017;20(3):299-303. doi: 10.1038/nn.4500.
37. Ioannidis JP. Why most published research findings are false. PLoS medicine 2005;2(8):e124. doi: 10.1371/journal.pmed.0020124.
38. Cremers HR, Wager TD, Yarkoni T. The relation between statistical power and inference in fMRI. PLoS One 2017;12(11):e0184923. doi: 10.1371/journal.pone.0184923.
39. Yarkoni T. Big Correlations in Little Studies: Inflated fMRI Correlations Reflect Low Statistical Power-Commentary on Vul et al. (2009). Perspectives on psychological science : a journal of the Association for Psychological Science 2009;4(3):294-8. doi: 10.1111/j.1745-6924.2009.01127.x.
40. Crutzen R, Peters GY. Targeting Next Generations to Change the Common Practice of Underpowered Research. Front Psychol 2017;8:1184. doi: 10.3389/fpsyg.2017.01184.
41. Joyce KE, Hayasaka S. Development of PowerMap: a software package for statistical power calculation in neuroimaging studies. Neuroinformatics 2012;10(4):351-65. doi: 10.1007/s12021-012-9152-3.
42. Poldrack RA, Farah MJ. Progress and challenges in probing the human brain. Nature 2015;526(7573):371-9. doi: 10.1038/nature15692.
43. Tong F, Pratte MS. Decoding patterns of human brain activity. Annual review of psychology 2012;63:483-509. doi: 10.1146/annurev-psych-120710-100412.
44. Bassett DS, Sporns O. Network neuroscience. Nat Neurosci 2017;20(3):353-64. doi: 10.1038/nn.4502.
45. Calhoun VD, Miller R, Pearlson G, Adali T. The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery. Neuron 2014;84(2):262-74. doi: 10.1016/j.neuron.2014.10.015.
46. Cornier MA, Despres JP, Davis N, Grossniklaus DA, Klein S, Lamarche B, Lopez-Jimenez F, Rao G, St-Onge MP, Towfighi A, et al. Assessing adiposity: a scientific statement from the American Heart Association. Circulation 2011;124(18):1996-2019. doi: 10.1161/CIR.0b013e318233bc6a.
47. Rosenbaum M, Agurs-Collins T, Bray MS, Hall KD, Hopkins M, Laughlin M, MacLean PS, Maruvada P, Savage CR, Small DM, et al. Accumulating Data to Optimally Predict Obesity Treatment (ADOPT): Recommendations from the Biological Domain. Obesity (Silver Spring, Md) 2018;26 Suppl 2:S25-s34. doi: 10.1002/oby.22156.
48. LaBar KS, Gitelman DR, Parrish TB, Kim YH, Nobre AC, Mesulam MM. Hunger selectively modulates corticolimbic activation to food stimuli in humans. BehavNeurosci 2001;115(2):493-500.
49. Goldstone AP, de Hernandez CG, Beaver JD, Muhammed K, Croese C, Bell G, Durighel G, Hughes E, Waldman AD, Frost G, et al. Fasting biases brain reward systems towards high-calorie foods. EurJ Neurosci 2009;30(8):1625-35.
50. Frank S, Laharnar N, Kullmann S, Veit R, Canova C, Hegner YL, Fritsche A, Preissl H. Processing of food pictures: influence of hunger, gender and calorie content. Brain Res 2010;1350:159-66.
51. Charbonnier L, van Meer F, Johnstone AM, Crabtree D, Buosi W, Manios Y, Androutsos O, Giannopoulou A, Viergever MA, Smeets PAM. Effects of hunger state on the brain responses to food cues across the life span. Neuroimage 2018;171:246-55. doi: 10.1016/j.neuroimage.2018.01.012.

52. Hernandez EA, Kahl S, Seelig A, Begovatz P, Irmler M, Kupriyanova Y, Nowotny B, Nowotny P, Herder C, Barosa C, et al. Acute dietary fat intake initiates alterations in energy metabolism and insulin resistance. The Journal of clinical investigation 2017;127(2):695-708. doi: 10.1172/jci89444.
53. Lichtman SW, Pisarska K, Berman ER, Pestone M, Dowling H, Offenbacher E, Weisel H, Heshka S, Matthews DE, Heymsfield SB. Discrepancy between self-reported and actual caloric intake and exercise in obese subjects. The New England journal of medicine 1992;327(27):1893-8. doi: 10.1056/nejm199212313272701.
54. Abdallah L, Chabert M, Louis-Sylvestre J. Cephalic phase responses to sweet taste. AmJClinNutr 1997;65(3):737-43.
55. Huettel S, Song A, McCarthy G. Functional Magnetic Resonance Imaging: {Sinauer Associates}, 2004.
56. Bennett C, Miller M, Wolford G. Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: an argument for multiple comparisons correction. NeuroImage 2009;47:S125.
57. Woo CW, Krishnan A, Wager TD. Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. Neuroimage 2014;91:412-9. doi: 10.1016/j.neuroimage.2013.12.058.
58. Hayasaka S, Nichols TE. Validating cluster size inference: random field and permutation methods. Neuroimage 2003;20(4):2343-56.
59. Hodgson K, Poldrack RA, Curran JE, Knowles EE, Mathias S, Göring HH, Yao N, Olvera RL, Fox PT, Almasy L. Shared genetic factors influence head motion during mri and body mass index. Cerebral Cortex 2016:1-8.
60. Nakamura Y, Goto TK, Tokumori K, Yoshiura T, Kobayashi K, Nakamura Y, Honda H, Ninomiya Y, Yoshiura K. Localization of brain activation by umami taste in humans. Brain Research 2011;1406(0):18-29.
61. Small DM, Gregory MD, Mak YE, Gitelman D, Mesulam MM, Parrish T. Dissociation of Neural Representation of Intensity and Affective Valuation in Human Gustation. Neuron 2003;39(4):701-11.
62. Friston KJ, Williams S, Howard R, Frackowiak RS, Turner R. Movement-related effects in fMRI time-series. Magn ResonMed 1996;35(3):346-55.
63. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 2012;59(3):2142-54.
64. Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage 2014;84:320-41. doi: 10.1016/j.neuroimage.2013.08.048.
65. Siegel JS, Power JD, Dubis JW, Vogel AC, Church JA, Schlaggar BL, Petersen SE. Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. Hum Brain Mapp 2014;35(5):1981-96. doi: 10.1002/hbm.22307.
66. Martin RE, Goodyear BG, Gati JS, Menon RS. Cerebral cortical representation of automatic and volitional swallowing in humans. Journal of Neurophysiology 2001;85(2):938-50.
67. Veldhuizen MG, Small DM. Modality-Specific Neural Effects of Selective Attention to Taste and Odor. Chemical Senses 2011;36(8):747-60. doi: 10.1093/chemse/bjr043.
68. Dipasquale O, Sethi A, Lagana MM, Baglio F, Baselli G, Kundu P, Harrison NA, Cercignani M. Comparing resting state fMRI de-noising approaches using multi- and single-echo acquisitions. PLoS One 2017;12(3):e0173289. doi: 10.1371/journal.pone.0173289.
69. Brooks JCW, Faull OK, Pattinson KTS, Jenkinson M. Physiological Noise in Brainstem fMRI. Frontiers in Human Neuroscience 2013;7:623. doi: 10.3389/fnhum.2013.00623.
70. Pruim RHR, Mennes M, Buitelaar JK, Beckmann CF. Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI. Neuroimage 2015;112:278-87. doi: 10.1016/j.neuroimage.2015.02.063.

71. Singer JDW, J.B. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. New York: Oxford University Press, 2003.
72. Geha PY, Aschenbrenner K, Felsted J, O'Malley SS, Small DM. Altered hypothalamic response to food in smokers. AmJClinNutr 2013;97(1):15-22.
73. Measelle JR, Stice E, Hogansen JM. Developmental trajectories of co-occurring depressive, eating, antisocial, and substance abuse problems in female adolescents. Journal of abnormal psychology 2006;115(3):524-38. doi: 10.1037/0021-843x.115.3.524.
74. Stice E, Yokum S. Gain in Body Fat Is Associated with Increased Striatal Response to Palatable Food Cues, whereas Body Fat Stability Is Associated with Decreased Striatal Response. J Neurosci 2016;36(26):6949-56. doi: 10.1523/jneurosci.4365-15.2016.
75. Stice E, Yokum S, Veling H, Kemps E, Lawrence NS. Pilot test of a novel food response and attention training treatment for obesity: Brain imaging data suggest actions shape valuation. Behaviour research and therapy 2017;94:60-70. doi: 10.1016/j.brat.2017.04.007.
76. Cornier MA, Salzberg AK, Endly DC, Bessesen DH, Tregellas JR. Sex-based differences in the behavioral and neuronal responses to food. Physiol Behav 2010;99(4):538-43.
77. Kishinevsky FI, Cox JE, Murdaugh DL, Stoeckel LE, Cook EW, III, Weller RE. fMRI reactivity on a delay discounting task predicts weight gain in obese women. Appetite 2012;58(2):582-92.
78. Stice E, Yokum S, Bohon C, Marti N, Smolen A. Reward circuitry responsivity to food predicts future increases in body mass: Moderating effects of DRD2 and DRD4. Neuroimage 2010;50(4):1618-25.
79. Winter SR, Yokum S, Stice E, Osipowicz K, Lowe MR. Elevated reward response to receipt of palatable food predicts future weight variability in healthy-weight adolescents. Am J Clin Nutr 2017;105(4):781-9. doi: 10.3945/ajcn.116.141143.
80. Ness A, Bruce J, Bruce A, Aupperle R, Lepping R, Martin L, Hancock L, Patrician T, Malley S, Selim N, et al. Pre-surgical cortical activation to food pictures is associated with weight loss following bariatric surgery. Surgery for obesity and related diseases : official journal of the American Society for Bariatric Surgery 2014;10(6):1188-95. doi: 10.1016/j.soard.2014.06.005.
81. Copas JB. Using regression models for prediction: shrinkage and regression to the mean. Statistical methods in medical research 1997;6(2):167-83. doi: 10.1177/096228029700600206.
82. Varoquaux G. Cross-validation failure: Small sample sizes lead to large error bars. Neuroimage 2017. doi: 10.1016/j.neuroimage.2017.06.061.
83. Skocik M, Collins J, Callahan-Flintoft C, Bowman H, Wyble B. I TRIED A BUNCH OF THINGS: THE DANGERS OF UNEXPECTED OVERFITTING IN CLASSIFICATION. bioRxiv 2016. doi: 10.1101/078816.
84. Poldrack RA, Gorgolewski KJ. Making big data open: data sharing in neuroimaging. Nat Neurosci 2014;17(11):1510-7. doi: 10.1038/nn.3818.
85. Fox PT, Lancaster JL, Laird AR, Eickhoff SB. Meta-analysis in human neuroimaging: computational modeling of large-scale databases. Annual review of neuroscience 2014;37:409-34. doi: 10.1146/annurev-neuro-062012-170320.
86. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. Large-scale automated synthesis of human functional neuroimaging data. Nature methods 2011;8(8):665-70. doi: 10.1038/nmeth.1635.
87. Salimi-Khorshidi G, Smith SM, Keltner JR, Wager TD, Nichols TE. Meta-analysis of neuroimaging data: a comparison of image-based and coordinate-based pooling of studies. Neuroimage 2009;45(3):810-23. doi: 10.1016/j.neuroimage.2008.12.039.
88. Gorgolewski KJ, Varoquaux G, Rivera G, Schwarz Y, Ghosh SS, Maumet C, Sochat VV, Nichols TE, Poldrack RA, Poline J-B, et al. NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. Frontiers in Neuroinformatics 2015;9(8). doi: 10.3389/fninf.2015.00008.

89. Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K. The WU-Minn Human Connectome Project: an overview. Neuroimage 2013;80:62-79. doi: 10.1016/j.neuroimage.2013.05.041.
90. Thompson PM, Stein JL, Medland SE, Hibar DP, Vasquez AA, Renteria ME, Toro R, Jahanshad N, Schumann G, Franke B, et al. The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. Brain imaging and behavior 2014;8(2):153-82. doi: 10.1007/s11682-013-9269-5.
91. Para AE, Sam K, Poublanc J, Fisher JA, Crawley AP, Mikulis DJ. Invalidation of fMRI experiments secondary to neurovascular uncoupling in patients with cerebrovascular disease. Journal of magnetic resonance imaging : JMRI 2017;46(5):1448-55. doi: 10.1002/jmri.25639.
92. Hege MA, Veit R, Krumsiek J, Kullmann S, Heni M, Rogers PJ, Brunstrom JM, Fritsche A, Preissl H. Eating less or more - Mindset induced changes in neural correlates of pre-meal planning. Appetite 2018;125:492-501. doi: 10.1016/j.appet.2018.03.006.
93. van Rijn I, de Graaf C, Smeets PA. It's in the eye of the beholder: selective attention to drink properties during tasting influences brain activation in gustatory and reward regions. Brain imaging and behavior 2017. doi: 10.1007/s11682-017-9710-2.
94. Veldhuizen MG, Small DM. Modality-specific neural effects of selective attention to taste and odor. ChemSenses 2011;36(8):747-60.
95. Grabenhorst F, Rolls ET. Selective attention to affective value alters how the brain processes taste stimuli. EurJNeurosci 2008;27(3):723-9.
96. Poldrack RA. Can cognitive processes be inferred from neuroimaging data? Trends Cogn Sci 2006;10(2):59-63.
97. Zhu JN, Wang JJ. The cerebellum in feeding control: possible function and mechanism. Cell MolNeurobiol 2008;28(4):469-78.
98. Frank S, Kullmann S, Veit R. Food related processes in the insular cortex. Front Hum Neurosci 2013;7:499. doi: 10.3389/fnhum.2013.00499.
99. Reid AT, Bzdok D, Genon S, Langner R, Muller VI, Eickhoff CR, Hoffstaedter F, Cieslik EC, Fox PT, Laird AR, et al. ANIMA: A data-sharing initiative for neuroimaging meta-analyses. Neuroimage 2016;124(Pt B):1245-53. doi: 10.1016/j.neuroimage.2015.07.060.
100. Smeets PA, Charbonnier L, van Meer F, van der Laan LN, Spetter MS. Food-induced brain responses and eating behaviour. ProcNutrSoc 2012;71(4):511-20.
101. Gorgolewski KJ, Alfaro-Almagro F, Auer T, Bellec P, Capota M, Chakravarty MM, Churchill NW, Cohen AL, Craddock RC, Devenyi GA, et al. BIDS apps: Improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. PLoS computational biology 2017;13(3):e1005209. doi: 10.1371/journal.pcbi.1005209.
102. Papies EK. Situating interventions to bridge the intention–behaviour gap: A framework for recruiting nonconscious processes for behaviour change. Social and Personality Psychology Compass 2017;11(7):e12323-n/a. doi: 10.1111/spc3.12323.
103. Stroebele N, De Castro JM. Effect of ambience on food intake and food choice. Nutrition 2004;20(9):821-38. doi: 10.1016/j.nut.2004.05.012.
104. van der Laan LN, Papies EK, Hooge IT, Smeets PA. Goal-directed visual attention drives health goal priming: An eye-tracking experiment. Health psychology : official journal of the Division of Health Psychology, American Psychological Association 2017;36(1):82-90. doi: 10.1037/hea0000410.
105. Papies EK, Hamstra P. Goal priming and eating behavior: enhancing self-regulation by environmental cues. Health psychology : official journal of the Division of Health Psychology, American Psychological Association 2010;29(4):384-8. doi: 10.1037/a0019877.
106. Papies EK. Health goal priming as a situated intervention tool: how to benefit from nonconscious motivational routes to health behaviour. Health psychology review 2016;10(4):408-24. doi: 10.1080/17437199.2016.1183506.

107. Harris JL, Bargh JA, Brownell KD. Priming effects of television food advertising on eating behavior. Health Psychol 2009;28(4):404-13.
108. Ouwehand C, Papies EK. Eat it or beat it. The differential effects of food temptations on overweight and normal-weight restrained eaters. Appetite 2010;55(1):56-60.
109. van der Laan LN, de Ridder DT, Viergever MA, Smeets PA. Activation in inhibitory brain regions during food choice correlates with temptation strength and self-regulatory success in weight-concerned women. Front Neurosci 2014;8:308. doi: 10.3389/fnins.2014.00308.
110. van der Laan LN, de Ridder DT, Charbonnier L, Viergever MA, Smeets PA. Sweet lies: neural, visual, and behavioral measures reveal a lack of self-control conflict during food choice in weight-concerned women. Front Behav Neurosci 2014;8:184. doi: 10.3389/fnbeh.2014.00184.
111. Charbonnier L, van der Laan LN, Viergever MA, Smeets PAM. Functional MRI of Challenging Food Choices: Forced Choice between Equally Liked High- and Low-Calorie Foods in the Absence of Hunger. PLoS ONE 2015;10(7):e0131727. doi: 10.1371/journal.pone.0131727.
112. van der Laan LN, de Ridder DT, Viergever MA, Smeets PA. Appearance matters: Neural correlates of food choice and packaging aesthetics. PLoS One 2012;7(7):e41738.
113. Hare TA, Camerer CF, Rangel A. Self-control in decision-making involves modulation of the vmPFC valuation system. Science 2009;324(5927):646-8.
114. Medic N, Ziauddeen H, Forwood SE, Davies KM, Ahern AL, Jebb SA, Marteau TM, Fletcher PC. The Presence of Real Food Usurps Hypothetical Health Value Judgment in Overweight People. eNeuro 2016;3(2). doi: 10.1523/eneuro.0025-16.2016.
115. Foerde K, Steinglass JE, Shohamy D, Walsh BT. Neural mechanisms supporting maladaptive food choices in anorexia nervosa. Nat Neurosci 2015;18(11):1571-3. doi: 10.1038/nn.4136.
116. Schwebel DC, McClure LA, Severson J. Usability and Feasibility of an Internet-Based Virtual Pedestrian Environment to Teach Children to Cross Streets Safely. Virtual reality 2014;18(1):5-11. doi: 10.1007/s10055-013-0238-5.
117. Gasco J, Holbrook TJ, Patel A, Smith A, Paulson D, Muns A, Desai S, Moisi M, Kuo YF, Macdonald B, et al. Neurosurgery simulation in residency training: feasibility, cost, and educational benefit. Neurosurgery 2013;73 Suppl 1:39-45. doi: 10.1227/neu.0000000000000102.
118. Nederkoorn C, Guerrieri R, Havermans RC, Roefs A, Jansen A. The interactive effect of hunger and impulsivity on food intake and purchase in a virtual supermarket. International journal of obesity (2005) 2009;33(8):905-12. doi: 10.1038/ijo.2009.98.
119. Riva G, Gaggioli A, Villani D, Preziosa A, Morganti F, Corsi R, Faletti G, Vezzadini L. NeuroVR: an open source virtual reality platform for clinical psychology and behavioral neurosciences. Studies in health technology and informatics 2007;125:394-9.
120. Waterlander WE, Scarpa M, Lentz D, Steenhuis IH. The virtual supermarket: an innovative research tool to study consumer food purchasing behaviour. BMC public health 2011;11:589. doi: 10.1186/1471-2458-11-589.
121. Lenggenhager B, Tadi T, Metzinger T, Blanke O. Video ergo sum: manipulating bodily self-consciousness. Science 2007;317(5841):1096-9. doi: 10.1126/science.1143439.
122. Burke RR, Harlam BA, Kahn BE, Lodish LM. Comparing Dynamic Consumer Choice in Real and Computer-simulated Environments. Journal of Consumer Research 1992;19(1):71-82. doi: 10.1086/209287.
123. van Herpen E, van den Broek E, van Trijp HCM, Yu T. Can a virtual supermarket bring realism into the lab? Comparing shopping behavior using virtual and pictorial store representations to behavior in a physical store. Appetite 2016;107:196-207. doi: 10.1016/j.appet.2016.07.033.
124. Waterlander WE, Jiang Y, Steenhuis IH, Ni Mhurchu C. Using a 3D virtual supermarket to measure food purchase behavior: a validation study. Journal of medical Internet research 2015;17(4):e107. doi: 10.2196/jmir.3774.

125. Campo KG, E; Guerra, F. Computer Simulated Shopping Experiments for Analyzing Dynamic Purchasing Patterns: Validation and Guidelines. J Empir Gen Mark Sci 1999;4(2).
126. Kundu P, Voon V, Balchandani P, Lombardo MV, Poser BA, Bandettini PA. Multi-echo fMRI: A review of applications in fMRI denoising and analysis of BOLD signals. Neuroimage 2017;154:59-80. doi: 10.1016/j.neuroimage.2017.03.033.
127. Yun SD, Shah NJ. Whole-brain high in-plane resolution fMRI using accelerated EPIK for enhanced characterisation of functional areas at 3T. PLoS One 2017;12(9):e0184759. doi: 10.1371/journal.pone.0184759.

Table 1. Overview of requirements and recommendations for nutritional neuroimaging[1]

Requirement/recommendation [section]	**Level[2]**
Participant description	
Report age	M
Report gender and test for possible effects	M
Report race and ethnicity	R
Report handedness and account for non-righthandedness in analyses	M
Report socio-economic status	R
Report physical activity level	R
Report use of relevant medication, tobacco, alcohol and caffeine	R
Report menstrual cycle phase and how this was accounted for in the analysis	HR
Report BMI or age-adjusted BMI and test for possible effects	M
Report further adiposity measures, e.g. % body fat, waist-hip ratio	R
Report a measure of dietary restraint	R
Report a measure of stress	R
Report personality traits such as reward sensitivity and impulsivity [2.2]	R
Eating disorder scales [section 2.2]	R
Report weight history; weight lost or gained in the weeks before brain imaging	HR
Report time since last meal	M

Requirement/recommendation [section]	**Level[2]**
Standardize the last meal before brain imaging	R
Report appetite ratings	HR
Report thirst ratings	R
Study design/procedures	
Describe the hunger state(s) and how they were achieved	M
Report food stimulus details including macronutient composition and energy content	M
For pre- versus post feeding studies motivate why fasted and fed conditions could not be completed on separate days to avoid order effects	M
fMRI task	
Mandatory items in the COBIDAS checklist (S1)[1]	M
Provide a power calculation [2.3]	HR
Report the task instructions	M
Report the number and timing of the task events and how their order was randomized and/or optimized	M
Describe the stimuli used and how they were matched e.g. on visual characteristics	M
Report stimulus liking and where appropriate intensity [2.4]	M
For taste stimuli: report temperature, volume, flow rate, swallowing instructions [2.4]	M
For olfactory stimuli: report temperature, flow rate and sniffing instructions	M

Requirement/recommendation [section]	Level[2]
fMRI data analysis	
Mandatory items in the COBIDAS checklist (S1)[1]	M
Indicate how correction for multiple comparisons was done and how the threshold used was determined	M
Test multiple ROIs with a single combined ROI mask	M
Use appropriate covariates, such as stimulus liking, gender, menstrual cycle phase, BMI	HR
Include blood parameters as covariates, if available [2.2]	R
Statistical inference/interpretation	
Avoid reverse inference [3.2]	HR
Be as specific as possible in the degree of overlap when comparing activated brain regions with regions found in other studies [3.3]	HR

[1] General requirements and recommendations for reporting neuroimaging methods can be found in the COBIDAS checklist (Online Supporting Material S1).

[2] Level: M = Mandatory; HR = highly recommended; R = recommended.

Supplemental Checklist S1

This is a PDF form version of Appendix D from http://www.humanbrainmapping.org/COBIDASreport and the below preprint, published under a CC-BY 4.0 international license. https://creativecommons.org/licenses/by/4.0/

Reference:
Nichols, T. E., Das, S., Eickhoff, S. B., Evans, A. C., Glatard, T., Hanke, M., Kriegeskorte, N., Milham, M. P., Poldrack, R. A., Poline, J.-B., Proal, E., Thirion, B., Van Essen, D. C., White, T., Yeo, B. T. T. (2016). Best Practices in Data Analysis and Sharing in Neuroimaging using MRI. bioRxiv doi: 10.1101/054262.

Appendix D. Itemized lists of best practices and reporting items

This section contains checklists for practices and items to report. Each item has been included because it is an essential piece of information needed to understand, evaluate and reproduce an experiment. Authors should strive to include all these items, but items marked as “Mandatory” are particularly crucial, and a published work cannot be considered complete without such information.

Authors are required to check all mandatory items that apply (Y or N/A).

Table D.1. Experimental Design Reporting

Aspect	Notes	Mandatory
Number of subjects	*Elaborate each by group if have more than one group.*	
Subjects approached		N
Subjects consented		N
Subjects refused to participate	Provide reasons.	N
Subjects excluded	Subjects excluded after consenting but before data acquisition; provide reasons.	N
Subjects participated and analyzed	Provide the number of subjects scanned, number excluded after acquisition, and the number included in the data analysis. If they differ, note the number of subjects in each particular analysis.	Y ☐
Inclusion criteria and descriptive	*Elaborate each by group if have more than one group.*	
Age	Mean, standard deviation and range.	Y ☐
Sex	Absolute counts or relative frequencies.	Y ☐
Race & ethnicity	Per guidelines of NIH or other relevant agency.	N
Education, SES	Education is essential for studies comparing patient and control groups; complete SES reporting less important for single-group studies, but still useful. Specify measurement instrument used; may be parental SES and education if study has minors.	Y ◯ N/A ◯
IQ	Specify measurement instrument used.	N

Handedness	Absolute or relative frequencies; basis of handedness-attribution (self-report, EHI, other tests). Important for fMRI, may be less important for structural MRI.	Y ☐
Exclusion criteria	Describe any screening criteria, including those applied to "normal" sample such as MRI exclusion criteria.	Y ☐
Clinical criteria	Detail the area of recruitment (in- vs. outpatient setting, community hospital vs. tertiary referral center etc.) as well as whether patients were currently in treatment.	Y ◯ N/A ◯
Clinical instruments	Describe the instruments used to obtain the diagnosis and provide tests of intra- or inter-rater reliability. Clarify whether a "clinical diagnosis" or "inventory diagnosis" was used (if applicable). State the diagnostic system (ICD, DSM etc) that was used.	Y ◯ N/A ◯
Matching strategy	If applicable.	Y ◯ N/A ◯
Population & recruitment strategy	Population from which subjects were drawn, and how and where recruitment took place, e.g., schools, clinics, etc. If possible, note if subjects are research-naive or have participated in other studies before.	Y ☐
Subject scanning order	With multiple groups, information on ordering and or balance over time; especially report relative to scanner changes/upgrades. (Ideally, use randomized or interleaved order to avoid bias due to scanner changes/upgrades.)	Y ◯ N/A ◯
Neurocognitive measures	All measures collected on subjects should be described and reported.	Y ☐
Ethical considerations		
Ethical approval	Describe approval given, including the particular institutional review board, medical ethics committee or equivalent that granted the approval. When data is shared, describe the ethics/institutional approvals required from either the author (source) or recipient.	Y ☐
Informed consent	Record whether subjects provided informed consent or, if applicable, informed assent.	Y ☐

Design specifications		
Design type	Task or resting state. Event-related or block design. (See body text for usage of 'block design' terminology.)	Y ◯ N/A ◯
Condition & stimuli	Clearly describe each condition and the stimuli used. Be sure to completely describe baseline (e.g. blank white/black screen, presence of fixation cross, or any other text), especially for resting-state studies. When possible provide images or screen snapshots of the stimuli.	Y ◯ N/A ◯
Number of blocks, trials or experimental units	Specify per session, and if differing by subject, summary statistics (mean, range and/or standard deviation) of such counts.	Y ◯ N/A ◯
Timing and duration	Length of each trial or block (both, if trials are blocked), and interval between trials. Provide the timing structure of the events in the task, whether a random/jittered pattern or a regular arrangement; any jittering of block onsets.	Y ◯ N/A ◯
Length of the experiment	Describe the total length of the scanning session, as well as the duration of each run. (Important to assess subject fatigue.)	Y ☐
Design optimization	Whether design was optimized for efficiency, and how.	Y ◯ N/A ◯
Presentation software	Name software, version and operating system on which the stimulus presentation was run. When possible, provide code used to drive experiment.	Y ◯ N/A ◯
Task specification		
Condition	Enumerate the conditions and fully describe and reference each. Consider using a shorthand name, e.g. AUDSTIM, VISSTIM, to refer to each condition, to clarify the distinction between a specific modeled effect and a psychological construct. Naming should reflect the distinction between instruction periods and actual stimuli, and between single parameters and contrasts of parameters.	Y ◯ N/A ◯
Instructions	Specify the instructions given to subjects for each condition (ideally the exact text in supplement or appendix). For resting-state, be sure to indicate eyes-closed, eyes-open, any fixation. Describe if the subjects received any	Y ☐

	the task, and state if there was a familiarization / training inside or outside the scanner.	
Stimuli	Specifics of stimuli used in each run. For example, the unique number of stimuli used, and whether/how stimuli were repeated over trials or conditions.	Y ◯ N/A ◯
Randomization	Describe block or event ordering as deterministic, or report manner of randomization, in terms of order and timing. If pseudo-randomized, i.e. under constraints, describe how and the criteria used to constrain the orders/timings.	Y ◯ N/A ◯
Stimulus presentation & response collection.	Specify the presentation hardware (e.g. back projection, in-room display, goggles, etc), and the response systems (e.g. button boxes, eye tracking, physiology). Note how equipment was synched to the scanner (e.g. scanner TTL, or manual sync.)	Y ◯ N/A ◯
Run order	Order in which tasks runs are conducted in the scanner.	Y ◯ N/A ◯
Power analysis		
Outcome	Specify the type of outcome used as the basis of power computations, e.g. signal in a pre-specified ROI, or whole image voxelwise (or cluster-wise, peak-wise, etc.).	Y ☐
Power parameters	Specify ● Effect size (or effect magnitude and standard deviation separately). ● Source of predicted effect size (previous literature with citation; pilot data with description, etc). ● Significance level (e.g. uncorrected alpha 0.05 for an ROI, or FWE-corrected significance ● Target power (typically 80%). ● Any other parameters set (e.g., for spatial methods a brain volume and smoothness may be needed to be specified).	Y ☐
Behavioral performance		

Variables recorded	State number of type of variables recorded (e.g. correct button press, response time).	Y ◯ N/A ◯
Summary statistics	Summaries of behavior sufficient to establish that subjects were performing the task as expected. For example, correct response rates and/or response times, summarized over subjects (e.g. mean, range and/or standard deviation).	Y ◯ N/A ◯

Table D.2. Acquisition Reporting

Aspect	**Notes**	**Mandatory**
Subject preparation		
Mock scanning	Use of an MRI simulator to acclimate subjects to scanner environment. Report type of mock scanner and protocol (i.e. duration, types of simulated scans, experiments).	N
Special accommodations	For example, for pediatric scanning, presence of parent/guardian in the room.	Y ◯ N/A ◯
Experimenter personnel	Whether a single or multiple experimenters interacted with the subjects.	N
MRI system description		
Scanner	Provide make, model & field strength in tesla (T).	Y ☐
Coil	Receive coil (e.g. “a 12-channel phased array coil”, but more details for a custom coil) and (if nonstandard) transmit coil. Additional information on the gradient system, e.g. gradient strength (if non-standard for the make and model, or switchable).	Y ☐
Significant hardware modifications	For example, special gradient inserts/sets.	N

Software version	Highly recommended when sharing vendor-specific protocols or exam cards, as version may be needed to correctly interpret that information.	N
MRI acquisition		
Pulse sequence type	For example, gradient echo, spin echo, etc.	Y ☐
Imaging type	For example, echo planar imaging (EPI), spiral, 3D. Number of shots (if multi-shot); partial Fourier scheme & reconstruction method (if used);	Y ☐
Essential sequence & imaging parameters.	For all acquisitions: ● Echo time (TE). ● Repetition time (TR). o For multi-shot acquisitions, additionally the time per volume. ● Flip angle (FA). ● Acquisition time (duration of acquisition). Functional MRI: ● Number of volumes. ● Sparse sampling delay (delay in TR) if used. Inversion recovery sequences: ● Inversion time (TI). B0 field maps: ● Echo time difference (dTE). Diffusion MRI: ● Number of directions. o Direction optimization, if used and type. ● b-values. ● Number of b=0 images. ● Number of averages (if any). ● Single shell, multi-shell (specify equal or unequal spacing). ● Single- or dual-spin-echo, gradient mode (serial or parallel). ● If cardiac gating used. Imaging parameters:	Y ☐

	• Field of view.
	• In-plane matrix size, slice thickness and interslice gap, for 2D acquisitions.
	• Slice orientation:
	  ○ Axial, sagittal, coronal or oblique.
	  ○ Angulation: If acquistion not aligned with scanner axes, specify angulation to AC-PC line (see Slice position procedure).
	• 3D matrix size, for 3D acquisitions.	

Phase encoding	Specify phase encoding direction (e.g. as A/P, L/R, or S/I). For 3D, specify “partition encode” (aka slice) direction. Phase encoding reversal: Mention if used (aka “blip-up/blip-down”).	Y ☐
Parallel imaging method & parameters	Report: • Method, e.g. SENSE, GRAPPA or other parallel imaging method, and acceleration factor. • Matrix coil mode, and coil combining method (if non-standard).	Y ◯ N/A ◯
Multiband parameters	Multiband factor and field-of-view shift (only if applicable).	Y ◯ N/A ◉
Readout parameters	Receiver bandwidth, readout duration, echo spacing.	N
Fat suppression	For anatomical scans, whether it was used or not.	Y ☐
Shimming	Any specialized shimming procedures.	Y ◯ N/A ◯
Slice order & timing	For fMRI acquisitions, interleaved vs. sequential ordering and direction (ascending/descending), location of 1st slice; any specialized slice timing.	Y ☐
Slice position procedure	For example, landmark guided vs. auto-alignment.	N
Brain coverage	Report whether coverage was whole-brain, and whether cerebellum and brainstem were included. If not whole-brain, note the nature of the partial area of coverage. If axial and co-planar with AC-PC line, the volume coverage in terms of Z in mm.	Y ☐

Scanner-side preprocessing	Including: ● Reconstruction matrix size differing from acquisition matrix size. ● Prospective-motion correction (including details of any optical tracking, and how motion parameters are used). ● Signal inhomogeneity correction. ● Distortion-correction.	Y ◯ N/A ◯
Scan duration	In seconds	N
Other non-standard procedures	Including: ● Turning off the cold head(s) (e.g. during EEG/fMRI or spectroscopy measurements). ● Reduce sound pressure by limiting the gradient slew rate.	N
T1 stabilization	Number of initial “dummy” scans acquired and then discarded by the scanner.	Y ◯ N/A ◯
Diffusion MRI gradient table	Also referred to as the b-matrix (but not to be confused with the 3×3 matrix that describes diffusion weighting for a single diffusion weighted measurement).	N
Perfusion: Arterial Spin Labelling MRI	● ASL Labelling method (e.g. continuous ASL (CASL), pseudo-continuous ASL (PCASL), Pulsed ALS (PASL), velocity selective ASL (VSASL)). ● Use of background suppression pulses and their timing. ● For either PCASL or CASL report: ○ Label Duration. ○ Post-labeling delay (PLD). ○ Location of the labeling plane. ● For PCASL also report: ○ Average labeling gradient. ○ Slice-selective labeling gradient. ○ Flip angle of B1 pulses. ○ Assessment of inversion efficiency; QC used to ensure off-resonance artifacts not problematic, signal obtained over whole brain. ● For CASL also report:	Y ◯ N/A ◯

	○ Use of a separate labeling coil. ○ Control scan/pulse used. ○ B1 amplitude. ● For PASL report ○ TI. ○ Labeling slab thickness. ○ Use of QUIPSS pulses and their timing. ● For VSASL ○ TI. ○ Choice of velocity selection cutoff (“VENC”).	
Perfusion: Dynamic Susceptibility Contrast MRI	Specify: ● Number of baseline volumes. ● Type, name and manufacturer of intravenous bolus (e.g. gadobutrol, Gadavist, Bayer). ● Bolus amount and concentration (e.g. 0.1 ml/kg and 0.1 mmol/kg). ● Injection rate (e.g. 5 ml/s). ● Post-injection of saline (e.g. 20 ml). ● Injection method (e.g. power injector).	Y ◯ N/A ◯
Preliminary quality control		
Motion monitoring	For functional or diffusion acquisitions, any visual or quantitative checks for severe motion; likewise, for structural images, checks on motion or general image quality.	Y ◯ N/A ◯
Incidental findings	Protocol for review of any incidental findings, and how they are handled in particular with respect to possible exclusion of a subject’s data.	N

Table D.3. Preprocessing Reporting

Aspect	Notes	Mandatory
Software	For each software used, be sure to include version and revision number.	Y ☐
Software citation	Include URL and Research Resource Identifier for each software used.	N
T1 stabilization	Number of initial "dummy" scans discarded as part of preprocessing (if not already performed by scanner).	Y ◯ N/A ◯
Brain extraction	If performed, report: ● Name of software/method (e.g., BET, recon-all in FreeSurfer, etc). ● Parameter choices (e.g. BET's fractional intensity threshold). ● Any manual editing applied to the brain masks.	Y ◯ N/A ◯
Segmentation	For structural images, method used to extract gray, white, CSF and other tissue classes.	Y ☐
Slice time correction	If performed, report: ● Name of software/method. ● Whether performed after or before motion correction. ● Reference slice. ● Interpolation type and order (e.g., 3rd order spline or sinc).	Y ◯ N/A ◯
Motion correction	Report: ● Name of software/method. ● Use of non-rigid registration, and if so the type of transformation. ● Use of motion susceptibility correction (fieldmap-based unwarping), as well as the particular software/method. ● Reference scan (e.g. 1st scan or middle scan). ● Image similarity metric (e.g. normalized correlation, mutual information, etc).	Y ◯ N/A ◯

	• Interpolation type (e.g., spline, sinc), and whether image transformations are combined to allow a single interpolation. • Use of any slice-to-volume registration methods, or integrated with slice time correction.	
Gradient distortion correction	(If not already described as part of motion susceptibility correction.)	Y ◯ N/A ◯
Diffusion MRI eddy current correction	Report: • Name of software/method, and if integrated with motion correction • Image similarity / cost function. • Type of transformation (e.g. rigid body, affine) and whether constrained only along the phase encode direction. • Note if gradient table (b-matrix) is then re-oriented. • Volumetric change applied for eddy current along the phase-encode axis (by the Jacobian determinant).	Y ◯ N/A ◉

Diffusion estimation	For all methods, report
	• Model, parameterisation and number of free parameters.
	• Estimation method.
	• Outlier handling approach.
	• Some evidence of fit quality; e.g sample of slices of diffusion weighted data, or residual maps.
	Items to note for particular approaches:
	• Tensor or Kurtosis.
	  ○ Any parameter constraints, like cylindrical symmetry.
	• Multi-compartmental models.
	  ○ Compartments of the model.
	• Orientation distribution function.
	  ○ Parametric (model) or nonparametric (basis function) model.
	  ○ Whether orientation distribution function or fibre orientation density is reported.
	  ○ For spherical deconvolution, note how the canonical fibre response function is derived (e.g. from the data themselves, or simulated data).
		Y ◯ N/A ◉

Diffusion processing	Report: ● Summary measures computed (FA, MD, AD, RD, MK, AK, RK, etc.). ● Whether a track based or voxel-wise method is used. ● Threshold used to define analysis voxels. ● Use of population reference track atlas vs. custom atlas (specify set of subjects used to create atlas). ● Standard deviation map (across subjects).	N
Diffusion tractography	Report: ● Name of software/method. ● Step size, turning angle and stopping criteria. ● For ROI based analysis, definition of ROIs (e.g. specify the images used to draw ROIs; manual, semi-automatic or automatic definition of ROIs). ● For tracking, note step-size, turning angle, any anatomical constraints imposed, and stopping criteria. ● If a measure of path probability / "connectivity" is extracted, clearly define this measure.	Y ◯ N/A ◯
Perfusion: Arterial Spin Labeling	Report modelling/post-processing scheme: ● For subtraction, specify whether simple subtraction, running, sinc-subtraction, etc. ● For quantitative model, specify model used, number of free parameters.	Y ◯ N/A ◉
Perfusion: Dynamic Susceptibility Contrast MRI	● How concentration time curves are calculated, e.g. use of T1 corrections (if short TR) or corrections for leakage. ● Selection of arterial input function (e.g. manual or automatic with reference to method). ● Deconvolution method (kinetic model) to estimate residue function (e.g. SVD or parametric model). ● Details of parameter calculations (e.g. CBF, CBV, MTT, TTP, Tmax).	Y ◯ N/A ◉
Function-structure (intra-subject) coregistration	Report: ● Name of software/method. ● Type of transformation (rigid, nonlinear); if nonlinear, type of transformation	Y ◯ N/A ◯

	• Cost function (e.g., correlation ratio, mutual information, boundary-based registration, etc). • Interpolation method (e.g., spline, linear). Note this step might not be necessary if direct T2* to a functional template registration is used.	
Distortion correction	Use of any distortion correction due to field or gradient nonlinearity.	Y ◯ N/A ◯
Intersubject registration	Report: • Name of software/method (e.g., FSL flirt followed by fnirt, FreeSurfer, Caret, Workbench, etc) • Whether volume and/or surface based registration is used (if not already clearly implied). • Image types registered (e.g. T2* or T1). • Any preprocessing to images; e.g. for T1, bias field correction, or segmentation of gray matter; for T2*, single image (specify image) or mean image. • Template space (e.g., MNI, Talairach, fsaverage, FS_LR), modality (e.g., T1, T2*), resolution (e.g., 2mm, fsaverage5, 32k_FS_LR), and the specific name of template image used; note the domain of the template if not whole brain, i.e. cortical surface only, cerebellum only, CIFTI 'grayordinates' (cortical surface vertices + subcortical gray matter voxels), etc. • Additional template transformation for reporting; e.g., if using a template in MNI space, but reporting coordinates in Talairach, clearly note and report method used (e.g., Brett's mni2tal, Lancaster's icbm_spm2tal). • Choice of warp (rigid, nonlinear); if nonlinear, transformation type (e.g., B-splines, stationary velocity field, momentum, non-parametric displacement field); if a parametric transformation is used, report resolution, e.g., 10x10x10 spline control points. • Use of regularization, and the parameter(s) used to set degree of regularization.	Y ☐

	• Interpolation type (e.g., spline, linear); if projection from volume to surface space, how were voxels sampled from the volume (e.g., trilinear; nearest neighbor; ribbon-constrained specifying inner and outer surface used). • Cost function (e.g., correlation ratio, mutual information, SSD). • Use of cost-function masking.	
Intensity correction	Bias field corrections for structural MRI, but also correction of odd versus even slice intensity differences attributable to interleaved EPI acquisition without gaps.	Y ◯ N/A ◯
Intensity normalization	Scan-by-scan or run-wide scaling of image intensities before statistical modelling. E.g. SPM scales each run such that the mean image will have mean intracerebral intensity of 100; FSL scales each run such that the mean image will have an intracerebral mode of 10,000.	N

Artifact and structured noise removal	Use of physiological noise correction method. Report:
	• Name of software/method used (e.g. CompCor, ICA-FIX, ICA-AROMA, etc.).
	• If using a nuisance regression method, specify regressors used; for each type, include key details, as follows:
	  ○ Motion parameters.
	  ■ Expansion basis and order (e.g. 1st temporal derivatives; Volterra kernel expansion)
	  ○ Tissue signals.
	  ■ Tissue type (e.g., whole brain, gray matter, white matter, ventricles).
	  ■ Tissue definition (e.g., a priori seed, automatic segmentation, spatial regression).
	  ■ Signal definition (e.g., mean of voxels, first singular vector, etc.).
	  ○ Physiological signals
	  ■ e.g., heart rate variability, respiration.
		Y ◯ N/A ◯

	■ Modeling choices (e.g. RETROICOR, cardiac and/or respiratory response functions) and number of computed regressors.	
Volume censoring	Remediation of problem scans, also known as “scrubbing” or “de-spiking”. Report: ● Name of software/method. ● Criteria (e.g., frame-by-frame displacement threshold, percentage BOLD change). ● Use of censoring or interpolation; if interpolation, method used (e.g., spline, spectral estimation).	Y ◯ N/A ◯
Resting state fMRI feature	Creation of summary measure like ALFF, fALFF, ReHo. For ALFF, fALFF report: ● Lower and upper band pass frequencies. For ReHo, report: ● Neighborhood size used to compute local similarity measures (e.g. 6, 18 or 26). ● Similarity measure (e.g. Kendall’s coefficient of concordance).	Y ◯ N/A ◉
Spatial smoothing	If this preprocessing step is performed, report: ● Name of software/method. ● Size and type of smoothing kernel. ● Filtering approach, e.g., fixed kernel or iterative smoothing until fixed FWHM. ● Space in which smoothing is performed (i.e. native volume, native surface, MNI volume, template surface).	Y ◯ N/A ◯
Quality control reports	Summaries of subject motion (e.g. mean framewise displacement), image variance (e.g. DVARS), and note of any other irregularities found (e.g. motion or poor SNR not sufficiently severe to warrant exclusion). Should be included with any publically shared data.	N

Table D.4. Statistical Modeling & Inference

Aspect	Notes/Ontology	Mandatory
Mass univariate analyses		
Dependent variable: Data submitted to statistical modeling	Report the number of time points, number of subjects; specify exclusions of time points / subjects, if not already specified in experimental design.	Y ◯ N/A ◯
Dependent variable: Spatial region modeled	If not “Full brain”, give a specification of an anatomically or functionally defined mask.	Y ◯ N/A ◯
Independent variables	For first level fMRI, specify: ● Event-related design predictors. ○ Modeled duration, if other than zero. ○ Parametric modulation. ● Block Design predictors. ○ Note whether baseline was explicitly modeled. ● HRF basis, typically one of: ○ Canonical only. ○ Canonical plus temporal derivative. ○ Canonical plus temporal and dispersion derivative. ○ Smooth basis (e.g. SPM “informed” or Fourier basis; FSL’s FLOBS). ○ Finite Impulse Response model. ● Drift regressors (e.g. DCT basis in SPM, with specified cut-off). ● Movement regressors; specify if squares and/or temporal derivative used. ● Any other nuisance regressors, and whether they were entered as interactions (e.g. with a task effect in 1st level fMRI, or with group effect). ● Any orthogonalization of regressors, and set of other regressors used to orthogonalize against. For second level fMRI or general group model, specify: ● Group effects (patients vs. controls).	Y ◯ N/A ◯

	• Clearly state whether or not covariates are split by group (i.e. fit as a group-by-covariate interaction). • Other between subject effects (age, sex; for VBM, total GM or ICV). For group model with repeated measures, specify: • How condition effects are modeled (e.g. as factors, or as linear trends). • Whether subject effects are modeled (i.e. as regressors, as opposed to with a covariance structure).	
Model type	Some suggested terms include: • "Mass Univariate". • "Multivariate" (e.g. ICA on whole brain data). • "Mass Multivariate" (e.g. MANOVA on diffusion or morphometry tensor data). • "Local Multivariate" (e.g. "searchlight"). • "Multivariate, intra-subject predictive" (e.g. classify individual trials in event-related fMRI). • "Multivariate inter-subject predictive" (e.g. classify subjects as patient vs. control). • "Representational Similarity Analysis".	Y ☐

Model settings	The essential details of the model. For mass-univariate, first level fMRI, these include:
	• Drift model, if not already specified as a dependent variable (e.g. locally linear detrending of data & regressors, as in FSL).
	• Autocorrelation model (e.g. global approximate AR(1) in SPM; locally regularized autocorrelation function in FSL).
	
	For mass-univariate second level fMRI these include:
	• Fixed effects (all subjects' data in one model).
	• Random or mixed-effects model, implemented with:
	  ○ Ordinary least squares (OLS, aka unweighted summary statistics approach; SPM default, FSL FEAT's "Simple OLS").
	  ○ weighted least squares (i.e. FSL FEAT's "FLAME 1"), using voxel-wise estimate of between subject variance.
		Y ☐

	○ Global weighted least squares (i.e. SPM's MFX). With any group (multi-subject) model, indicate any specific variance structure, e.g. ● Un-equal variance between groups (and if globally pooled, as in SPM). ● If repeated measures, the specific covariance structure assumed (e.g. compound symmetric, or arbitrary; if globally pooled). For local-multivariate report: ● The number of voxels in the local model. ● Local model used (e.g. Canonical Correlation Analysis) with any constraints (e.g. positive weights only).	
Inference: Contrast/effect	● Specification of the precise effect tested, often as a linear contrast of parameters in a model. When possible, define these in terms of the task or stimulus conditions instead of psychological concepts (See *Task Specification* in *Experimental Design Reporting*). ● Provide tables/figures on main effects (e.g. in supplement), not just differences or interactions. For example, an inference on a difference of two fMRI conditions, A-B, doesn't indicate if both A & B induced positive changes; likewise, to fully interpret an interaction requires knowledge of the main effects. ● Indicate any use of any omnibus ANOVA tests. ● All contrasts explored as part of the research should be fully described in the methods section, whether or not they are considered in the results. ● If performing a two-sided test via two one-sided tests, double the one-sided p-values to convert them into two-sided p-values. For example, if looking at both a contrast [-1 1] and [1 -1] together, each with cluster-forming threshold p=0.001, double the FWE cluster p-values from each contrast to obtain two-sided inferences.	Y ☐
Inference: Search region	● Whole brain or "small volume"; carefully describe any small volume correction used for each contrast. ● If a small-volume correction mask is defined anatomically, provide named anatomical regions from a publicly available ROI atlas.	Y ☐

	● If small-volume correction mask is functionally defined, clearly describe the functional task and identify any risk of circularity. ● All small-volume corrections should be fully described in the methods section, not just mentioned in passing in the results.	
Inference: Statistic type	Typically one of: ● Voxel-wise (aka peak-wise in SPM). ● Cluster-wise. ○ Cluster size. ○ Cluster mass. ○ Threshold-free Cluster Enhancement (TFCE). For cluster size or mass, report: ● Cluster-forming threshold. For all cluster-wise methods, report: ● Neighborhood size used to form clusters (e.g. 6, 18 or 26). For TFCE, report: ● Use of non-default TFCE parameters.	Y ☐
Inference: P-value computation	Report if anything but standard parametric inference used to obtain (uncorrected) P-values. If nonparametric method was used, report method (e.g. permutation or bootstrap) and number of permutations/samples used.	Y ◯ N/A ◯
Inference: Multiple testing correction	For mass-univariate, specify the type of correction and how it is obtained, especially if not the typical usage. Usually one of: ● Familywise Error. ○ Random Field Theory (typical). ○ Permutation. ○ Monte Carlo. ○ Bonferroni. ● False Discovery Rate. ○ Benjamini & Hochberg FDR (typical). ○ Positive FDR. ○ Local FDR. ○ Cluster-level FDR.	Y ☐

	• None/Uncorrected. If permutation or Monte Carlo, report the number of permutations/samples. If Monte Carlo, note the brain mask and smoothness used, and how smoothness was estimated.	
Functional connectivity		
Confound adjustment & filtering	Report: • Method for detecting movement artifacts, movement-related variation, and remediation (e.g. 'scrubbing', 'despiking', etc). • Use of global signal regression, exact type of global signal used and how it was computed. • Whether a high- or low-pass temporal filtering is applied to data, and at which point in the analysis pipeline. Note, any temporal regression model using filtered data should have it's regressors likewise filtered.	Y ◯ N/A ◉
Multivariate method: Independent Component Analysis	Report: • Algorithm to estimate components. • Number of components (if fixed), or algorithm for estimating number of components. • If used, method to synthesize multiple runs. • Sorting method of IC's, if any. • Detailed description of how components were chosen for further analysis.	Y ◯ N/A ◯
Dependent variable definition	For seed-based analyses report: • Definition of the seed region(s). • Rationale for choosing these regions. For region-based analyses report: • Number of ROIs. • How the ROI's are defined (e.g. citable anatomical atlas; auxiliary fMRI experiments); note if ROIs overlap. • Assignment of signals to regions (i.e. how a time series is obtained from each region, e.g. averaging or first singular vector) • Note if considering only bilateral (L+R) merged regions.	Y ◯ N/A ◯

	• Note if considering only interhemispheric homotopic connectivity.	

Functional connectivity measure/ model	Report:
	• Measure of dependence used, e.g. Pearson's (full) correlation, partial correlation, mutual information, etc; also specify:
	  ○ Use of Fisher's Z-transform (Yes/No) and, if standardised, effective N is used to compute standard error (to account for any filtering operations on the data).
	  ○ Estimator used for partial correlation.
	  ○ Estimator used for mutual information.
	• Regression model used to remove confounding effects (Pearson or partial correlation).
		Y ◯ N/A ◯

Effectivity connectivity	Report: • Model. • Algorithm used to fit model. • If per-subject model, method used to generalize inferences to population. • Itemize models considered, and method used for model comparison.	Y ◯ N/A ◯
Graph analysis	Report the 'dependent variable' and 'functional connectivity measure' used (see above). Specify either: • Weighted graph analysis or, • Binarized graph analysis is used, clarifying the method used for thresholding (e.g. a 10% density threshold, or a statistically-defined threshold); consider the sensitivity of your findings to the particular choice of threshold used. Itemise the graph summaries used (e.g. clustering coefficient, efficiency, etc), whether these are global or per-node/per-edge summaries. In particular with fMRI or EEG, clarify if measures applied to individual subject networks or group networks.	Y ◯ N/A ◯
Multivariate modelling & predictive analysis		

Independent variables	Specify:
	● Variable type (discrete or continuous).
	● Class proportions in classification settings.
	● Variable dimension.
	  o For whole-brain prediction, this is a voxel count.
	  o For searchlight analyses, the exact number of voxels in the search region, not just a radius.
	  o Provide dimension before and after any feature selection and/or dimension reduction.
	If available, report on population stratification:
	● Information on how the target values relate to the population (e.g. male/female frequency or age distribution by group).
	● Specify how this is taken into account in the predictive model.
		Y ○ N/A ◉

Features extraction and dimension reduction	Specify the use of any: ● Feature transformation. ● Feature selection. ● Dimension reduction. When these techniques are data-driven, specify the procedures used to learn the parameters involved.	Y ○ N/A ◉
Model	For traditional multivariate analyses, report: ● Type of model, e.g. MANOVA. ● Assumptions made on the covariance structure, e.g. independence, or a common arbitrary covariance between groups. ● Statistic used to assess significance, e.g. Wilk's lambda, Hotelling-Lawley trace, etc. For predictive models, report: ● Type of model, e.g. Linear discriminant analysis, support vector machines, logistic regression, etc. ● For kernel-based methods (i.e. SVM) report type of kernel used, type and number of parameters needed to be estimated.	Y ○ N/A ◉
Learning method	Report:	Y ○ N/A ○

	• Figure-of-merit optimised. • Fitting method. • Parameter settings, those fixed and those estimated; specify how fixed parameter values were chosen. • How the convergence of the learning method is monitored.	
Training procedure	Describe: • Pipeline structure applied uniformly to all cases (e.g. that could be independently applied to a new case). • Method for hyper-parameter setting. • Data splitting (cross validation).	Y ◯ N/A ◉
Evaluation metrics: Discrete response	Describe the evaluation metrics that are to be computed. Always compute: • Accuracy. • If group sizes unequal, balanced (or average) accuracy. When there are only 2 classes, and one can be labeled "positive": • Precision (1 − false discovery rate). • Recall (sensitivity). • False positive rate (1-specificity). • F1 (incorporates both precision and recall). • Receiver operating characteristic (ROC) curves, e.g. summarised by area under the curve (AUC); AUC for only high specificity (e.g. false positive rates no greater than 10%) are also useful. When there are 3 or more classes: • Report the confusion matrix.	Y ◯ N/A ◉
Evaluation metrics: Continuous response	"Prediction R^2", the percentage of variance explained by prediction, computed as one minus the ratio of prediction sum-of-squares to total sum-of-squares. (Note this *is not* the squared correlation coefficient between true and predicted values).	Y ◯ N/A ◯
Evaluation metrics: Representational similarity analysis	Report the Kendall Tau statistic for each candidate model considered.	Y ◯ N/A ◉

Evaluation metrics: Significance	When possible use formal test to obtain P-value to assess whether evaluation metric is “significant” or consistent with noise.	Y ◯ N/A ◉
Fit interpretation	Procedure used to interpret the fit of the classifier, identifying the relative importance of the features (e.g. the weight vector in linear discriminant).	N

Table D.5. Results Reporting

Aspect	**Notes/Ontology**	**Mandatory**
Mass univariate analysis		
Effects tested	Provide a complete list of tested and omitted effects.	Y ◯ N/A ◯
Extracted data	• Define how voxels/elements were selected; if region is based on the same data, clarify how circularity was accounted for. • For any summary reported, give units. Ideally these are as interpretable as possible (e.g. percent change). • If reporting R^2 (coefficient of determination) clarify how nuisance variability is considered. For instance, in task fMRI the vast majority of variance is explained by slow temporal drift, and R^2 values for an effect of interest will be vastly different if computed with or without counting drift in the total variance.	Y ◯ N/A ◯
Tables of coordinates	Provide one table of coordinates including: • Contrast / effect to which it refers. • XYZ coordinate (with coordinate system, MNI, Talairach, noted in caption; also clarify whether peak or center-of-mass location). • Anatomical region (in caption or body text, describe source of labels, e.g. subjective, atlas, etc).	Y ◯ N/A ◯

	• P-value forming basis of inference (e.g. voxel-wise FWE corrected P; or cluster-wise FDR corrected P). • T/Z/F statistic (with degrees of freedom in table caption) • In caption, state whether coordinates are from whole brain, or from a specific constrained volume. • If cluster-wise inference is used, the cluster size. Report in mm^3 or, if in voxels, be explicit about the size of voxels. If a cluster statistic other than size is used (e.g. mass) it should be listed as well. • In caption or body text, note criterion for peak per cluster reporting; e.g. "one peak per cluster listed", or "up to 3 per cluster that are at least 8mm apart" (SPM default), etc.	
Thresholded maps	For each effect, provide images of maps of significant regions, ensuring that each caption describes: • Type of inference and the correction method, as well as form of any sub-volume corrections applied when computing corrected significance. • Include color bars; when presenting multiple maps in a figure, use a common color bar to ensure the results are comparable.	N
Unthresholded maps	Share, via supplementary material or repository: • Unthresholded statistic maps. • Optionally, the thresholded statistic maps. • Optionally, the effect size map (e.g. % BOLD change, % GM change).	Y ◯ N/A ◯
Extracted data	State whether data extracted from an ROI (e.g. to compute an effect size) is defined based on independent data, as otherwise it is susceptible to bias. If ROIs are circularly defined, best not to provide any statistical summary (i.e. P-values, R^2, etc).	Y ◯ N/A ◯
Spatial features	Report the • Size of the analysis volume in voxels, mm. • Spatial smoothness of noise (e.g. FWHM) and Resel count (if using Random Field Theory).	Y ◯ N/A ◯

Functional connectivity		
ICA analyses	Report the total number of components (especially when estimated from the data and not fixed). Report the number of these analyzed and the reason for their selection.	Y ◯ N/A ◯
Graph analyses: Null hypothesis tested	For graph-based methods, carefully state what is the null hypothesis of the test and how the statistic distribution under the null is computed.	Y ◯ N/A ◉
Multivariate modelling & predictiveanalysis		
Optimised evaluation metrics	Report the values obtained for the evaluation metrics chosen (see Evaluation Metrics, above), as well as any P-values to justify above-chance performance.	Y ◯ N/A ◯

Table D.6. Data Sharing [[Propose to omit for AJCN]]

Aspect	**Notes**	**Mandatory**
Reporting a data sharing resource		
Material shared	List types of images and non-imaging data provided. Report on the completeness of the data (e.g., number of subjects where all types of imaging, demographic, and behavioral data is available).	Y
URL, access information	Provide: • Stable URL or DOI. • Specific instructions on how to gain access. Specifically mention whether application must be vetted for particular intended research use (e.g. to preclude multiple users investigating the same question), or whether a research collaboration must be established.	Y

	• Cost of access.	
Ethics compliance	Confirm that the ethics board of the host institution generating the data approves the sharing of the data made available. Clarify any constraints on uses of shared data, for example, whether users downloading the data also need ethics approval from their own institution.	Y
Documentation	Provide URL to documentation, and specify its scope (e.g. worked examples, white papers, etc).	N
Data format	Report the format of the image data shared, e.g. DICOM, MINC, NIFTI, etc.	Y
Ontologies	Data organization structures, including Data Dictionaries and Schemas. Is the software using an established ontology?	N
Visualization	Availability of in-resource visualization of the imaging or non-imaging data.	N
De-identification	How, if at all, data are de-identified.	N
Provenance and history	Availability of detailed provenance of preprocessing and analysis of shared data.	N
Interoperability	Ability of a repository to work in a multi-database environment, availability of API's and ability to connect to analysis pipelines.	N
Querying	Mechanisms available for constructing queries on the repository (e.g. SQL, SPARQL).	N
Versioning	How users can check version of downloaded data and compare it to the current version at a later time.	N

Table D.7. Reproducibility

Aspect	Notes/Ontology	Mandatory
Documentation		
Tools used	Tool names, versions, and URLs.	Y ◯ N/A ◉
Infrastructure	Machine CPU model, operating system version, any use of parallelization.	Y
Workflow	Use of a workflow system, its version and URL.	N
Provenance trace	State whether detailed provenance information is available.	N
Literate program implementing results	Provide a URL linking to the relevant resource; for example, an ipython notebook implementing key analyses.	N
English language version	As the scientific lingua franca, documentation should be provided in English in addition to any other languages.	N
Archiving		
Tools availability	Note if tools are publically available.	N
Virtual appliances	Note if a virtual environment to facilitate a repeated analysis is available.	N
Citation		
Data	Provide permanent identifier if possible.	N
Workflow	Provide permanent identifier if possible.	N

SUPPLEMENTAL MATERIAL S2

Choosing and matching food-related stimuli

1. Visual Stimulation

The most frequently employed method for assessing brain response to food is the display of food images on a screen. This approach has several advantages. First, unlike oral and olfactory stimulation, the paradigms are relatively simple to create and require no specialized delivery systems. Second, outside of the scanner the sight of food is an indication of food availability and an important exogenous catalyst for promoting behaviors to acquire the food. Learning about the neural systems supporting these behaviors and the variables influencing these systems is important for understanding food choice, food craving and incentive motivation. However, in the scanning environment it is difficult to visually present actual food items. Food pictures provide a reasonable proxy but it is recommended that the pictures be made relevant to food availability and that the participant be made aware of this association. For example, responses in appetitive circuits (e.g. amygdala, orbitofrontal cortex and striatum) are enhanced when participants understand that the observed items will be made available for consumption after the scan (1, 2). Further, the anticipation of eating may interact with many variables of interest. For example, restrained eaters show increased food intake at a taste test when anticipating eating a subsequent meal (3).

A third advantage of using food pictures is the ease with which variables can be manipulated such as portion size, energy density, macronutrient content etc. This flexibility results from a greater ability to acquire and manipulate the images and from faster trial times, allowing greater number of presentations for inter-trial averaging and consequently the assessment of a greater number of factors. This advantage also promotes the creation of parametric designs and greater generalization because the researcher is not limited to the number of “channels” available in liquid and odor delivery devices (typically between 2 and 10). Ease and speed of image presentation also facilitates more involved designs where behavior is manipulated in the scanner, such as bidding for, or choosing between items.

Fourth, using food pictures avoids complications related to nutrient metabolism, satiation and post-ingestive signals that occur when participants are asked to repeatedly taste and ingest liquid stimuli over the course of the fMRI task.
Importantly, many factors that can be manipulated to create advantage can also be disadvantageous if not properly considered and controlled. For example, noise can be introduced by collating images that vary in macronutrient content, portion size, caloric load, familiarity, or healthfulness, because many neural circuits of interest are strongly affected by these variables. Likewise, a researcher might be interested in the influence of liking on brain response and sort responses by liking ratings. However, if all liked images are carbohydrate and all disliked images fat then it is impossible to determine whether macronutrient or liking drives differential brain response (see e.g.(4)) . Care should also be taken to equate images on visual perceptual parameters such as contrast, size, color etc.
Finally, it is imperative to use images of foods that are representative of the participants' diet. This is because the value of foods and their ability to recruit brain circuits is strongly tied to their nutritional properties (5-7), which are conveyed by metabolic signals to the brain. As such, foods that have been previously consumed by participants become calorie/nutrient-predictive stimuli capable of eliciting conditioned brain responses, whereas unfamiliar food images will not.

1.1 What should the control stimulus be?

It is also critical to choose an appropriate visual control stimulus. In comparisons of food with nonfood stimuli, low level visual features, such as luminance and contrast should be matched as should stimulus liking and familiarity. In making comparisons between food stimuli it is also important to consider whether one should match portion size, macronutrient content, actual or estimated energy density, actual or estimated cost and perceived healthiness. The appropriate control stimulus may depend on the research question, but generally contrasting high- versus low-calorie food images would provide the best comparison for studies interested in food reward.

1.2 Online resources for various image sets

There are several (food image) sets available online and it is recommended to use these if possible. Examples are the Food-pics database (http://eat.sbg.ac.at/resources/food-pics, (8)) and the Full4Health Image collection (http://nutritionalneuroscience.eu/index.php/11-resources/32-f4h-image-collection, (9)). A more detailed overview is given at the end of this document in Table S1.

2. Olfactory stimulation

Food aromas are potent cues, as anyone who has passed a French bakery on an empty stomach can attest. A meta-analysis comparing visual, olfactory and oral food-cue paradigms found that visual stimulation led to the most extensive and robust activations, with olfactory and oral stimulation as shared runner-ups (10). Surprisingly few studies have directly compared the impact of visual versus olfactory food stimulation. The olfactory cortex is in the limbic system and highly integrated with regions involved in valuation, interoception, drive and memory (11), and as such may have a privileged role in driving food seeking behavior (12). In addition, there are receptors for gut peptides, such as ghrelin, on neurons in the olfactory bulb and evidence that manipulation of these peptides influences olfactory perception (13-16). Hence, the olfactory system is more tightly integrated with physiology regulating metabolism than the visual system and an important target of investigation in relation to food seeking and consumption.

Another attractive feature of the olfactory system is that odors not only indicate food availability, but also food receipt, as olfaction is an integral part of the flavor percept (17). This means that one can use the same physical stimulus as a distal cue of food availability and a proximal cue of food receipt, which is important given the evidence for distinct circuits for anticipatory versus consummatory food reward (18). Further, unlike food pictures, which provide a representation of a food, food aromas, like the sight of real food, indicate availability. This is important because, while presentation of actual food items in the scanner is difficult, olfactometers enable precise delivery of odorants so that sensation can be time-locked with the BOLD response (19-22). With all of these advantages, the primary reason that aromas are not used more frequently is that odor delivery in the scanner is expensive and requires a significant level of expertise to run and maintain odor delivery devices (i.e. olfactometers). Moreover, the few commercial olfactometers that are available are prohibitively expensive and the assembly of one's own device requires a high level of engineering

expertise. However, if one is considering taking on these challenges, olfactometers can be found for purchase here: http://www.burghart-mt.de/en/, http://www.osmicenterprises.com/index.html. There are also several published papers describing the steps and equipment necessary to make your own olfactometer (19-22). Several special considerations for odorant delivery are discussed below.

2.1 To sniff or not to sniff

Olfactory sensation depends upon breathing and sniffing (23). Thus, it is important to instruct subjects to sniff in concert with odorant delivery. This is often accomplished with a "count-down cue" to time the sniff ("three, two, one, sniff") with the olfactometer programmed to deliver the odorant at the end of countdown (18). Another method is to use an airflow sensor at the nostrils and trigger delivery based on sniff initiation or breath inhalation at the nose (24). Although it has been argued convincingly that sniffing is a necessary part of the olfactory percept (25), sniffing is not strictly necessary for olfactory stimulation. Passive diffusion of volatiles to the olfactory epithelium has also been achieved by asking participants to effectively eliminate airflow from breathing in the nasopharynx by practising velopharyngeal closure (26). However, it has been noted that olfactory stimulation during velopharyngeal closure might not effectively activate all brain areas involved in processing olfactory information. This may be due to the fact that sniffing is an integral part of olfactory perception (7, 8). If sniffing is employed, it is an important factor to control, as sniffing itself results in neural activity in olfactory cortex (27), and may be accompanied by movement (if a participant interprets sniffing as a big inhale of breath, rather than small short inhalation). This can be done by measuring sniff vigor and volume with an olfactory mask coupled to a spirometer (28) and standard MRI equipment for measuring breathing rate.

2.2 Orthonasal versus retronasal stimulation

Orthonasal olfaction is associated with sensing foods at a distance and is dependent upon the odorant entering the external nares and flowing across the olfactory epithelium from front to back. In contrast, retronasal olfaction is associated with sensing foods being consumed and is dependent upon volatiles entering the nasopharynx from the oral cavity and flowing across the olfactory epithelium from

back to front (29). This is an important consideration because the direction and dynamics of odorant flow across the epithelium is thought to play a role in olfactory coding (30). In addition, orthonasal and retronasal olfaction map on to different aspects of ingestive behavior (anticipation versus consumption). Direct comparison of ortho and retronasal olfaction is complicated because eating is associated with other sensations (e.g. taste, temperature) and mouth movements. However, Hummel and colleagues created a delivery device where tubes are inserted into the nose with one ending at the external nares and another at the nasopharynx to simulate retronasal delivery (31). In so doing, the same physical stimulus can be used to stimulate both orthonasal and retronasal olfaction and differential BOLD response measured to the same odorant (32).

2.3 Design efficiency

A single presentation of an odor in an event-related design requires, at minimum, about 13 seconds, which limits the total number of events that can be presented. Regardless of the design, odor delivery should be short and inter-trial intervals relatively long (ideally even 30 s) because the uniquely rapid habituation to odors is an important consideration, as noted by Poellinger et al. (33). Block designs have higher power, as inter-stimulus time can be shorter and more presentations can be achieved (compare trial duration of 3 s in an on-off block design (34) to between 13-35 s in event-related designs (28, 35). Here habituation is dealt with by using an on-off design with pauses of no odorant delivery during an odor block. In determining how long each block should be it is also important to keep in mind that different cortical areas show different habituation patterns in response to odors (33).

2.4 What should the control stimulus be?

Since sniffing is associated with activation of many regions of interest, it is important to measure responses to odorless sniffs. Here a critical issue is ensuring that the air stream (or ambient air) is not contaminated. This means that the tubes carrying the air to the participants must be cleaned or replaced frequently. When making comparisons between food and non-food odors it is important to consider whether odorants are purely olfactory, such as phenylethal alcohol or if they contain a trigeminal component. Nonfood odors can also produce taste-like sensations. For example, many floral aromas are described as sweet, and to some may even be

edible (36). For this reason, it can be useful to have participants rate odorant edibility. Finally, odors are notoriously difficult to name. Differences in nameability between food and nonfood odors may lead to unanticipated confounds. We therefore recommend familiarizing participants with the odorants, providing their labels and measuring discriminability.

3. Oral stimulation

There are several advantages of using oral, rather than (or in addition to) visual and olfactory stimulation. First, the experience of pleasure derived from eating depends heavily on flavor perception, which results from the integration of distinct oral sensations of taste, retronasal olfaction, oral somatosensation and possibly chemesthesis (37). Individual differences in sensitivity of and preference for particular flavors (e.g. sweet concentration preference) and textures (e.g. fat sensing) play an important role in ingestive behavior. Therefore, examination of oral sensation is critical to understanding the neural circuits regulating feeding.

Second, oral stimulation occurs during food consumption and represents a distinct aspect of ingestive behavior from food acquisition. This is an important point because appetitive learning is driven by the generation of errors between predictions/actions and outcomes (38). Whereas visual and olfactory cues provide information important for prediction and action, oral sensory information provides information about outcome. As such, measuring brain response to both oral and extra-oral stimulation provides a more comprehensive assessment of so-called “food reward circuits” and can be valuable for interpreting findings. For example, response in the dorsal striatum to consuming small drops of milkshake is often negatively associated with body mass index (BMI), which has lead researchers to conclude that these striatal “reward” responses are hypo-responsive in obesity (39-42). However, striatal response to high-calorie food images correlates positively with BMI (43-48). This suggests that a more accurate interpretation is that BMI is associated with amplified prediction signals coupled with blunted outcome signals (49).

A third, and relatively unexplored advantage of assessing oral stimulation is that nutrients can be consumed and metabolized. This process is associated with a cascade of events including gastric secretions, hormone release and gut-to-brain neural signaling (i.e. the generation of vagal afferent signals) that are critical for associating food stimuli with their nutritive value (50-53). This provides the

opportunity to study the dynamic gut-brain axis, which is emerging as a major factor in understanding metabolism and ingestive behavior in health and in disease. One major hurdle towards this aim is the lack of information on timing, which makes it impossible to time-lock post-oral or metabolic events with brain response. Moreover, because internal state is changing over the course of the scanning session, it is important to measure variables related to internal state, like hunger, fullness and thirst.

Perhaps the biggest disadvantage of oral stimulation is that, to date, it has only been feasible to deliver liquids in the fMRI scanner. This limitation results because of difficulty with delivery and movement. The logistics of delivering a food item, even as small as a blueberry or an M&M™ to a subject lying in an fMRI scanner bore are not trivial. The bore is narrow and head coils often obscure or bar the mouth area. One could conjure an image of a reverse vacuum-like device, but this would need to be non-magnetic and designed so that it poses no risk of choking. However, even if one could successfully deliver the food item the movement caused by chewing has serious consequences for data quality. A basic assumption in fMRI analysis is that a given voxel corresponds to a given volume of brain tissue across time (54). Even small movements can lead to significant displacement and thereby reduce signal to noise, known as “partial volume effects”. However, more problematic is the fact that the movement associated with chewing is directly related to the event of interest. Thus, the data from a given voxel will be derived from two correlated sources, mouth movement causing displacements and BOLD response related to eating. Solving these issues would bring about a major step forward and is an important direction for research and development.

Notably, early studies that used water bolus methodology to measure regional cerebral blood flow with PET did not have the same magnitude of constraint. Voxels were larger, making small displacements less detrimental to SNR, temporal resolution was poorer, making precisely timed phasic stimulus delivery unnecessary, and participants were only inserted into the PET camera up to their forehead. In one study this allowed experimenters to hand-feed subjects squares of chocolate and measure brain response over a 60-sec window (rather poor temporal resolution compared to the 1-3-sec typical of current fMRI) as participants let the chocolate melt in their mouths (55). This eating experience is arguably more pleasurable than consuming small drops of liquid. Unfortunately, PET fell out of favor because it

requires the use of a radioligand, is extremely expensive, and the poor temporal resolution posed significant limitations for studies of rapid cognitive and perceptual operations. While the temporal disadvantage may not be as problematic for feeding research the cost and radiation exposure keep PET beyond the scope of most research programs. Although, MRI may be used to study blood-flow related responses to food with the use of arterial spin labeling techniques (ASL, (56)) delivery of solid food remains challenging. This is not an insignificant limitation given the importance of actions in motivated behavior and habitual responding, and bearing in mind the established literature documenting differences in oral sensation and metabolism in the consumption of liquid versus solid energy sources (57). Although these considerations are important, it is worth noting that liquid delivery is also not akin to drinking since only very small boluses of liquid are delivered at a time. Thus each “food” event is limited primarily to stimulation of oral sensation. Special considerations for oral stimulation are discussed below.

3.1 To swallow or not to swallow

Swallowing is an integral part of the act of eating, but it also introduces movement, which degrades data quality. Therefore, it is worth considering whether measures should be taken to limit swallowing or to de-correlate it from the event of interest. For example, if a researcher is interested in taste intensity perception then they could opt for a design where very small quantities of liquid are sprayed into the mouth negating the need to swallow (58), where the liquids are sucked out of the back of the mouth (59), or where participants are asked to postpone swallowing until a cue is presented, decoupling it from the onset of taste perception (60). One caveat associated with these methods is that taste buds are distributed across the entire oral epithelium, not only on the tongue, including the palate and pharynx (61), therefore taste stimulation is not comprehensive. However, if a researcher wanted to study the act of eating then a swallow is necessary, since only then is a food consumed. In addition, retronasal olfaction is a critical part of flavor (17) and is dependent on a swallow to move volatiles from oral to nasal cavity via the nasophayrnx (58, 62, 63).

3.2 Choosing a taste task

Generally, in fMRI studies there is the risk of participants falling asleep during passive stimulation tasks. This is often counteracted by engaging them with a task, such as making a perceptual rating. With oral stimulation somnolence is less of a risk, as participants are engaged in managing small drops of liquids dripping into the mouth while in a supine position. Moreover, asking participants to engage in a task hinders detection of sensory responses because taste representation is sparse and the insular “taste” cortex is multimodal. More specifically, unlike visual, motor, somatosensory, and auditory cortex where most neurons are engaged by the sensory input or motor output, many neurons in taste cortex do not respond to taste (64). For this reason, this multimodal cortex has been proposed as better defined as ingestive cortex (65). In addition, insular cortex is engaged by attention to body states (66). The combination of sparse taste representation and attentional activation may be equal or higher than sensory activation in gustatory cortex (35, 67, 68). Perceptual judgments (such as rating pleasantness or intensity) also influence the location of activation within gustatory cortex (69-71). There is also evidence that in the absence of a task the flow of stimulus information differs. For example during passive tasting there is stronger connectivity between the amygdala and the insula compared to performing a detection, identification or pleasantness rating task (70). Therefore, if the goal of the experiment is to understand a process related to sensation it is best to deliver the liquids passively.

Another important consideration is whether to deliver a cue to alert the participant to the impeding stimulus delivery. If no cue is provided and several different stimuli are used (e.g. milkshake and tasteless), the stimuli are generally unpredictable and subject to the generation of prediction errors. However, if a cue signals the identity of a forthcoming stimulus (e.g. a picture of a milkshake or water) then no error signal is generated. This is important because dopamine release is integral to error signal generation and individual variance in dopamine signaling may influence the sensory response (72). A similar situation occurs when no cue is used but the timing of delivery is random.

Another consideration is the quality of a cue used. If information about the stimulus is conveyed by the cue this too can have an important influence on the response (73-75).

3.3 What should the control stimulus be?

Early studies made the intuitive choice to use water as a stimulus e.g. (76, 77). However, water has a “taste” (78) and has an important physiological significance. As such, comparison of taste (e.g. sweet, sour, salty, bitter) minus water, may fail to isolate gustatory cortex. Subsequent studies then showed that water activates gustatory cortex as effectively as taste (79, 80). Water can also be a reinforcing stimulus itself, especially under a thirsty state. Two alternatives to water have been proposed. First, Frey and Petrides asked participants to move their mouths and swallow as if they were tasting (79). This method was successful in producing greater response to taste vs. mouth movement in chemosensory cortex. A second option developed by O’Doherty et al. is to administer a solution that contains the main molecular components of saliva (bicarbonate sodium and potassium chloride) (81). Of note, it is important not to describe the solution as “artificial saliva” but rather as “tasteless” or “control” to avoid negative responses. Tasteless solutions have become the gold standard. However, they do not have the typical viscosity associated with saliva and many participants report that they perceive taste. For this reason, it is best to create individualized tasteless solutions based on a “two alternative forced choice” procedure with several concentrations. Here, the subject is asked to “choose the solution that tastes most like nothing, or has the least taste” (see e.g. (82)). The concentration of the chosen tasteless solution may differ between individuals by a factor 8, which means that a single average concentration should not be used, but a determination of each participant’s tasteless is important.

References

1. Malik S, McGlone F, Dagher A. State of expectancy modulates the neural response to visual food stimuli in humans. Appetite 2011;56(2):302-9. doi: S0195-6663(11)00019-5 [pii] 10.1016/j.appet.2011.01.005.
2. Blechert J, Klackl J, Miedl SF, Wilhelm FH. To eat or not to eat: Effects of food availability on reward system activity during food picture viewing. Appetite 2016;99(Supplement C):254-61. doi: https://doi.org/10.1016/j.appet.2016.01.006.
3. Ruderman AJ, Belzer LJ, Halperin A. Restraint, anticipated consumption, and overeating. Journal of Abnormal Psychology 1985;94(4):547.
4. Charbonnier L, van der Laan LN, Viergever MA, Smeets PAM. Functional MRI of Challenging Food Choices: Forced Choice between Equally Liked High- and Low-Calorie Foods in the Absence of Hunger. PLoS ONE 2015;10(7):e0131727. doi: 10.1371/journal.pone.0131727.
5. de Araujo IE, Lin T, Veldhuizen MG, Small DM. Metabolic regulation of brain response to food cues. Current biology : CB 2013;23(10):878-83. doi: 10.1016/j.cub.2013.04.001.
6. Tang DW, Fellows LK, Dagher A. Behavioral and neural valuation of foods is driven by implicit knowledge of caloric content. Psychological science 2014;25(12):2168-76.

7. Veldhuizen MG, Babbs RK, Patel B, Fobbs W, Kroemer NB, Garcia E, Yeomans MR, Small DM. Integration of Sweet Taste and Metabolism Determines Carbohydrate Reward. Current Biology 2017;27(16):2476-85.e6. doi: https://doi.org/10.1016/j.cub.2017.07.018.
8. Blechert J, Meule A, Busch NA, Ohla K. Food-pics: an image database for experimental research on eating and appetite. Front Psychol 2014;5:617. doi: 10.3389/fpsyg.2014.00617.
9. Charbonnier L, van Meer F, van der Laan LN, Viergever MA, Smeets PA. Standardized food images: A photographing protocol and image database. Appetite 2016;96:166-73. doi: 10.1016/j.appet.2015.08.041.
10. Huerta CI, Sarkar PR, Duong TQ, Laird AR, Fox PT. Neural bases of food perception: Coordinate-based meta-analyses of neuroimaging studies in multiple modalities. Obesity 2014;22(6):1439-46.
11. Gottfried JA. Smell: Central Nervous Processing. Edtion ed. In: Hummel T, Welge-Lüssen A, eds. Taste and Smell An update. Basel: Karger, 2006:44-69.
12. Stevenson RJ. An Initial Evaluation of the Functions of Human Olfaction. Chemical Senses 2010;35(1):3-20. doi: 10.1093/chemse/bjp083.
13. Tong J, Mannea E, Aimé P, Pfluger PT, Yi C-X, Castaneda TR, Davis HW, Ren X, Pixley S, Benoit S. Ghrelin enhances olfactory sensitivity and exploratory sniffing in rodents and humans. Journal of Neuroscience 2011;31(15):5841-6.
14. Palouzier-Paulignan B, Lacroix M-C, Aimé P, Baly C, Caillol M, Congar P, Julliard AK, Tucker K, Fadool DA. Olfaction Under Metabolic Influences. Chemical Senses 2012;37(9):769-97. doi: 10.1093/chemse/bjs059.
15. Sun X, Veldhuizen MG, Wray AE, de Araujo IE, Sherwin RS, Sinha R, Small DM. The neural signature of satiation is associated with ghrelin response and triglyceride metabolism. Physiology & behavior 2014;136:63-73. doi: http://dx.doi.org/10.1016/j.physbeh.2014.04.017.
16. Sun X, Veldhuizen MG, Babbs AE, Sinha R, Small DM. Perceptual and Brain Response to Odors Is Associated with Body Mass Index and Postprandial Total Ghrelin Reactivity to a Meal. Chemical Senses 2016;41(3):233-48. doi: 10.1093/chemse/bjv081.
17. Small DM. Flavor is in the brain. Physiology & behavior 2012;107(4):540-52.
18. Small DM, Veldhuizen MG, Felsted J, Mak YE, McGlone F. Separable Substrates for Anticipatory and Consummatory Food Chemosensation. Neuron 2008;57(5):786-97.
19. Lorig TS, elmes DG, Zald DH, Pardo JV. A computer-controlled olfactometer for fMRI and electrophysiological studies of olfaction. Behavior Research Methods, Instruments, & Computers 1999;31(2):370-5. doi: 10.3758/bf03207734.
20. Johnson BN, Sobel N. Methods for building an olfactometer with known concentration outcomes. Journal of Neuroscience Methods 2007;160(2):231-45.
21. Lundström JN, Gordon AR, Alden EC, Boesveldt S, Albrecht J. Methods for building an inexpensive computer-controlled olfactometer for temporally-precise experiments. International Journal of Psychophysiology 2010;78(2):179-89. doi: https://doi.org/10.1016/j.ijpsycho.2010.07.007.
22. Sommer JU, Maboshe W, Griebe M, Heiser C, Hörmann K, Stuck BA, Hummel T. A mobile olfactometer for fMRI-studies. Journal of Neuroscience Methods 2012;209(1):189-94. doi: https://doi.org/10.1016/j.jneumeth.2012.05.026.
23. Sela L, Sobel N. Human olfaction: a constant state of change-blindness. Exp Brain Res 2010;205(1):13-29. doi: 10.1007/s00221-010-2348-6.
24. Hoffmann-Hensel SM, Freiherr J. Intramodal Olfactory Priming of Positive and Negative Odors in Humans Using Respiration-Triggered Olfactory Stimulation (RETROS). Chemical senses 2016;41(7):567-78.
25. Mainland J, Sobel N. The Sniff Is Part of the Olfactory Percept. Chemical Senses 2006;31(2):181-96. doi: 10.1093/chemse/bjj012.

26. Kobal G. Pain-related electrical potentials of the human nasal mucosa elicited by chemical stimulation. Pain 1985;22(2):151-63. doi: https://doi.org/10.1016/0304-3959(85)90175-7.
27. Sobel N, Prabhakaran V, Desmond J, Glover G, Goode R, Sullivan E, Gabrieli J. Sniffing and smelling: separate subsystems in the human olfactory cortex. Nature 1998;392(6673):282-6.
28. Zelano C, Bensafi M, Porter J, Mainland J, Johnson B, Bremner E, Telles C, Khan R, Sobel N. Attentional modulation in human primary olfactory cortex. Nature neuroscience 2005;8(1):114-20.
29. Rozin P. "Taste-smell confusions" and the duality of the olfactory sense. Perception and Psychophysics 1982;31(4):397-401.
30. Scott JW, Sherrill L, Jiang J, Zhao K. Tuning to odor solubility and sorption pattern in olfactory epithelial responses. J Neurosci 2014;34(6):2025-36. doi: 10.1523/jneurosci.3736-13.2014.
31. Heilmann S, Hummel T. A new method for comparing orthonasal and retronasal olfaction. Behavioral Neuroscience 2004;118(2):412-9.
32. Small DM, Gerber JC, Mak YE, Hummel T. Differential neural responses evoked by orthonasal versus retronasal odorant perception in humans. Neuron 2005;47(4):593-605.
33. Poellinger A, Thomas R, Lio P, Lee A, Makris N, Rosen BR, Kwong KK. Activation and habituation in olfaction--an fMRI study. Neuroimage 2001;13(4):547-60.
34. Bensafi M, Iannilli E, Gerber J, Hummel T. Neural coding of stimulus concentration in the human olfactory and intranasal trigeminal systems. Neuroscience 2008;154(2):832-8.
35. Veldhuizen MG, Small DM. Modality-Specific Neural Effects of Selective Attention to Taste and Odor. Chemical Senses 2011;36(8):747-60. doi: 10.1093/chemse/bjr043.
36. Veldhuizen MG, Nachtigal D, Teulings L, Gitelman DR, Small DM. The insular taste cortex contributes to odor quality coding. Frontiers in Human Neuroscience 2010;4:12.
37. Auvray M, Spence C. The multisensory perception of flavor. Consciousness and cognition 2008;17(3):1016-31.
38. O'Doherty JP. Reward representations and reward-related learning in the human brain: insights from neuroimaging. Current opinion in neurobiology 2004;14(6):769-76. doi: https://doi.org/10.1016/j.conb.2004.10.016.
39. Stice E, Spoor S, Bohon C, Small DM. Relation between obesity and blunted striatal response to food is moderated by TaqIA A1 allele. Science 2008;322(5900):449-52. doi: 322/5900/449 [pii] 10.1126/science.1161550.
40. Frank GK, Reynolds JR, Shott ME, Jappe L, Yang TT, Tregellas JR, O'Reilly RC. Anorexia nervosa and obesity are associated with opposite brain reward response. Neuropsychopharmacology 2012;37(9):2031-46. doi: 10.1038/npp.2012.51.
41. Green E, Jacobson A, Haase L, Murphy C. Reduced nucleus accumbens and caudate nucleus activation to a pleasant taste is associated with obesity in older adults. Brain research 2011;1386:109-17. doi: 10.1016/j.brainres.2011.02.071.
42. Babbs RK, Sun X, Felsted J, Chouinard-Decorte F, Veldhuizen MG, Small DM. Decreased caudate response to milkshake is associated with higher body mass index and greater impulsivity. Physiol Behav 2013;121:103-11. doi: 10.1016/j.physbeh.2013.03.025.
43. Stoeckel LE, Weller RE, Cook EW, 3rd, Twieg DB, Knowlton RC, Cox JE. Widespread reward-system activation in obese women in response to pictures of high-calorie foods. Neuroimage 2008;41(2):636-47. doi: 10.1016/j.neuroimage.2008.02.031.
44. Bruce AS, Holsen LM, Chambers RJ, Martin LE, Brooks WM, Zarcone JR, Butler MG, Savage CR. Obese children show hyperactivation to food pictures in brain networks linked to motivation, reward and cognitive control. IntJObes(Lond) 2010;34(10):1494-500.
45. Dimitropoulos A, Tkach J, Ho A, Kennedy J. Greater corticolimbic activation to high-calorie food cues after eating in obese vs. normal-weight adults. Appetite 2012;58(1):303-12. doi: 10.1016/j.appet.2011.10.014.
46. Frankort A, Roefs A, Siep N, Roebroeck A, Havermans R, Jansen A. Reward activity in satiated overweight women is decreased during unbiased viewing but increased when imagining

taste: an event-related fMRI study. Int J Obes (Lond) 2012;36(5):627-37. doi: 10.1038/ijo.2011.213.
47. Martin LE, Holsen LM, Chambers RJ, Bruce AS, Brooks WM, Zarcone JR, Butler MG, Savage CR. Neural mechanisms associated with food motivation in obese and healthy weight adults. Obesity(SilverSpring) 2010;18(2):254-60.
48. Stice E, Yokum S, Bohon C, Marti N, Smolen A. Reward circuitry responsivity to food predicts future increases in body mass: moderating effects of DRD2 and DRD4. Neuroimage 2010;50(4):1618-25. doi: 10.1016/j.neuroimage.2010.01.081.
49. Kroemer NB, Small DM. Fuel not fun: Reinterpreting attenuated brain responses to reward in obesity. Physiology & Behavior 2016;162:37-45. doi: http://dx.doi.org/10.1016/j.physbeh.2016.04.020.
50. Tellez LA, Medina S, Han W, Ferreira JG, Licona-Limón P, Ren X, Lam TT, Schwartz GJ, de Araujo IE. A Gut Lipid Messenger Links Excess Dietary Fat to Dopamine Deficiency. Science 2013;341(6147):800-2. doi: 10.1126/science.1239275.
51. Tellez LA, Han W, Zhang X, Ferreira TL, Perez IO, Shammah-Lagnado SJ, Van Den Pol AN, De Araujo IE. Separate circuitries encode the hedonic and nutritional values of sugar. Nature neuroscience 2016;19(3):465-70.
52. Frank S, Linder K, Fritsche L, Hege MA, Kullmann S, Krzeminski A, Fritsche A, Schieberle P, Somoza V, Hinrichs J, et al. Olive oil aroma extract modulates cerebral blood flow in gustatory brain areas in humans. Am J Clin Nutr 2013. doi: 10.3945/ajcn.113.062679.
53. Frank S, Linder K, Kullmann S, Heni M, Ketterer C, Cavusoglu M, Krzeminski A, Fritsche A, Haring HU, Preissl H, et al. Fat intake modulates cerebral blood flow in homeostatic and gustatory brain areas in humans. Am J ClinNutr 2012;95(6):1342-9.
54. Friston KJ, Williams S, Howard R, Frackowiak RS, Turner R. Movement-related effects in fMRI time-series. Magnetic resonance in medicine 1996;35(3):346-55.
55. Small DM, Zatorre RJ, Dagher A, Evans AC, Jones-Gotman M. Changes in brain activity related to eating chocolate: from pleasure to aversion. Brain 2001;124(Pt 9):1720-33.
56. Alsop DC, Detre JA, Golay X, Gunther M, Hendrikse J, Hernandez-Garcia L, Lu H, Macintosh BJ, Parkes LM, Smits M, et al. Recommended implementation of arterial spin-labeled perfusion MRI for clinical applications: A consensus of the ISMRM perfusion study group and the European consortium for ASL in dementia. Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine 2014. doi: 10.1002/mrm.25197.
57. Dhillon J, Running CA, Tucker RM, Mattes RD. Effects of food form on appetite and energy balance. Food Quality and Preference 2016;48(Part B):368-75. doi: https://doi.org/10.1016/j.foodqual.2015.03.009.
58. Eldeghaidy S, Marciani L, Pfeiffer JC, Hort J, Head K, Taylor AJ, Spiller RC, Gowland PA, Francis S. Use of an Immediate Swallow Protocol to Assess Taste and Aroma Integration in fMRI Studies. Chemosensory Perception 2011;4(4):163-74. doi: 10.1007/s12078-011-9094-4.
59. Nakamura Y, Goto TK, Tokumori K, Yoshiura T, Kobayashi K, Nakamura Y, Honda H, Ninomiya Y, Yoshiura K. Localization of brain activation by umami taste in humans. Brain Research 2011;1406(0):18-29.
60. Small DM, Gregory MD, Mak YE, Gitelman D, Mesulam MM, Parrish T. Dissociation of Neural Representation of Intensity and Affective Valuation in Human Gustation. Neuron 2003;39(4):701-11.
61. Lindemann B. Receptors and transduction in taste. Nature 2001;413(6852):219-25.
62. Buettner A, Beer A, Hannig C, Settles M. Observation of the Swallowing Process by Application of Videofluoroscopy and Real-time Magnetic Resonance Imaging—Consequences for Retronasal Aroma Stimulation. Chemical Senses 2001;26(9):1211-9. doi: 10.1093/chemse/26.9.1211.

63. Rabe S, Linforth RST, Krings U, Taylor AJ, Berger RG. Volatile Release from Liquids: A Comparison of In Vivo APCI-MS, In-mouth Headspace Trapping and In vitro Mouth Model Data. Chemical Senses 2004;29(2):163-73. doi: 10.1093/chemse/bjh021.
64. Scott TR, Plata-Salaman CR. Taste in the monkey cortex. Physiology and Behavior 1999;67(4):489-511.
65. de Araujo I, Geha P, Small D. Orosensory and Homeostatic Functions of the Insular Taste Cortex. Chemosensory Perception 2012;5(1):1-16. doi: 10.1007/s12078-012-9117-9.
66. Avery JA, Kerr KL, Ingeholm JE, Burrows K, Bodurka J, Simmons WK. A common gustatory and interoceptive representation in the human mid-insula. Human brain mapping 2015;36(8):2996-3006.
67. Veldhuizen MG, Bender G, Constable RT, Small DM. Trying to Detect Taste in a Tasteless Solution: Modulation of Early Gustatory Cortex by Attention to Taste. Chemical Senses 2007;32(6):569-81. doi: 10.1093/chemse/bjm025.
68. Veldhuizen MG, Douglas D, Aschenbrenner K, Gitelman DR, Small DM. The Anterior Insular Cortex Represents Breaches of Taste Identity Expectation. The Journal of Neuroscience 2011;31(41):14735-44. doi: 10.1523/jneurosci.1502-11.2011.
69. Grabenhorst F, Rolls ET. Selective attention to affective value alters how the brain processes taste stimuli. European Journal of Neuroscience 2008;27(3):723-9.
70. Bender G, Veldhuizen MG, Meltzer JA, Gitelman DR, Small DM. Neural correlates of evaluative compared with passive tasting. European Journal of Neuroscience 2009;30(2):327-38.
71. van Rijn I, de Graaf C, Smeets PAM. It's in the eye of the beholder: selective attention to drink properties during tasting influences brain activation in gustatory and reward regions. Brain Imaging and Behavior 2017. doi: 10.1007/s11682-017-9710-2.
72. Kroemer NB, Sun X, Veldhuizen MG, Babbs AE, de Araujo IE, Small DM. Weighing the evidence: Variance in brain responses to milkshake receipt is predictive of eating behavior. Neuroimage 2016;128:273-83. doi: 10.1016/j.neuroimage.2015.12.031.
73. Plassmann H, O'Doherty J, Shiv B, Rangel A. Marketing actions can modulate neural representations of experienced pleasantness. Proceedings of the National Academy of Sciences of the United States of America 2008;105(3):1050-4. doi: 10.1073/pnas.0706929105.
74. Yeomans MR, Chambers L, Blumenthal H, Blake A. The role of expectancy in sensory and hedonic evaluation: The case of smoked salmon ice-cream. Food quality and preference 2008;19(6):565-73.
75. Veldhuizen MG, Nachtigal DJ, Flammer LJ, de Araujo IE, Small DM. Verbal descriptors influence hypothalamic response to low-calorie drinks. Molecular metabolism 2013.
76. Small DM, Jones-Gotman M, Zatorre RJ, Petrides M, Evans AC. A role for the right anterior temporal lobe in taste quality recognition. JNeurosci 1997;17(13):5136-42.
77. Small DM, Jones-Gotman M, Zatorre RJ, Petrides M, Evans AC. Flavor processing: more than the sum of its parts. Neuroreport 1997;8(18):3913-7.
78. Bartoshuk LM, McBurney DH, Pfaffmann C. Taste of Sodium Chloride Solutions after Adaptation to Sodium Chloride: Implications for the "Water Taste". Science 1964;143(3609):967-8. doi: 10.1126/science.143.3609.967.
79. Frey S, Petrides M. Re-examination of the human taste region: a positron emission tomography study. European Journal of Neuroscience 1999;11(8):2985-8.
80. Zald DH, Pardo JV. Cortical activation induced by intraoral stimulation with water in humans. Chemical Senses 2000;25(3):267-75.
81. O'Doherty J, Rolls ET, Francis S, Bowtell R, McGlone F. Representation of pleasant and aversive taste in the human brain. Journal of Neurophysiology 2001;85(3):1315-21.

82. Veldhuizen MG, Bender G, Constable RT, Small DM. Trying to detect taste in a tasteless solution: modulation of early gustatory cortex by attention to taste. ChemSenses 2007;32(6):569-81.

Supplemental Table S1. Online food image resources

Name	Nr images Food/non-food	Categories	Culture	Ratings/measures	URL	Publication[1]
Databases developed for research						
F4H Image Collection	377/41	Sweet, savory, high and low calorie, non-foods (office utensils)	European (Netherlands, Scotland, Greece, Germany, Sweden, Hungary)	Liking, perceived calories and healthiness	http://nutritionalneuroscience.eu/index.php/resources/f4h-image-collection	Charbonnier et al. 2016 DOI 10.17605/OSF.IO/CX7TP
FRIDa	295/582	Natural, transformed, rotten food + various non-foods	Mediterranean	Valence, arousal, familiarity typicality ambiguity, perceived calorie-content, perceived immediate-edibility, perceived level of transformation	https://foodcast.sissa.it/neuroscience/	Foroni et al. 2013
Food.pics	896/314	sweet and savoury foods, high and low calorie, warm and cold dishes, processed and raw foods	German, North-American	palatability, desire to eat, complexity, recognizability, valence, arousal	http://eat.sbg.ac.at/resources/food-pics	Blechert et al. 2014

Name	Nr images Food/non-food	Categories	Culture	Ratings/measures	URL	Publication[1]
OLAF	96	Sweet high fat, salty high fat, fruit, veggies	Spanish	valence, arousal, dominance craving	https://zenodo.org/record/10202	Miccoli et al. 2014
Databases developed for AI learning[2]						
PFID		Fast food	United States		http://pfid.rit.albany.edu/	
Food-101	101.000	1000 categories			https://www.vision.ee.ethz.ch/datasets_extra/food-101/	
UEC FOOD 256	256	Food	Japanese	N/A	http://foodcam.mobi/dataset256.html	
Food-5K Food-11	2500/16643	Food and non-food	?	N/A	http://mmspg.epfl.ch/food-image-datasets	

[1] Blechert, J., A. Meule, N. A. Busch and K. Ohla (2014). "Food-pics: an image database for experimental research on eating and appetite." Frontiers in Psychology 5(617).
Charbonnier, L., F. van Meer, L. N. van der Laan, M. A. Viergever and P. A. Smeets (2016). "Standardized food images: A photographing protocol and image database." Appetite 96: 166-173.
Foroni, F., G. Pergola, G. Argiris and R. Rumiati (2013). "The FoodCast research image database (FRIDa)." Frontiers in Human Neuroscience 7(51).
Miccoli, L., R. Delgado, S. Rodríguez-Ruiz, P. Guerra, E. García-Mármol and M. C. Fernández-Santaella (2014). "Meet OLAF, a Good Friend of the IAPS! The Open Library of Affective Foods: A Tool to Investigate the Emotional Impact of Food in Adolescents." PLOS ONE 9(12): e114515.
[2] AI databases contain typical and less typical and noisy images, but may be a good source to select appropriate images from.