Title: A Methodological Checklist for fMRI Drug Cue Reactivity Studies: Development and Expert Consensus

Authors: Hamed Ekhtiari[1#], Mehran Zare-Bidoky[2,3], Arshiya Sangchooli[3], Amy C. Janes[4], Marc J. Kaufman[4], Jason Oliver[5], James J Prisciandaro[6], Torsten Wüstenberg[7], Raymond F. Anton[6], Patrick Bach[8], Alex Baldacchino[9], Anne Beck[7], James Bjork[10], Judson Brewer[11], Anna Rose Childress[12], Eric Claus[13], Kelly E. Courtney[14], Mohsen Ebrahimi[3], Francesca M. Filbey[15], Dara Ghahremani[16], Peyman Ghobadi Azbari[17,3], Rita Z. Goldstein[18], Anneke Goudrian[19], Erica Grodin[16], Paul Hamilton[20], Colleen A. Hanlon[21], Peyman Hasani Abharian[22], Andreas Heinz[7], Jane E. Joseph[23], Falk Kiefer[24], Arash Khojasteh Zonoozi[3,25], Hedy Kober[26], Rayus Kuplicki[1], Qiang Li[27], Edythe D. London[16], Joseph McClernon[5], Hamid R. Noori[28, 29], Max M. Owens[30], Martin Paulus[1], Irene Perini[20], Marc Potenza[26], Stephane Potvin[31], Lara Ray[16], Joseph P. Schacht[32], Dongju Seo[26], Rajita Sinha[26], Michael N. Smolka[33], Rainer Spanagel[34], Vaughn R. Steele[26], Elliot Stein[35], Sabine Steins Loeber[36], Susan F. Tapert[14], Antonio Verdejo-Garcia[37], Sabine Vollstädt-Klein[8, 38], Reagan Wetherill[12], Stephen J. Wilson[39], Katie Witkiewitz[40], Kai Yuan[41], Xiaochu Zhang[42], Anna Zilverstand[43]

1. Laureate Institute for Brain Research, Tulsa, OK, USA
2. Shahid-Sadoughi University of Medical Sciences, Yazd, Iran
3. Iranian National Center for Addiction Studies (INCAS), Tehran University of Medical Sciences, Tehran, Iran
4. Department of Psychiatry, McLean Hospital, Harvard Medical School, Belmont, MA, USA
5. Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC, USA
6. Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, SC, USA
7. Department of Psychiatry and Psychotherapy, Charité Campus Mitte, Charité - Universitätsmedizin Berlin, Charitéplatz, Berlin, Germany
8. Central Institute of Mental Health, Mannheim, Heidelberg University, Germany
9. St Andrews University Medical School, University of St Andrews, UK
10. Department of Psychiatry, Virginia Commonwealth University, VA, USA
11. Brown University, RI, USA
12. University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA, USA
13. Mind Research Network, Albuquerque, New Mexico, USA
14. Department of Psychiatry, University of California, San Diego, La Jolla, California, United States of America
15. Center for Brain Health, School of Behavioral and Brain Sciences, University of Texas at Dallas, Dallas, TX, USA
16. University of California, Los Angeles, Los Angeles, CA, USA
17. Department of Biomedical Engineering, Shahed University, Tehran, Iran
18. Departments of Psychiatry & Neuroscience, Icahn School of Medicine at Mount Sinai , New York, New York, USA
19. Department of Psychiatry, Amsterdam University Medical Center, University of Amsterdam, The Netherlands
20. Linköping University, Linköping, Sweden

21. Department of Cancer Biology, Wake Forest School of Medicine, Winston Salem, NC, USA
22. Brain and Cognition Clinic, Institute for cognitive Science Studies, Tehran, Iran
23. Department of Neuroscience, Medical University of South Carolina, Charleston, SC, USA
24. Department of Addictive Behaviour and Addiction Medicine, CIMH Mannheim, Germany
25. Mashhad University of Medical Sciences, Mashhad, Iran
26. Yale University, New Haven, CT, USA
27. Department of Radiology, Tangdu Hospital, Fourth Military Medical University, China
28. Max Planck Institute for Biological Cybernetics, Tuebingen, Germany
29. McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA, USA
30. Department of Psychiatry, University of Vermont, Burlington, VT, USA
31. Centre de recherche de l'Institut Universitaire en Santé Mentale de Montréal, Montreal, Canada
32. University of Colorado Anschutz Medical Campus, Aurora, CO, USA
33. Department of Psychiatry, Technische Universität Dresden, Dresden, Germany
34. Institute of Psychopharmacology, Central Institute of Mental Health, Germany
35. National Institute on Drug Abuse, Baltimore, MD, USA
36. Department of Clinical Psychology and Psychotherapy, University of Bamberg, Bamberg, Germany
37. Turner Institute for Brain and Mental health, Monash University, Australia
38. Mannheim Center for Translational Neurosciences (MCTN), Medical Faculty of Mannheim, University of Heidelberg, Mannheim, Germany
39. Department of Psychology, The Pennsylvania State University, University Park, PA, USA
40. Department of Psychology, University of New Mexico, Albuquerque, NM, USA
41. School of Life Science and Technology, Xidian University, Xi'an, Shaanxi 710071, People's Republic of China
42. University of Science and Technology of China, China
43. Department of Psychiatry & Behavioral Sciences, University of Minnesota, Minneapolis, MN, USA

#Corresponding Author:
Hamed Ekhtiari, MD, PhD,
Laureate Institute for Brain Research,
6655 South Yale Ave. Tulsa, OK 74136,
Tel/fax: 918.502.5100,
email: hekhtiari@laureateinstitute.org

Abstract (314 words)

Background: Cue reactivity is one of the most frequently used paradigms in functional magnetic resonance imaging (fMRI) studies of substance use disorders (SUDs). Although there have been many promising results elucidating the neurocognitive mechanisms of SUDs and SUD treatments, heterogeneities in participant characteristics, task design, craving assessment, scanning preparation and analysis decisions limit rigor and reproducibility in the field of fMRI of drug cue reactivity (FDCR), hampering clinical translation and synthesis by systematic reviews and meta-analyses. The aim of this consensus paper and Delphi study is to outline the important methodological aspects of FDCR studies and present a list of items and recommendations that should be taken into account when designing FDCR studies and reporting their results.

Methods: Fifty-five FDCR scientists from around the world participated. First, an initial checklist of items deemed important in FDCR studies was developed by a group of members from the ENIGMA Addiction Consortium based on a systematic review. Then, using a modified Delphi consensus method, all experts were asked to comment on, revise or add items to the initial checklist. Subsequently, experts were asked to rate the importance of the items.

Results: Thirty-seven items were proposed in the first round. After the commenting phase, seven new items suggested by experts were added and six were removed. The final 38 items that reached a defined consensus threshold in the rating phase were classified under seven categories and are considered important for conducting and reporting in any FDCR study.

Conclusion: This paper proposes a list of items and additional recommendations that researchers in the field of FDCR are encouraged to note and report when designing an FDCR study and reporting its results. Along with the presentation of a quality control checklist with Yes/No ratable items, various challenges in moving towards greater homogeneity in FDCR research and widespread use of FDCR to investigate SUDs and develop clinically relevant biomarkers are discussed.

Key Terms: fMRI, Cue Reactivity, Addiction, Craving, Biomarker, Harmonization, Checklist

Main Body: 8748 words

Seven Tables and Three Figures

1. Introduction

Substance use disorders (SUDs) affect hundreds of millions of individuals and are responsible for a substantial global burden of disease (Degenhardt et al., 2018). To improve translational research, as well as treatment and prevention, researchers and clinicians need a better understanding of the underlying neurocognitive mechanisms of SUDs (Hamed Ekhtiari, Faghiri, Oghabian, & Paulus, 2016). There is also a need for better brain-based biomarkers to study the course and treatment response in SUD (S. J. Moeller & Paulus, 2018). A powerful method for investigating brain function among people with SUDs is task-based functional magnetic resonance imaging (fMRI) of drug cue reactivity (FDCR) paradigms (Hamed Ekhtiari, Nasseri, Yavari, Mokri, & Monterosso, 2016). In FDCR studies, subjects are exposed to drug-associated cues in one or more sensory modalities while undergoing fMRI. These paradigms are popular among researchers, and to date, 318 published studies have used this method (Based on a database available at (H Ekhtiari, 2020a)).The results of these studies can help in understanding the neurobiology of SUDs, diagnostic classification of people with SUDs, discovering intervention targets, understanding the temporal evolution of the disease process, and monitoring the effectiveness of treatments and treatment outcomes; for more details see (Carmichael et al., 2018).

Despite the promising results of FDCR studies, multiple methodological heterogeneities limit rigor and reproducibility, hindering knowledge production and clinical translation by undermining generalizability and the ability to optimally conduct comparative reports and meta-analyses (Carmichael et al., 2018). There are many sources of potentially significant heterogeneity, including characteristics of participants, types of cues, durations of cue exposure, analysis methods, etc., such that the field would benefit from the establishment of best/standardized practices to enhance the generalizability of FDCR study outcomes and guide future research.

There are multiple ways of achieving more consistency, clarity, replicability and generalizability across FDCR studies: (1) Preregistered replicable protocols: Study protocols define the structure of a study and can include the sequence of different imaging sessions, data acquisition settings, and other methodological details (B. Casey et al., 2018; A. J. Jasinska, E. A. Stein, J. Kaiser, M. J. Naumer, & Y. Yalachkov, 2014); (2) Published drug cue databases: Drug cues in FDCR studies can be validated and standardized in terms of their average effects on arousal and valence, including affect and craving and activations in relevant brain areas/networks. They can also be matched to control stimuli in multiple respects. One way of achieving this goal would be the sharing and utilization of standardized cue databases. The first openly accessible database with 360 cues is a recently validated methamphetamine and opioid cue database (H. Ekhtiari, Kuplicki, Pruthi, & Paulus, 2020); (3) Data-analysis guides and pre-registered and standardized analysis pipelines: Preprocessing and analysis pipelines have significant effects on fMRI study results (Lindquist, 2020). Researchers can utilize credible recommendations, e.g., by the Committee on Best Practice in Data Analysis and Sharing (COBIDAS). Pre-

registration and open sharing of pipelines would also help in this regard, and moving towards consistent software and toolboxes is recommended (Esteban et al., 2020), and (4) Extant checklists: Many itemized checklists and recommendations have been developed to address different elements of research design and reporting in fMRI studies in differing degrees of specificity (for some instances, see (Bossuyt et al., 2015; Collins, Reitsma, Altman, & Moons, 2015; Gagnier et al., 2013; Moher, Liberati, Tetzlaff, Altman, & Group, 2009; O'Brien, Harris, Beckman, Reed, & Cook, 2014; Schulz, Altman, Moher, & Group, 2010; Tong, Sainsbury, & Craig, 2007; von Elm et al., 2008). Regarding fMRI analysis specifically, the COBIDAS proposes a checklist with the goal of enhancing the reporting of MRI studies (Thomas E. Nichols et al., 2017). Unfortunately, no checklist with clear recommendations for FDCR research design and reporting exists.

Most authoritative research checklists and guidelines represent consortium efforts. This expert consensus development helps to elucidate the research process and its various aspects and clarify opinion on the importance of these aspects. Furthermore, consortium involvement substantiates the claim of the checklist to represent a diversity of opinions in the field (Hsu & Sandford, 2007). One of the most common methods of achieving expert consensus is the Delphi technique. In the Delphi process, experts in the field approach consensus on a matter by participating in a series of commenting and/or item rating rounds with feedback (Jorm, 2015). An example of the use of this method in addiction sciences is a 2019 study to determine the significance of Research Domain Criteria in addiction medicine (Yucel et al., 2019).

The purpose of the present study was to develop an itemized checklist for FDCR researchers to use in future studies. Such a checklist should include items that are most important in study design and reporting to facilitate data sharing, enable future meta-analyses, increase replicability and validity, and increase the transparency of FDCR studies (Jorm, 2015). Using Delphi consensus technique, we aimed to develop this checklist based on the consensus of FDCR experts. Furthermore, we aimed for this to be a consensus paper, representing the views of experts who participated in the Delphi process discussing why and how various categories within the checklist affect FDCR research.

2. Methods

2.1. Scope of the Checklist

The items included in the checklist were those identified as being specifically important in FDCR studies. This checklist was developed to act as a standalone tool for describing methodological details considered to influence results of FDCR studies. The authors also detailed additional recommendations for each item that should be considered in order to increase the quality of reporting. The checklist can be used to increase transparency,

support replicability, improve quality of data acquisition, facilitate future data sharing between labs and make increasingly sophisticated meta-analyses possible.

2.2. Contributors

The contributions to this project were organized on two levels: a steering committee (SC) and a larger expert panel (EP). This method was chosen as it enables a small and collaborative group of leaders to flexibly and rapidly make decisions and resolve conflicts within the SC and lead the project to fruition. This approach also ensured that the voices of a much wider and more diverse group of international experts meaningfully impact the consensus process.

2.2.1. Steering Committee (SC)

The SC consisted of 14 individuals, including Anna Rose Childress, Hamed Ekhtiari, Rita Goldstein, Andreas Heinz, Amy Janes, Jane Joseph, Hedy Kober, F Joseph McClernon, Martin Paulus, Lara Ray, Rajita Sinha, Elliot Stein, Reagan Wetherill, and Anna Zilverstand. This group grew out of the ENIGMA addiction working group (https://www.enigmaaddictionconsortium.com) after a series of meetings in which substantial heterogeneity in FDCR studies, poor reporting of methods (insufficient for replication), and disagreements over the importance of various methodological parameters were discussed along with strategies to amend the situation. Furthermore, the initial members of the SC were asked to identify additional members chosen based on their scientific expertise and contributions to the FDCR literature.

The SC members outlined the scope of the Delphi project (Diamond et al., 2014) and its important questions, developed and approved the initial checklist of important methodological parameters, processed the comments and revisions, and led the authorship of this manuscript, all based on consensus.

2.2.2. Expert Panel (EP)

The panel of experts for this Delphi study was chosen based primarily on a systematic review of 318 published addiction-related FDCR studies in English (H Ekhtiari, 2020a). The main inclusion criteria were: a) appearing among the authors of at least 4 papers in the systematic review database; and, b) holding first, last or corresponding authorship position in at least one of the 318 papers. In addition, the members of the SC were asked to nominate candidates in the field of FDCR for inclusion within the EP. All SC members agreed on the list of experts before the invitation process.

All chosen experts received an email briefly outlining the importance, structure, and goals of this Delphi study, and were asked to state whether they wished to participate. To invite new participants, each candidate was contacted by email, and if there was no answer, two reminders were sent within roughly two-week intervals.

Those who decided to enroll received a further email with more details about how their feedback would be collected and used in the Delphi study, and then they formally entered the Delphi process. 76 EP candidates were contacted by email, 21 did not respond to the email, 6 had incorrect email addresses, 4 explicitly declined to participate, and 45 accepted to join the EP. Providing the study participants with information is not necessary for Delphi studies, which did not rely on explicit information and published data (Forsman et al., 2015; Jorm, 2015). Therefore, in this study, participants were asked to primarily rely on their prior knowledge of FDCR task design and methodology during the Delphi process; although they were provided with the list of the 318 studies included in the aforementioned systematic review so they could have viewed the relevant articles if needed.

2.3. Procedure

A general schematic of the methodology and its various stages is depicted in Figure 1.

2.3.1. Checklist Development Phase

To simplify consensus development and facilitate the process of finalizing a comprehensive but concise list of important methodological aspects of FDCR studies, the SC decided to begin the feedback rounds after developing a basic set of categories, items and their associated recommendations. Each item included one concise point of an aspect in the category in which it appeared (the final list of categories and items are available in the Tables 2-7 in the results section). There could also be some additional recommendations associated with each item. This basic structure evolved based on the initial feedback of the SC and a consideration of the methodological parameters commonly observed to be important to the studies included in the aforementioned systematic review. Upon completion, the items in the checklist questionnaire were pilot tested by rating 5 randomly selected FDCR papers with Yes/No ratings on whether the item was reported in the paper or not. Using data from the pilot testing analysis, the SC reworded and/or combined items that could not be easily given a Yes/No rating for inclusion in the revision phase.

2.3.2. Checklist Revision Phase

In the revision phase, 45 EP and 14 SC members were sent the checklist and were asked to add comments and suggest revisions to the existing items and their associated additional recommendations. They were also asked to suggest new items that they feel were overlooked, along with an explanation of why they thought the item should be included. They also were informed that there was no limit to the number of new items they could suggest. 41 members of the EP responded. 10 SC members also added additional comments in this phase. Overall, we reached a response rate of 85% across all participants (EP and SC).

In this revision phase, members of the EP and SC answered a short questionnaire (Chipchase et al., 2012) assessing their basic demographic information (age, sex, highest academic degree, country of residence, primary

affiliation/place of work), primary field of research (psychiatry, psychology, pharmacology, neuroscience, cognitive science, etc.), primary place of work (university, hospital, business, independent research institute, etc.), length of time spent in addiction medicine, and length of time spent specifically researching FDCR. These questions were asked to ensure we included a diverse field of experts (Table 1).

Comments for each item were processed by the SC. During processing, repetitive comments were removed, items with unclear meaning were reworded and those outside the scope of the study were removed (Hasson, Keeney, & McKenna, 2000) so that a list of clear and unique single-point notes extracted from the comments was obtained.

The notes obtained after the processing of comments were of three kinds: first, proposed changes to an existing item or its associated recommendations; second, adding or removing items; and, third, general changes or critiques regarding the checklist. The decision of applying or rejecting each note were made by the SC.

The modified version was sent once more to the SC and EP, and the members were asked to comment on the new changes. After receiving and applying their comments, the final version was approved by the SC members.

2.3.3. Checklist Rating Phase

In the second round, participants from the SC and EP were sent the edited checklist along with the newly added items. The participants were asked to rate each item in terms of importance in the methodology of FDCR studies, from 1 to 5 (87.5% completed the entire survey). The exact question was: “To facilitate visibility, replication and data sharing, how important is it to report this item?”. Also, for each additional recommendation, we asked: "Do you support the inclusion of this additional note as a recommendation to be considered in fMRI drug-cue reactivity studies?". Out of 59 members of the SC and EP, 49 (83%) participated in the rating phase.

To avoid a non-neutral center rating and encourage deliberation, ratings were termed “not important”, “slightly important”, “moderately important”, “highly important” and “extremely important”. The participants were allowed not to rate an item if they chose not to do so. The inclusion of each additional recommendation for each item could be rated “Yes” or “No”.

2.4. Data Analysis

All statistical analyses were conducted using RStudio (RStudio Version 3.4.1). For the rating phase, the average rating and the number of responses were calculated. Based on the distribution of the ratings, it was calculated whether items passed either of two importance thresholds. The more stringent threshold was a rating of 4 or 5 by at least 80% of participants (threshold 1, preregistered H Ekhtiari, 2020a), and the less stringent threshold was a rating of 3 or higher by at least 70% of participants (threshold 2) (dotted lines in Figure 2). For additional

recommendations, we defined those with a "Yes" rating by more than 50% of respondents as key ENIGMA ACRI checklist recommendations.

2.5. Ethical Considerations

To ensure informed autonomy, all contributors were informed about the study's aims and methods in the invitation email. Further notes within the questionnaire and emails during each round provided extra details, though the general study design and purpose remained unchanged. Members of both the SC and EP were invited to view the study's evolving open science foundation (OSF) page (H Ekhtiari, 2020a). All contributors were informed that they could terminate their participation whenever they wished. To ensure confidentiality, contributors were kept anonymous during both rounds of the Delphi survey, and comments and ratings were anonymized to all except the lead authors. Neither responding to the basic information collected nor commenting on and rating the checklist items was deemed to require the disclosure of personal information.

3. Results

3.1. Characteristics of Steering Committee and Expert Panel and Response Rates

Of the original 14 SC members and 45 EP members who accepted the invitation, 51 (86.4%) respondents completed the revision round of the ENIGMA ACRI Delphi questionnaire. In the rating phase, 49 (83%) sent back complete responses. Four members of the EP responded to neither the revision nor the rating phase and therefore, were subsequently removed from the EP.

The characterization of the SC and EP is provided in Table 1, which shows SC members were overall older than the EP without any significant difference (51.1±9.1 vs. 45.3±9.4); Sixty percent (5 SC, 28 EP) of respondents were male. Most respondents hold a PhD (79% SC, 80% EP) and MD PhD degrees (21% SC, 10% EP) and reported their primary field of research predominantly in neuroscience (29% SC, 44% EP) and psychiatry (43% SC, 34% EP). The professional affiliation of respondents were primarily universities (57% SC, 80% EP), hospitals (21% SC, 10% EP) and independent research institutes (14% SC, 10% EP). EP and SC members represented many SUD cohorts employing cue-reactivity (e.g., methamphetamine, cocaine, opioid, alcohol, tobacco, and gambling).

3.2. Delphi Process Results

A schematic of the entire study process and checklist development stages can be viewed in Figure 1.

3.2.1. Checklist Development Phase

After the systematic review of 318 articles, an initial list of suggestions for the overall structure of the checklist and important items was developed. This list consisted of 42 items in 5 categories, including 13 General Task

Information items, 9 Drug Cue Information items, 9 Control Cue Information items, 6 Craving Assessment Inside Scanner items and 5 Craving Assessment Outside Scanner items. After the discussions within the SC members, this initial draft was developed into a checklist with 7 categories and 37 items, including 8 Participant Characteristic items, 4 General fMRI Information items, 5 General Task Information items, 6 Cue Information items, 5 Craving Assessment Inside Scanner items, 4 Craving Assessment Outside Scanner items and 5 Pre- and Post-Scanning Considerations items. In addition, based on the SC inputs, a column with 27 additional recommendations corresponding to the different items were added to this checklist.

3.2.2. Revision Phase

Based on SC and EP comments on the checklist, one Participant Characteristic item, one Cue Information item, one Craving Assessment Inside Scanner item, one Craving Assessment Outside Scanner item and two Pre- and Post-Scanning Considerations items were excluded. New items were refined and added to the ENIGMA ACRI checklist following suggestions made by respondents to the "please suggest extra variable" question. Additional Participant Characteristic items were "Psychiatric Profile" and "Substance Use Profile-Main Drug". The additional General Task Information items were about "Temporal Information of the Event/Block Duration" and "Data and Resource-Sharing". The additional Pre- and Post-Scanning Considerations item was about "Other Tasks and Procedures in the Imaging Session". Also, one item was split into two items: item 4—Advanced Demographics I and item 5—Advanced Demographics II. Thus, in the rating round, there were 11 Participant Characteristic items, 4 General fMRI Information items, 7 General Task Information items, 5 Cue Information items, 4 Craving Assessment Inside Scanner items, 3 Craving Assessment Outside Scanner items and 4 Pre- and Post-Scanning Considerations items. The 22 additional recommendations were also expanded to 75, of which 69 were item-specific recommendations and 6 were category-specific recommendations. All the comments received in the revision phase are provided in an anonymized database in the project's OSF page.

3.2.3. Rating Phase

Rating phase results can be viewed in Figure 2. Respondents had a high rate of agreement on most checklist items and all items reached the less stringent threshold (over 70% of participants select "extremely important", "highly important", or "moderately important" rating). Most of the items also met the more stringent threshold of the consensus (over 80% of participants select "extremely important" or "highly important" rating). The following items (marked with † in Figure 2 and supplementary material 1) did not reach the most stringent *a priori* threshold of the consensus: *Advanced Demographics I, Advanced Demographics II, Handedness, Substance Use Profile-Main Drug, Substance Use Profile-Other Drug, Data and Resource-Sharing, Sources of Cues-Development, Drug and Neutral/Control Cue Content, Neutral/Control Matching to Drug-Cues for Physical Features, Craving Assessment Inside Scanner-Technology, Craving Assessment Outside Scanner-Time Points, Pre-*

scanning Training and Familiarization, Other Tasks and Procedures in the Imaging Session, and *Post-scanning Craving Management.* The result of the "Yes/No" rating of the 75 additional recommendations is presented in Figure 3. The result shows that 69 (92%) of recommendations reached the 50% threshold, but the following 6 did not (8%): *Interviewer Qualification, Motivation to Quit, Socio-economic Status, Body Mass Index, Menstrual Status, and Sleepiness/Alertness.* With the exception of revisions for minor grammatical and typographical errors, the checklist was not changed in the rating phase and no item or category changes were made as a priori planned (H Ekhtiari, 2020b). The actual ratings of the ENIGMA ACRI checklist are outlined in Tables 2 to 7. The final checklist is available as pdf or excel files in supplementary material 1 to 4.

4. Discussion

We developed a checklist resulting from a consensus process that represents the views of participating scientists regarding the important methodological aspects of conducting an FDCR study and reporting its results. These methodological aspects include seven distinct categories of core items and additional recommendations. The available literature for the role, some of these items and recommendations are playing in the results of FDCR studies, their interpretation, and syntheses across studies remain scant. Thus, the focus here was primarily on capturing and representing expert opinion on best practices in the field. This checklist should be seen only as a starting point to further gain empirical insight on the effect of these methodological details in the FDCR research. In addition, we hope the development of this initial checklist will set a standard for research practices and encourage scientific authorities in other areas of task-based fMRI to follow similar pathways to promote harmonization and transparency across different areas of functional human brain mapping (Verdejo-Garcia et al., 2019). As a secondary effect, it is hoped that journal reviewers and editors will apply aspects of this checklist in the peer review process for publication of relevant FDCR articles.

4.1. Participant Characteristics

Age and sex/gender passed our more stringent consensus threshold. In terms of age, FDCR studies can typically be divided into two major categories, those involving adolescents/emerging-adults (e.g.,Karoly et al., 2019; Rubinstein et al., 2011) versus those involving adults (e.g., Claus, Blaine, Filbey, Mayer, & Hutchison, 2013; Francesca M. Filbey, Schacht, Myers, Chavez, & Hutchison, 2009). This distinction is important in part due to development of the cortical circuitry that provides top-down control over bottom-up limbic systems that continue to mature throughout adolescence to early adulthood (B. J. Casey, Jones, & Hare, 2008). In addition, it is likely that age is correlated with years of substance use (James J. Prisciandaro et al., 2014), and it is generally thought that reactivity/neurocircuitry adaptations occur over time, leading to a source of confusion/confound

in the interpretation of studies. Within each of these developmental categories, there have been few reports of associations of age (in years) with FDCR, perhaps partly due to restriction of range. Because older adults have been routinely excluded from MRI studies that do not focus on aging, and taking into account shared neurodegenerative impacts of addiction and biological aging (Cheng et al., 2013), there is relatively little known about FDCR in older adults (e.g., > 65-years). In terms of sex/gender, multiple studies have demonstrated sex-/gender-related differences in FDCR, particularly in participants who smoke cigarettes (Francis J. McClernon, Kozink, & Rose, 2008; Wetherill et al., 2013), individuals with cocaine dependence (Joseph et al., 2019; Potenza et al., 2012), and those with gambling (Kober et al., 2016) and gaming (Dong, Wang, Du, & Potenza, 2018; Dong, Wang, Wang, Du, & Potenza, 2019; Dong, Zheng, et al., 2018) disorders, which may depend, in part, on menstrual cycle phase in women (T. R. Franklin, Jagannathan, Ketcherside, Spilka, & Wetherill, 2019).

Additional baseline demographics that passed the less stringent consensus threshold included education/intelligence, handedness, and race/ethnicity. These were rated as relatively less important than age and sex/gender because of a lack of published evidence for their association with FDCR. It is perhaps not surprising that education/intelligence has not been found to be reliably associated with FDCR, given the often low cognitive demands of a typical FDCR task (i.e., passively perceiving sensory stimuli). Although handedness can be a critical consideration in fMRI studies of cognition (e.g., language and memory (Cuzzocreo et al., 2009)), it does not appear to play a major role in the lateralization of FDCR. In the case of race/ethnicity, it is possible that the literature as a whole has not provided sufficient opportunity to detect associations between FDCR and race/ethnicity (which could be driven entirely by un-modeled environmental/contextual variables), as studies have historically contained too few non-white/Hispanic participants to provide adequate statistical power to detect such associations.

In terms of clinical characteristics, pattern/severity of substance use, addiction treatment status, last use and abstinence status, psychiatric profile, and study inclusion/exclusion criteria passed our more stringent consensus threshold. For example, in people who use cocaine, greater FDCR has been positively associated with addiction severity (e.g., Goldstein et al., 2009; Volkow et al., 2006) and predictive of relapse (Kosten et al., 2006; J. J. Prisciandaro, Myrick, Henderson, McRae-Clark, & Brady, 2013). Perhaps unsurprisingly, self-reported craving has also been associated with FDCR across a variety of drugs (Agnes J. Jasinska, Elliot A. Stein, Jochen Kaiser, Marcus J. Naumer, & Yavor Yalachkov, 2014). Though both treatment seekers and non-treatment seekers similarly demonstrate activation to drug cues in the ventral striatum (Chase, Eickhoff, Laird, & Hogarth, 2011), treatment seekers have lower activation to drug cues in a variety of non-limbic (e.g., frontal, cingulate, temporal) brain regions relative to non-treatment seekers (James J. Prisciandaro et al., 2014). This difference may be attributable to the expected availability of drug reward following cue exposure (Wertz & Sayette, 2001;

Wilson, Sayette, & Fiez, 2004), an additional variable of potential interest to consider for future consensus checklists. Abstinence has also been associated with increased drug cue reactivity (e.g., in dorsolateral PFC and occipital cortex) in cigarette-smokers (Engelmann et al., 2012) and (e.g., in the midbrain) in individuals with cocaine use disorder (Scott J Moeller et al., 2018) but needs further study. Although individuals with acute co-occurring psychiatric illness are typically excluded from FDCR studies, studies that include individuals with lifetime psychiatric illness, or more commonly subclinical symptoms of psychiatric disorders like depression and anxiety, may consider the potential enhancing effect of psychopathology on FDCR (Coffey et al., 2010). Finally, all study inclusion/exclusion criteria, including those already discussed, must be carefully considered. As just one example, psychiatric medications have been shown to alter FDCR (e.g., Goudriaan, Veltman, van den Brink, Dom, & Schmaal, 2013); information concerning psychiatric medications should be provided to readers in a standardized manner (e.g., in Chlorpromazine equivalents for neuroleptic medication), and attempts should be made to prevent or at least examine the potential impact of all medication classes on FDCR results via appropriate randomization and/or analytic strategies.

Additional clinical characteristics that passed our less stringent consensus threshold included substance use route of administration and other (i.e., co-occurring) drug use. FDCR studies often isolate participants by route of administration either purposefully or through convenience sampling (e.g., demographic homogeneity due to geographic location of participant recruitment). Nonetheless, care (e.g., in cue representation and covariate analysis) should be taken when combining groups of individuals who use the same drug (e.g., opioids) but self-administer it via different routes (e.g., intravenous vs. oral; McHugh, Fulciniti, Mashhoon, & Weiss, 2016) within the same sample or study. Though researchers typically aim to isolate a single or “primary” drug in FDCR studies, other drug use should also be considered, as sensory cues of the “primary” drug may nonetheless trigger neurobehavioral responses to multiple drugs, particularly when such drugs are commonly used simultaneously (e.g., cannabis and alcohol Clayton, Bailey, & Liu, 2019).

Another potentially important participant characteristic is genetics. This factor was not considered important for inclusion in this checklist by our participating experts, perhaps because the influence of genes on various aspects of FDCR remains understudied. Nonetheless, polymorphisms in dopaminergic, GABAergic, glutamatergic, cholinergic, opioidergic and other genes may affect FDCR results (e.g. (Bach et al., 2015; Blaine, Claus, Harlaar, & Hutchison, 2013; Chen et al., 2015; F. M. Filbey et al., 2008; Janes et al., 2012; Jorde et al., 2014; Kareken et al., 2010; Kuhn et al., 2016; F. J. McClernon, Hutchison, Rose, & Kozink, 2007; S. J. Moeller et al., 2013; J. P. Schacht et al., 2017; J. P. Schacht, Voronin, Randall, & Anton, 2018; Xu et al., 2013; Yang, Balodis, Lacadie, Xu, & Potenza, 2016). As FDCR methods are harmonized and more data sharing can occur, we

recommend that FDCR studies bank subject DNA for future genotyping so DNA will be available to support analyses including those involving polygenic risk scoring.

4.2. General fMRI Information

General fMRI detail reporting was considered extremely important by more than 80% of raters (Figure 2 and Table 3), and collectively, this category had the highest mean rating of all 7 reporting categories. Similarly, for additional recommendations items (Figure 3 and Table 3), the General fMRI Information category had the highest proportion of elements (89%) recommended by at least 75% of raters. This strong consensus is not surprising because these FDCR elements robustly influence data quality and variability. Below we discuss select items in each subcategory (acquisition, preprocessing, processing and reporting) to illustrate key points.

It was recommended with near unanimity that FDCR data acquisition details be reported using detailed checklists (e.g., COBIDAS Report T. E. Nichols et al., 2017 and/or Poldrack et al., 2008). Detailed reporting can increase experimental design consistency, can assist investigators new to the field in implementing robust methods, and can increase FDCR replicability and enable data sharing and meta-analyses. For example, it is very important to report hardware details that could affect fMRI signals in different ways across the brain, such as the number of head-coil channels (e.g., 32 versus 8). Indeed, a “coil-bias” effect has been documented by several studies: one study determined that a 32-channel coil was more sensitive than an 8-channel coil for detecting cortical surface signals during a finger-tapping paradigm but less sensitive for detecting subcortical activations (Albrecht et al., 2010). A more recent and comprehensive study investigating coil bias determined that head-coil channel number affects volumetric and diffusion measures as well as resting-state BOLD signal measures, with channel number strongly affecting BOLD signals in posterior visual and default mode network areas (Panman et al., 2019). Also, while most current FDCR studies are conducted on 3 Tesla systems, other factors will need to be considered in future as more studies are conducted at higher magnetic field strengths. For example, a preliminary (bioRxiv) communication compared fMRI results on a monetary incentive task in 8 subjects scanned both at 7 and 3 Tesla (Colizoli, De Gee, Van Der Zwaag, & Donner, 2020). The study reported that 7 T scans yielded higher effects than 3 T scans in small subcortical nuclei relevant to FDCR studies, including the substantia nigra, ventral tegmentum, and locus coeruleus.

Detailed reporting of preprocessing parameters using the structured checklists noted above was unanimously endorsed. Preprocessing parameters such as the spatial smoothing Full-Width Half Maximum (FWHM) value should be reported because they affect statistical inferences. In this regard, a meta-analysis of fMRI tasks involving rewarding stimuli revealed that the spatial smoothing value affects apparent nucleus accumbens volumes and anatomical positions (Sacchet & Knutson, 2013). There was near unanimity in the endorsement of reporting of artifact detection methods and motion thresholds for data exclusion. There was substantial but

lower agreement (79%) regarding reporting of group motion parameters during FDCR drug versus neutral cue blocks, which, if differing by group could confound data analyses. This version of the checklist did not explicitly include denoising protocols, which when applied can affect task-related fMRI data by reducing noise and signal (Mayer et al., 2019). Future checklist versions might consider including denoising procedures, which hopefully will evolve to more selectively attenuate noise.

For data processing pipeline procedures, there was near unanimity (98-100%) for most elements, including recommendations to report on single-subject and group-level processing steps, nature of GLM analyses (random, mixed, fixed), whether covariates or demeaning are used, software tools used, multiple comparisons corrections applied, and regions of interest specifications, if applicable (e.g., manually drawn, atlas-based, dataset-determined). Reporting of pre-registration of data processing methods and of effect sizes were considered important but lower priorities. The sample sizes commonly used in task-based fMRI research tend to generate small-to-medium effect sizes (Cohen's d<0.8, Poldrack et al., 2017). However, it seems likely that effect size reporting will be considered a higher priority in future.

There was greater variability across fMRI Data Reporting elements with more than 80% of raters endorsing detailed reporting of second-level maps or activation foci within groups, whole-brain contrasts, beta-weights during craving and neutral conditions, and inclusion of whole-brain maps even in studies not using standard analytic methods, to facilitate data comparisons across studies. Other reporting elements were considered somewhat lower priorities, including providing non-thresholded statistical maps and stating whether data have been or will be deposited in publicly available repositories, which can be challenging given inconsistencies in repository reporting requirements. Most (78%) raters recommended that reporting go beyond the use of checklists by providing as much experimental detail as possible. Undoubtedly, over time, as more data are aggregated in meta-analyses and as additional factors are determined to impact FDCR data effect sizes, such factors will be added to the reporting checklist.

4.3. General Task Information

An adequate description of the task is essential for assessing the appropriateness of the analytical approaches and for understanding the results. As such, it is not surprising that this category was classified as almost as important as the "Participant characteristics" and "General fMRI information" sections (Figures 2 and 3 and Table 4). Specifically, for a transparent report, it is not sufficient to assign the used design to a certain design category (e.g., blocked, event-related or mixed blocked/event design). Due to its fundamental implications for modeling and design efficiency, it is necessary to show the exact temporal structure of the task, specifically the order, the onset, the spacing and the duration of stimuli.

In this sense, the duration of stimulation not only defines the design category but also significantly influences the amplitude of the evoked hemodynamic response. Compared to simple cue presentation experiments, designs with more complex trial structure interactions between affective and cognitive trial components may influence cue-associated brain response and functional connectivity (Lorenz et al., 2013). A detailed description of the timing within the trial and the related modeling approach is therefore desirable. In order to sample hemodynamic responses in event-related designs optimally especially in the fast event related designs, it is necessary to jitter the inter stimulus intervals (ISI). The distribution of the resulting ISIs is important for the design efficacy and should thus be described in detail (Dale, 1999; Josephs & Henson, 1999).

Beyond this micro-timing information, information like the overall duration of the scanner session, the duration of the experimental paradigm, the start in relation to the onset of the scanning session and the position within the order of possible additional paradigms are also of interest since multi-paradigm fMRI experiments are known to be prone to carry over and order effects (A. J. Jasinska et al., 2014). Reporting should further mention whether and how the order and timing of stimulus presentation were optimized. If appropriate, all of this information could be provided in compact and understandable ways by means of graphic displays.

In the interest of a complete description of the experimental setup, we suggest providing detailed reporting of the technical details of stimulation, especially if a less common sensory modality was used (e.g., gustatory), including the physical stimulation parameters (e.g., substance concentration, presentation latencies, etc.) and the equipment used should be described in detail.

Whether participants are instructed to interact passively or actively with the cue, to allow or to regulate craving, is an important component of instruction, influencing the experimental setting. To enable the reader to judge the clarity of the instruction, the verbatim instructions given to the participant should be included. Especially in passive tasks, additional default processes could occur (Gusnard, Raichle, & Raichle, 2001), with potential harm for the specificity of statistical analyses. Therefore, the chosen activity level and possible attempts to quantify participants' compliance, attention, and vigilance should be described in detail.

Although 58% of the panel experts were of the opinion that the task code and stimuli sharing item (Table 4) should be included in the checklist, its importance was rated lower (3.31) compared to the other items. This is particularly surprising given the intense contemporary discussion about reproducibility in fMRI research (Poldrack et al., 2017). In our opinion, authors should still report whether they have used an open scientific platform to provide task-related data (stimuli, software) to the imaging community. Therefore, the manuscript should include, where appropriate, information on access points and conditions of access in accordance with the FAIR principles for data exchange (https://www.force11.org/fairprinciples).

4.4. Cue Information

This category was considered to be of moderate importance on the whole, but several items and recommendations received near-unanimous support (Figures 2 and 3 and Table 5). Item rating means ranged from 4.07 (for the description of the validation extent of the cues) to 4.77 (for the description of the sensory modality of cues). Multiple drug- or control-cue-related aspects of FDCR studies may affect study outcomes (A. J. Jasinska et al., 2014). The most important factor may be the description of the sensory modality of drug and control cues. Although cues in different sensory modalities often induce distinct brain activation profiles (Yalachkov, Kaiser, Gorres, Seehaus, & Naumer, 2013), many studies do not clearly describe the sensory modality of their utilized cues. Depending on the sensory modality, there are various parameters that may be considered for drug cues and control stimuli. For instance, for pictorial cues, it is recommended that authors provide details regarding picture luminance, complexity (including human presence), hue and saturation. For auditory cues, it is important to consider factors such as intensity and frequency (H. Ekhtiari et al., 2020; Lang, 2005; Manoliu et al., 2020).

Furthermore, these parameters may be used to "match" drug cues and control stimuli (or those belonging to other cue categories in a study). Matching is done to minimize the effects of these other factors on the differential activation patterns elicited by different cue types. Also, cues can be matched based on their standardized arousal, valence, or craving induction scores (H. Ekhtiari et al., 2020; Macatee, Carr, Afshar, & Preston, 2020; Manoliu et al., 2020).

Another important but often overlooked factor limiting replicability and interpretation of FDCR studies is confusion over the sources of utilized cues, how they were obtained or developed, and whether they have been validated (i.e., shown to elicit a certain range of arousal, affective, or craving related responses in individuals). Providing cue validation details was considered very important, but providing cue development details was not. Nevertheless, there was near-unanimous support that researchers should consider reporting the exact source of their cues and how their cues were developed from this source, if applicable. This suggests that participants broadly considered this a significant aspect of a FDCR study. Even in cases where authors are using cues developed or validated in another published study, it is still desirable to provide minimal development and validation details in addition to references.

Though not always optimal, using cues from already validated and widely used cue databases may save researchers considerable resources and improve consistency across studies. There have been recent attempts to develop large pictorial cue databases to address these issues (H. Ekhtiari et al., 2020; Macatee et al., 2020; Manoliu et al., 2020). These databases include cues that have been developed in a methodologically consistent manner, and whose craving and arousal elicitation effects have been formally studied. The best FDCR cue

databases include neutral stimuli as well drug cues that are matched according to various characteristics (Lang, 2005; Stritzke, Breiner, Curtin, & Lang, 2004). Newer databases with a greater focus on drug-cue reactivity studies have become available in recent years (Billieux et al., 2011) and large developing cue banks may even contain multiple drug cues and control stimuli types (H. Ekhtiari et al., 2020).

The exact content of the cues can also influence multiple dimensions of cue reactivity. Drug cues may depict the drugs themselves, drug paraphernalia, individuals preparing or using drugs, or spaces where drug use is likely. Differences in the content of cues (drug vs. drug use tools vs. drug use actions) may recruit different brain areas, and this may have implications for how these cues link to drug-seeking behavior (Zeng et al., 2018). It may be important to consider this aspect of cue selection when designing studies, as certain cue contents may be more appropriate to test some but not all hypotheses.

Additionally, among recommendations in this category, there was widespread agreement on the importance of describing substance delivery methods in studies where a substance is administered as a cue, prior cue exposure, and cue tailoring. Studies where a substance is directly administered (usually in small amounts) remain relatively rare in the field of FDCR as a whole. However, given the popularity of these paradigms in some fields (such as in tobacco use disorder and alcohol use disorder) and the large variety of substance delivery mechanisms used, it is recommended that researchers describe their delivery mechanisms in detail and cite the relevant literature when possible (Lindsey et al., 2013; J. P. Schacht, Anton, & Myrick, 2013; Wall et al., 2017). Prior exposure of participants to cues is also important. Some brain regions may rapidly habituate to specific drug cues, decreasing their reactivity to them, even in the absence of a reduction in self-reported craving (Hamed Ekhtiari, Rayus Kuplicki, Robin Aupperle, & Martin P Paulus, 2020). Lastly, personalized tailoring of cues presents unique challenges and opportunities in FDCR studies. While it potentially leads to maximal cue reactivity in all participants, it also leads to heterogeneous cues that present problems for generalizability and interpretation. It is recommended that authors specify whether tailoring was conducted (if there is room for misunderstanding), and to present precise details for how tailoring was conducted for each participant. While all individual cues in a study may be tailored (Seo et al., 2013), tailoring can be particularly applied based on the participant drug of choice in samples of individuals who use multiple drugs (Hanlon et al., 2018). Tailoring of drug-related messages meant to encourage drug-use cessation is another possibility (Chua et al., 2011). Tailoring for gender/race/ethnicity is another area that is not well explored yet.

4.5. Task-related Assessments

Integration of self-report, behavioral or physiological measures as part of FDCR is commonplace (Li et al., 2012; F Joseph McClernon et al., 2016; F Joseph McClernon, Hiott, Huettel, & Rose, 2005). Yet perhaps because fMRI is the primary focus of these papers, the methodological details of other task-related assessments (e.g., self-

reported drug use, craving/urge) would be standard to report in behavioral research papers are sometimes excluded. Details of items, ratings, and recommendations are presented in Figures 2 and 3 and Table 6. A recent review of opioid craving measurement identified many different instruments for assessing opioid craving; however, many had not been tested for reliability and validity (Kleykamp et al., 2019). Harmonization and validation of the instruments used for subjective reporting of drug craving should be considered a priority in the field.

The timing of additional task-related assessments received high ratings of importance overall, with universal agreement that reporting the time period considered for in-scanner tasks (i.e., urges while viewing the image versus afterward) is important. This information is critical for proper interpretation of the nature and magnitude of the response. There is evidence that the effects of imagery-based cue procedures on urge may persist for extended periods of time (e.g., 15-30 minutes) (Heishman, Lee, Taylor, & Singleton, 2010; Heishman, Saha, & Singleton, 2004), but the duration of effects from the brief image presentations commonly employed in FDCR are largely unknown. Indeed, given that many FDCR paradigms rely on random/pseudo-random presentation of interleaved images from varying categories, an implicit assumption of most research is that the duration of these effects is brief. Continued research on this topic examining the validity of this assumption is critical and could conceivably lead to the development of formal guidelines for such assessments depending upon the nature of the study, the cue modality employed, and the specific question being asked.

As with timing, there was near-universal agreement that detailed reporting of the content of both in-scanner and out-of-scanner assessments is important. This is perhaps particularly critical for in-scanner assessments where historically, research has relied more heavily on single-item measures and may not have been subjected to the same rigorous examination of psychometric properties common for traditional self-report measures (Franken, Hendriks, & van den Brink, 2002; Heishman, Singleton, & Liguori, 2001; Tiffany & Drobes, 1991). Although the general construct is frequently reported (e.g., urge, liking), reporting the exact phrasing is less common despite long-standing recognition that subtle differences in wording can impact participant interpretation and study outcomes (Kozlowski, Mann, Wilkinson, & Poulos, 1989; Kozlowski & Wilkinson, 1987). This issue will be particularly important as research continues to explore covariation of constructs with brain activation. Indeed, research has already shown that patterns of activation may be at least partly dependent upon urge strength (Wilson & Sayette, 2015). It should be recognized, however, that subjective “craving/urge” is highly variable and situation-specific (e.g. scanner versus bar). As such, brain activation to cues during fMRI might be less variable and, in fact, was one of the reasons for the initial development of FDCR paradigms.

There was also agreement about the importance of reporting hardware (e.g., button box and response pad) used for collection of these assessments. This may be particularly critical for research where response time is

examined as a primary or secondary outcome. An extensive body of literature documents the existence of substantial variability in the accuracy of data collection devices outside the scanner (Plant, Hammond, & Whitehouse, 2003; Plant & Turner, 2009; Segalowitz & Graves, 1990). To our knowledge, no similar evaluation of variability in the accuracy of common MRI-compatible devices has been conducted.

Comparatively fewer experts (61%) recommended the inclusion of other physiological measures relative to other topics under consideration. One likely reason is that to date, these measures rarely have been included in FDCR studies. Nonetheless, examination of heart rate, skin conductance and other peripheral physiological measures are standard in the broader drug cue reactivity (Carter & Tiffany, 1999). It is certainly plausible that changes in peripheral physiology could influence findings, particularly for certain types of imaging (e.g. arterial spin labeling). Moreover, inclusion of peripheral signals as covariates is becoming standard in resting scans in light of evidence showing it can alter connectivity maps (Chang, Cunningham, & Glover, 2009), and there is little reason to believe these concerns should not extend to task-based scans. While it may be premature to make formal recommendations for inclusion of peripheral measures at this time, continued exploration of this topic is critical and may reveal a need for inclusion in later instances.

4.6. Pre-Post Scanning Consideration

Of the pre/post scanning considerations, pre-scanning drug and smoking consumption was the only metric rated as moderately to extremely important by all reviewers (Figures 2 and 3 and Table 7). This is likely due to the impact that both abstinence and recent substance use can have on cue-induced craving and brain function. The length of abstinence also matters as studies generally support the idea that short-term abstinence enhances cue-reactivity relative to satiety (Gloria et al., 2009; Janes et al., 2009; Lou, Wang, Shen, & Wang, 2012; F. J. McClernon, Kozink, Lutz, & Rose, 2009; Parvaz, Moeller, & Goldstein, 2016), which mirrors preclinical findings (Lu, Grimm, Hope, & Shaham, 2004). In contrast, longer-term abstinence is associated with reduced cue-reactivity (Parvaz, Moeller, &Goldstein, 2016). Further, deprivation and cue presentations may have independent, interactive effects on subjective reports of craving (Bailey, Goedeker, & Tiffany, 2010), supporting the need to clearly indicate the conditions under which cue reactivity is evaluated. There is also a need to report the recency of other substance use and medications as they may influence subjective cue responses and the physiology underlying the fMRI signal.

Other recommendations include indicating whether participants have had prior cue exposure in the context of the study. This is important as habituation to emotionally evocative stimuli has been identified in specific brain regions (Fischer et al., 2003), yet not in all participant groups, particularly those who may be more reactive to the cue content (Siegle, Steinhauer, Thase, Stenger, & Carter, 2002). While within-session habituation is a potential confound (Hamed Ekhtiari, Rayus Kuplicki, Robin Aupperle, & Martin P. Paulus, 2020; Fischer et al.,

2003), cues continue to elicit subjective craving and comparable brain activity patterns over repeated sessions separated by longer durations (2-3 weeks) (T. Franklin et al., 2011; LaRowe, Saladin, Carpenter, & Upadhyaya, 2007; Joseph P Schacht et al., 2011). However, this finding has not been supported in all studies (Hamed Ekhtiari, Rayus Kuplicki, Robin Aupperle, & Martin P. Paulus, 2020), thus supporting the need to clearly report details surrounding previous cue presentations. Reporting drug expectancy is also recommended as recent work suggests that participant expectations influence cue-reactivity and related circuitry (Gu et al., 2016; McBride, Barrett, Kelly, Aw, & Dagher, 2006; Perry et al., 2019).

Several elements of pre/post scanning considerations did not reach a stringent consensus. Pre-scanning training and familiarization were ranked as highly important by ~60% of respondents, as some reviewers felt this was such a fundamental aspect of good scientific procedures that it was assumed that study participants were familiarized in some way with the task. Additionally, most cue-reactivity tasks involve passive exposure to cues, which unlike complex behavioral tasks, do not require extensive pre-scan training. However, such familiarization may also impact potential habituation and expectancy, which would support the need to report based on the discussion points above. The need to report other tasks and procedures in the imaging session was similarly ranked and did not reach a stringent consensus. It is plausible that the lack of reporting of other tasks may imply a singular focus on cue-reactivity with no potential influence for the other tasks. That said, reporting tasks that have the potential to influence cue-reactivity is considered best practice. Post-scanning craving management was rated the lowest element with less than 35% of the respondents ranking it as extremely/highly important, perhaps because it is viewed as more of an ethical consideration that would be considered by local institutional review boards rather than a factor that would impact cue-reactivity directly.

4.7. Conclusion and Future Directions

As demonstrated by the consensus of the experts participating in this study and the review of the literature summarized in the discussion above, FDCR studies have a vast methodological parameter space in which many impactful choices regarding study design and reporting can be made. The methodological heterogeneity makes replication and generalizability difficult, hampers data synthesis and clinical translation, and makes transparency in reporting methodological details an urgent need.

This paper presents the results of an international effort to develop a checklist of important items and recommendations that FDCR researchers can use to plan future studies or assess past work. The itemized and hierarchical structure of the checklist is meant to help researchers read and consider various parts as needed, and the ratable format makes it possible to use the checklist to score a FDCR study. Our ultimate hope is that this checklist will be used widely within the field, will foster homogeneity and transparency in FDCR research, and will facilitate data syntheses.

This is merely the first iteration of the checklist. Considering the rapid rate of progress in the field and based on feedback from the FDCR academic community, the checklist will be revised in later editions. Additionally, the checklist will be tested on a database of FDCR papers to ensure its validity. To ensure feasibility of the checklist application, we suggest considering the “items” (Supplementary Material 1) as a “must” in FDCR studies and “additional recommendations” as suggestions to improve the methodology design and report of FDCR studies (Supplementary Material 3). Finally, the adoption or neglect of the checklist by journal editors/reviewers and FDCR researchers around the world will determine its success or failure in the long term.

Funding Supports

Raymond F. Anton is supported by NIAAA P50 AA010761. Patrick Bach is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 402170461 – TRR 265 (Heinz et al., Addict Biol. 2019). Anne Beck is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 402170461 – TRR 265-project C02 (Heinz et al., Addict Biol. 2019). Kelly E. Courtney is supported by the California Tobacco-Related Disease Research Grant Program of the University of California grant number T30IP0962. Hamed Ekhtiari is supported by the Laureate Institute for Brain Research (LIBR), Warren K. Family Foundation, Oklahoma Center for Advancement of Science and Technologies (OCAST, #HR18-139) and Brain and Behavior Foundation (NARSAD Young Investigator Award #27305). Francesca M. Filbey is supported by the National Institute on Drug Abuse grants R01 DA030344 and R21DA044465. Rita Z. Goldstein is supported by NIDA grants R01DA041528, R01DA048301 and R01DA047851 and NCCIH grant R01AT010627. Erica Grodin is supported by the National Institute on Alcohol Abuse and Alcoholism grant F32AA027699. Andreas Heinz is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 402170461 – TRR 265 (Heinz et al., Addict Biol. 2019). Amy C. Janes is supported by R01DA039135 and K02DA042987. Marc J. Kaufman is supported with R01 DA041866. Hamid R Noori is supported by the Bundesministerium für Bildung und Forschung (e:Med program: FKZ: 01ZX1503 and 01ZX1909B), and the Deutsche Forschungsgemeinschaft (TRR 265-A05). Stéphane Potvin is holder of Eli Lilly Canada Chair on schizophrenia research. Joseph P. Schacht is supported by NIAAA grants R01 AA027765 and R01 AA026859. Dongju Seo is supported by R01AA026844 and K08AA023545. Vaughn R. Steele is partially funded by the National Institute on Drug Abuse (NIDA) grant K12 DA000167. Susan F. Tapert is supported by NIAAA grants U01 AA021692 and U24 AA021695, and NIDA grants U01 DA041089 and U24 DA041147. Sabine Vollstädt-Klein is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation): Project ID-402170461 (Heinz et al., Addict Biol. 2019), Project ID-421888313, Project-ID 437718741 and Project-ID 324164820. Reagan Wetherill is supported by K23AA023894, R01DA040670, and R21HL144673. Stephen J. Wilson is supported by R01DA041438 and R21DA045853. Katie Witkiewitz is supported by the National Institute on Alcohol Abuse and Alcoholism grants R01AA023665 and R01AA022328. Kai Yuan is supported by the National Natural Science Foundation of China (Grant No. 81871426). Anna Zilverstand is supported by NIDA grants P30 DA048742 and R01DA047851. The views presented in this manuscript represent those of the authors and not necessarily those of the funding agencies.

Declaration of Competing Interest

Raymond F. Anton is Chair of the American Society of Clinical Psychopharmacology's Alcohol Clinical Trials Initiative (ACTIVE Group), which has received support in the past 36 months by: Alkermes, Amygdala Neurosciences, Arbor Pharmaceuticals, Dicerna, Ethypharm, Indivior, Lundbeck, Mitsubishi, and Otsuka. He has been a recent consultant for Allergan, Alkermes, Dicerna, Insys, Laboratorio Pharmaceutico and Life Technologies, as well as receiving grant funding from Laboratorio Pharmaceutico. Amy C. Janes consults for Axial Biotherapeutics. Marc Potenza consulted for and advised the Addiction Policy Forum, Game Day Data, AXA, Idorsia and Opiant/Lakelight Therapeutics; received research support from the Mohegan Sun Casino and the National Center for Responsible Gaming (now the International Center for Responsible Gaming); consulted for legal and gambling entities on issues related to impulse-control and addictive disorders; given academic lectures in grand rounds, CME events, and other clinical/scientific venues; and generated books or chapters for publishers of mental health texts. Joseph P. Schacht has consulted for and received grant funding from Laboratorio Farmaceutico CT. Other authors declare no competing interests.

Tables and Figures

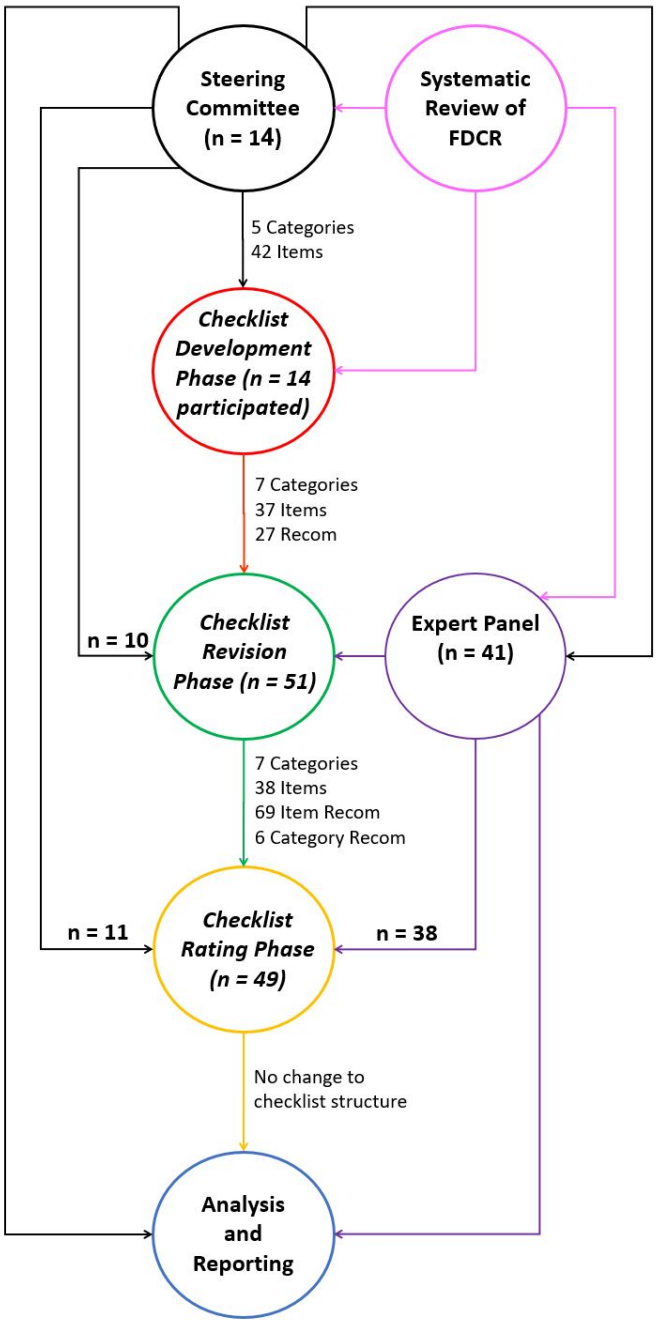

Figure 1: A schematic of the entire study methodology, The process has been roughly divided into distinct stages including the selection of the SC (in black), using the results of an earlier mentioned systematic review to choose the initial checklist items and expert committee candidates (in pink), checklist development phase (in red), expert panel selection (in purple), checklist commenting and revision phase (in green), checklist rating phase (in yellow), data analysis and Delphi process finalization (in blue). The number of contributors to each section is displayed by "n =". To the left of the main graph, an overview of the structure of the checklist at each stage is presented.

Table 1: Demographic and academic information for SC and EP members

	Steering Committee (n=14)	*Expert Panel (n=41)*
Gender		
Male	5	28
Female	9	13
Other	0	0
Age (years)		
Mean±SD	*51.1±9.1*	*45.3±9.4*
≤ 30	0	0
31–40	2	12
41–50	5	17
≥51	6	11
No response	1	1
Highest Academic Degree		
Bachelor of Science	0	0
Master of Science	0	0
Doctor of Medicine (MD)	0	4
Doctor of Philosophy (PhD)	11	33
Doctor of Medicine and Philosophy (MD, PhD)	3	4
Country of Residence		
Australia	0	1
Canada	0	1
China	0	3
Germany	1	9
Iran	0	1
Netherland	0	1
Sweden	0	2
United Kingdom	0	1
United States	13	22
Primary Field of Research		
Cognitive Science	1	1

Neuroscience	4	18
Psychiatry	6	14
Psychology	2	6
Statistics	0	1
Others	1	1
Primary Place of Work		
Business/Industry	0	0
Hospital	3	4
Independent Research Institute	2	4
University	8	33
Others	1	0
Time Spent in Addiction Research (Years)		
Mean±SD	*25.1±10.3*	*17±8.6*
≤5	0	3
6–10	0	7
11–20	8	21
≥21	4	10
No response	1	0
Time Spent in FDCR Research (Years)		
Mean±SD	*16.6±5.4*	*10.8±5.7*
≤5	0	9
6–10	2	12
11–20	9	18
≥21	2	1
No response	1	0

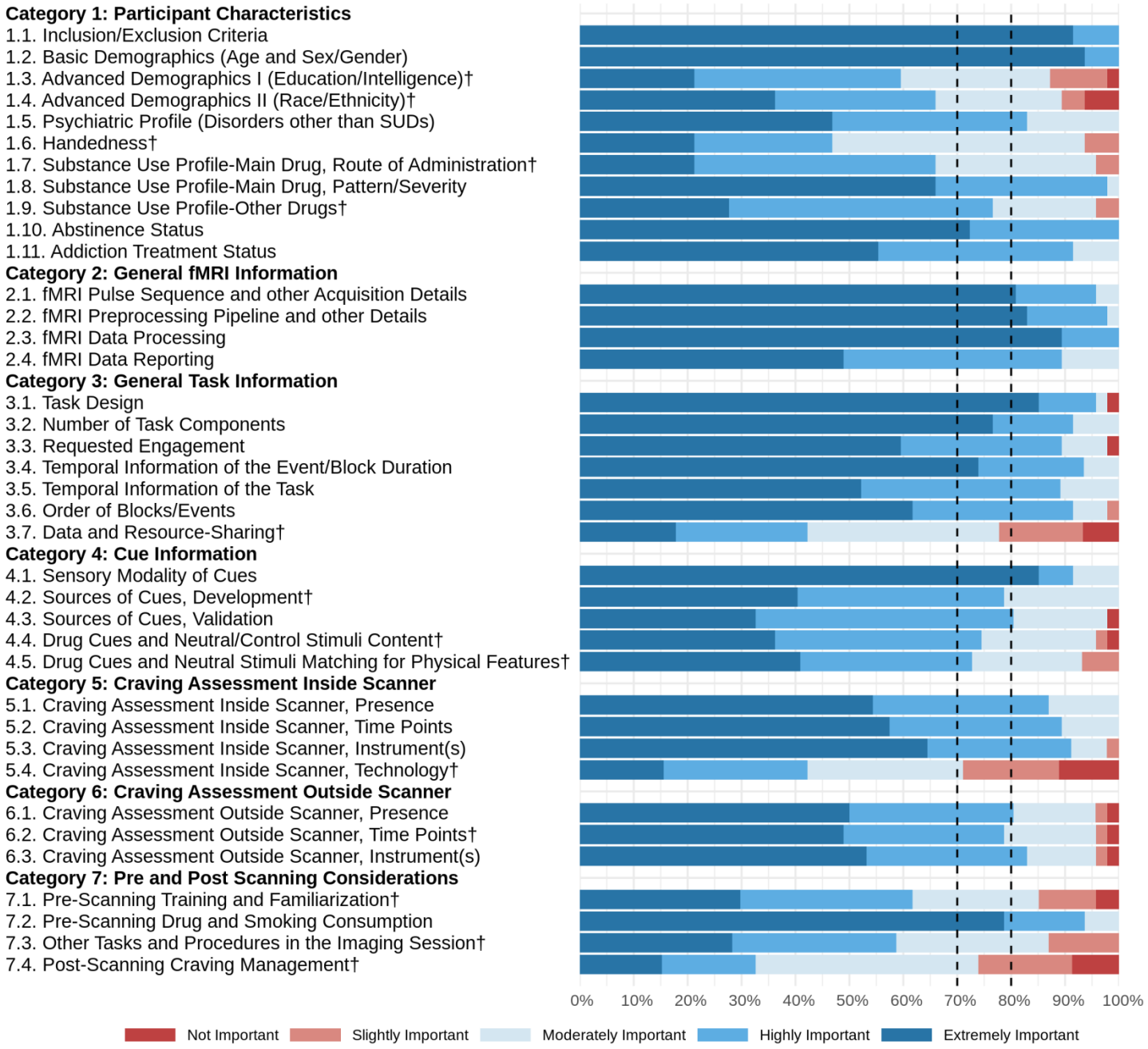

Figure 2: Ratings for 38 items in 7 categories. This figure depicts the rating of 49 raters (11 from the steering committee and 38 from the expert panel) for the checklist items. Each item was rated from 1-5 (not important- extremely important). All the items met threshold 1 and were rated as either moderately, highly, or extremely important by more than 70% of the raters. Also, 24 items reached the more stringent threshold 2 of being rated as either highly or extremely important by 80% raters (the ones which did not reach this threshold are marked with "†"). Items are represented by their summary in the figure. Full text of the items is provided in Tables 2-7.

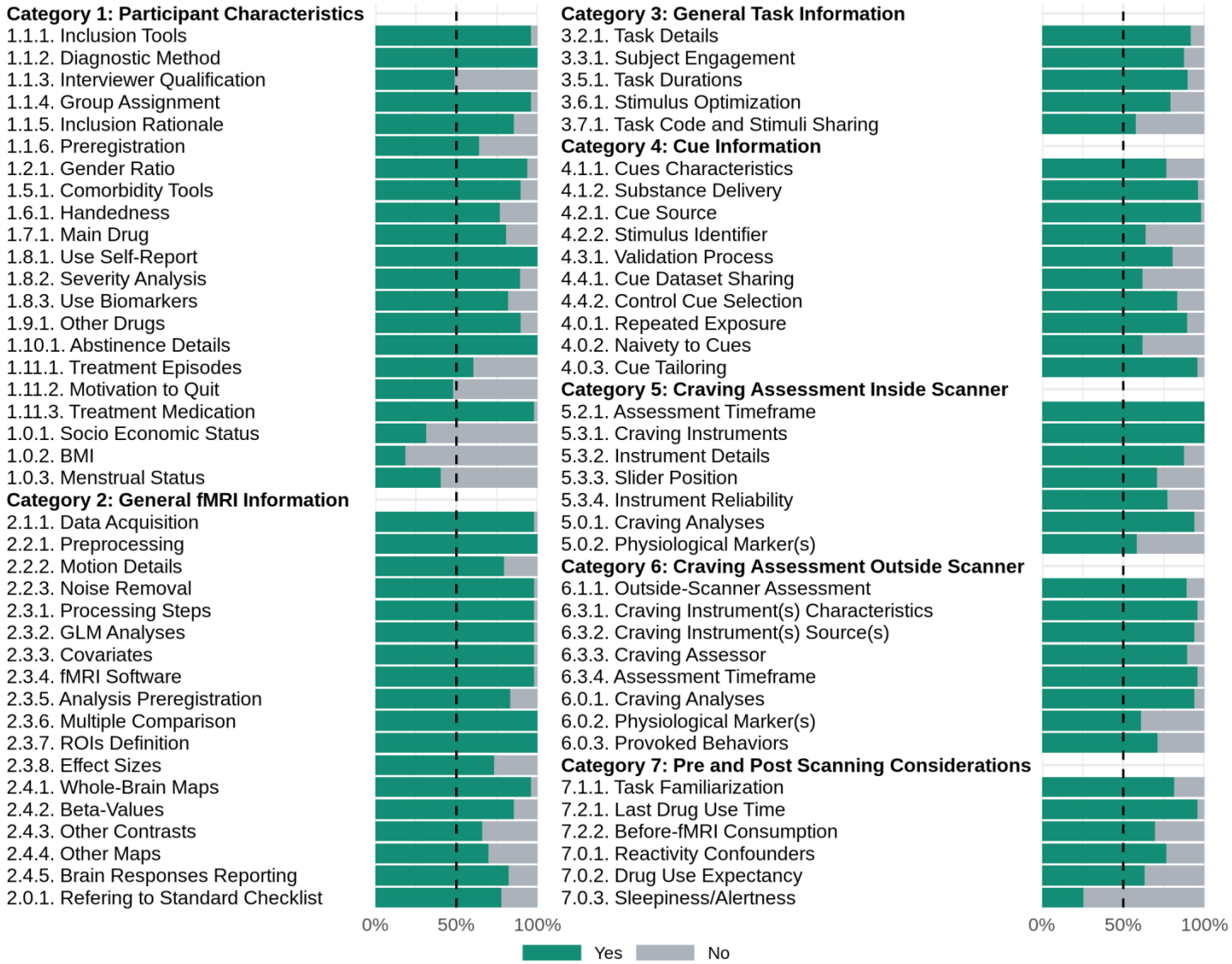

Figure 3: Ratings for 75 additional recommendations in 7 categories. This figure depicts the rating of 49 raters (11 from the steering committee and 38 from the expert panel) for the checklist additional recommendations. Each additional recommendation was rated either "Yes" or "No", on the question of whether it should be included as a recommendation. Recommendations are represented by their summary in the figure. Full text of the recommendations is provided in Tables 2-7.

Table 2. Items to report and recommendations in the participant characteristics category (category 1) of the checklist. Rating for items (1-5) are reported as mean (standard deviation) and rating for recommendations (Yes or No) are reported as frequency of Yes (percent of Yes reports).

Sub-Categories	Main Items to Report	Item Importance (1 to 5)	Specific Recommendation to Report	Recommendation Inclusion (Yes/No)
		Mean (SD)		Yes (Percent)
Inclusion/Exclusion Criteria	1.1. Inclusion and exclusion criteria for all participant groups	4.91 (0.28)	1.1.1. Include specific diagnostic criteria/measurement tools for conditions that were included and those that were excluded.	46 (96%)
			1.1.2. Clearly specify methods used to assess any diagnostic/dimensional criteria (e.g., SCID, MINI and their versions).	48 (100%)
			1.1.3. Report the qualification of the person who has applied these criteria (e.g. clinical psychologist, institute secretary, psychiatrist, etc.).	23 (49%)
			1.1.4. Report how participants were assigned to different groups in studies in which participants are assigned to more than one group.	46 (96%)
			1.1.5. Explain the rationale for criteria selected for recruitment (e.g., if only males are included).	41 (85%)
			1.1.6. Report whether methods for any additional subgroups and adjusted analyses were preregistered before or not (i.e., protocol paper, registration websites and etc.).	30 (64%)
Basic Demographics	1.2. Age and sex/gender for all participant groups	3.83 (0.82)	1.2.1. Report the number of male/female in the sample included in the reported analyses. There are studies which have reported the ratio in the recruited sample without reporting the ratio in the sample included in the analyses.	44 (94%)

Advanced Demographics I	1.3. Education or a measurement of intelligence for all participant groups	4.64 (0.53)		
Advanced Demographics II	1.4. Race or ethnicity for all participant groups	4 (0.81)		
Psychiatric Profile	1.5. Any categorical or dimensional measurement of psychopathologies other than substance use disorder	4.72 (0.45)	1.5.1. Report psychiatric comorbidities using diagnostic criteria (e.g., DSM) or questionnaires to assess the level of psychiatric comorbidities (for example, a quantitative assessment of depression or anxiety using various questionnaires).	42 (89%)
Handedness	1.6. Handedness for all participant groups	4.47 (0.65)	1.6.1. Use validated handedness inventories like the Edinburgh Handedness Inventory. The effect of handedness in the laterality of fMRI drug cue reactivity and its significance is still unclear. However, this effect can be explored with reproducible reporting of the handedness in the shared databases.	36 (77%)
Substance Use Profile-Main Drug	1.7. Route(s) of administration for the main substance (if it is obvious it doesn't need to be reported; i.e., there is only one route of administration for cigarette smokers or alcohol drinkers)	4.91 (0.28)	1.7.1. Report the breakdown of the main drug by type and route.	37 (80%)
Substance Use Profile-Main Drug	1.8. Current and lifetime use pattern/severity for the main drug of use for all participant groups	4.94 (0.25)	1.8.1. Report the exact measures and instruments used to assess current (e.g., last few days, last month, last 3 months or etc.) and life-time substance use (e.g., questions, questionnaires or lab tests).	46 (100%)
			1.8.2. Report whether/how derived variables from these severity measures have been used in fMRI drug cue reactivity analysis (whether they're used as	41 (89%)

			variables of interest or a regressed out variable, for example).	
			1.8.3. Include biological markers of drug use/severity (if available).	36 (82%)
Substance Use Profile-Other Drugs	1.9. Measures of current or lifetime use pattern/severity for drugs, other than the main drug of use, for all participant groups	4.3 (0.75)	1.9.1. Report the current and lifetime patterns and severity of use of other substances and potential use disorders.	43 (90%)
Abstinence Status	1.10. Days/hours/minutes since last use (duration of abstinence) and how abstinence was verified for all participant groups	3.62 (0.9)	1.10.1. Report a clear definition of abstinence, its assessment methods (e.g., timeline followback, urine toxicology, monitoring (i.e., breathalyzer or CO measures), clinical interviews, etc.) and the reference time point (i.e., recruitment or scanning).	47 (100%)
Addiction Treatment Status	1.11. Treatment status for all participant groups, (i.e., non-treatment seeking active users, treatment seeking active users, undergoing active treatment, treated and abstinent, relapsed after treatment, etc.)?	3.83 (0.82)	1.11.1. Specify the number and the nature of treatment episodes if participants have undergone multiple unsuccessful treatment episodes.	29 (60%)
			1.11.2. Report the level of motivation to discontinue substance use for active drug users.	22 (48%)
			1.11.3. Report whether they are on medication to treat their SUDs.	47 (98%)
General Recommendations			1.0.1. Probe and report a measure of income or sociodemographic status however the effect of this demographic dimension in fMRI drug cue reactivity is not explored yet.	15 (31%)
			1.0.2. Report BMI for all participant groups.	9 (19%)
			1.0.3. Report the menstrual status (e.g., days since the first day of last menstrual period (LMP) or menstrual phase/status) in female participants.	19 (40%)

Table 3. Items to report and recommendations in the general fMRI information category (category 2) of the checklist. Rating for items (1-5) are reported as mean (standard deviation) and rating for recommendations (Yes or No) are reported as frequency of Yes (percent of Yes reports).

Sub-Categories	Main Items to Report	Item Importance (1 to 5)	Specific Recommendation to Report	Recommendation Inclusion (Yes/No)
		Mean (SD)		Yes (Percent)
fMRI pulse sequence and other acquisition details	2.1. A section for fMRI data acquisition details	4.77 (0.52)	2.1.1. Report fMRI data acquisition details based on the available checklists (e.g., COBIDAS or Poldrack et.al., 2008, Neuroimage). FMRI data acquisition details might have explicit effects on drug cue reactivity results, e.g., number of head coil channels, as higher channels (32 compared to 8) might be associated with better SNR in the cortex with the cost of losing signal in the deep parts of the brain.	47 (98%)
fMRI preprocessing pipeline and other details	2.2. A section for fMRI preprocessing details	4.81 (0.45)	2.2.1. Report fMRI preprocessing details based on the available checklists (e.g., COBIDAS or Poldrack et.al., 2008, Neuroimage). There are items in the preprocessing steps that might have effects on fMRI drug cue reactivity results. For example, higher FWHM might be related to the loss of signal in small nuclei.	48 (100%)
			2.2.2. Report motion differences between participant groups (i.e., individuals with a SUD vs. controls) as higher motion during the drug related blocks compared to neutral blocks might act as a confounder.	38 (79%)
			2.2.3. Report quality control measures, artefact detection methods and the threshold to exclude participants with heavy movement.	47 (98%)
fMRI Data Processing	2.3. A section for MRI analyses and statistical modeling details	4.89 (0.31)	2.3.1. Report fMRI single subject level and group level processing steps based on the standard checklists (e.g., COBIDAS or Poldrack et.al., 2008, Neuroimage).	47 (98%)
			2.3.2. Report whether GLM analyses are random, mixed, or fixed effects for inclusion in future meta analyses.	47 (98%)

			2.3.3. Report all covariates used for each model and whether or not demeaning was done for covariates of interest.	47 (98%)
			2.3.4. Report any publicly available tool/software use (e.g., SPM, AFNI, FSL, etc.).	47 (98%)
			2.3.5. Report any attempt for preregistration of data processing methods.	39 (83%)
			2.3.6. Report methods that are used to control for multiple comparisons error and spatial autocorrelations.	47 (100%)
			2.3.7. Report the definition of the ROIs for studies using an ROI approach.	47 (100%)
			2.3.8. Provide effect sizes for all reported statistics.	33 (73%)
fMRI Data Reporting	2.4. A map of the basic whole brain response to drug cues	4.38 (0.68)	2.4.1. Report the second-level maps or activation foci therein of each study group singly, as well as group-difference map (e.g. between clinical group and control group) (if applicable) in the results or the supplements as a figure or table (foci coordinates and stats) with details on the thresholding measures and quantities. Even if the paper has other analyses (e.g. task-based connectivity), the whole brain maps of the craving>neutral contrast should be reported for comparison with other studies and future meta analyses.	45 (96%)
			2.4.2. Report beta-values for both conditions (craving and neutral) as an "activation" in the mPFC during craving could be explained by a deactivation in the control condition.	41 (85%)
			2.4.3. Report the contrast map for other included conditions ((e.g. multiple drug stimuli, affective images, other active control) If other conditions are included	31 (66%)
			2.4.4. Provide effect size map, non-thresholded statistical map, and the data in an accessible repository (e.g., OSF, NIMH/NIAAA data archive, GitHub, Neurovault, etc.).	32 (70%)

	2.4.5. It is understandable that researchers who are not using conventional whole brain GLM based methods (i.e., ICA, Graph Theory, PPI connectivity, ROI only analysis and etc.) or developing other innovative and non-conventional methods might face difficulties to report "whole brain response to drug cues". It is still recommended for these studies to consider strategies for reporting brain response to drug cues to make data/results aggregation and comparison possible.	37 (82%)
General Recommendations	2.0.1. Refer to standard checklists (e. g., COBIDAS) for items in this category. Items in the ENIGMA ACRI checklist are designed to be dichotomous (Yes or No), however, there is a continuum for the details to be reported. Provide as much detail as available.	28 (78%)

Table 4. Items to report and recommendations in the general task information category (category 3) of the checklist. Rating for items (1-5) are reported as mean (standard deviation) and rating for recommendations (Yes or No) are reported as frequency of Yes (percent of Yes reports).

Sub-Categories	Main Items to Report	Item Importance (1 to 5)	Specific Recommendation to Report	Recommendation Inclusion (Yes/No)
		Mean (SD)		Yes (Percent)
Task Design	3.1. General task structure (Event-related, Block or Mixed (events in blocks))	4.77 (0.7)		
Number of Task Components	3.2. Number of runs (if more than one), blocks (for block-designed studies), and events (including drug cues, neutral/control cues, fixations, etc.)	4.68 (0.63)	3.2.1. Explicitly define terms such as "block", "event", "session", "run" etc., with reference to standard checklists (e.g., COBIDAS) given the ambiguity surrounding these terms.	44 (92%)
Requested Engagement	3.3. Instructions to the study participants on how to engage with the cues	4.45 (0.83)	3.3.1. Report the details of the given instructions on how to engage (interact) with cues and provide the exact text of the instruction. The interactions may be passive viewing (if there was explicitly no instruction or if they were asked to do nothing), free craving, attentive viewing, rating or classifying each cue, spatial cueing, inhibiting craving, etc.	42 (88%)
Temporal information of the event/block duration	3.4. Duration of each cue (for both event and blocked-design tasks) and the total block duration (for blocked-design tasks)	4.67 (0.6)		
Temporal Information of the Task	3.5. Total task duration	4.41 (0.69)	3.5.1. Report the duration of all sections of the task between the cues/events/blocks and within them.	43 (90%)

Order of Blocks/Events	3.6. Order of block types (e.g., drug, control) (for blocked-designs) or event types (e.g., drug, control) (for event-related designs) (The order can be fully randomized (randomized and different between subjects), pseudorandomized (identical between subjects, but randomized once for the order of events/blocks), or not randomized (fixed order like neutral-drug-neutral-drug for all subjects)	4.51 (0.72)	3.6.1. Report if the stimulus presentation was optimized using any software (e.g. genetic algorithm, optseq).	38 (79%)
Data and resource-sharing	3.7. Sharing policy for the behavioral task code or source images	3.31 (1.14)	3.7.1. Provide the task code and the code used for generating these sequences (i.e. GitHub or OSF platforms).	26 (58%)

Table 5. Items to report and recommendations in the cue information category (category 4) of the checklist. Rating for items (1-5) are reported as mean (standard deviation) and rating for recommendations (Yes or No) are reported as frequency of Yes (percent of Yes reports).

Sub-Categories	Main Items to Report	Item Importance (1 to 5)	Specific Recommendation to Report	Recommendation Inclusion (Yes/No)
		Mean (SD)		Yes (Percent)
Sensory Modality of Cues	4.1. Modality(ies) of utilized drug and neutral/control cues (The modalities can be word, picture, smell, taste, tactile, audio script, written script, imagination, silent video, audiovisual video, paraphernalia, substance itself, or mixed.)	4.77 (0.6)	4.1.1. Provide an overview of the range of values for important characteristics of chosen cues. In the case of visual cues, this could be in the form of describing the complexity, luminance, and hue of cues. For auditory cues this could consist of describing the volume and frequency, and for scripts it could be font and typeface.	36 (77%)
			4.1.2. Report the amount of the substance and its method of delivery (i.e., oral, IV), If the substance itself is administered as a cue (e.g. very small amounts of alcohol or cigarette smoke).	46 (96%)
Sources of Cues, Development	4.2. Source of drug and neutral/control cues	4.09 (0.84)	4.2.1. Report the exact source of acquiring the cues If the cues are newly developed, or cite the relevant references if they are from other already developed sources. If the stimulus set is newly developed, criteria used for stimulus selection should be specified (e.g., exclusion of people in images, paraphernalia only). If a subset of developed sources was used, indicate what criteria were used	47 (98%)

			for selecting this subset (could be a random selection).	
			4.2.2. List the stimulus identifiers in the appendix or supplementary material of the paper, if the cue sources include stimulus identifiers	30 (64%)
Sources of Cues, Validation	4.3. Extent of prior validation of drug and neutral/control cues used in the study (Drug and neutral/control cues in a study might be not validated, validated by assessing the craving induction of each cue individually using simple-item craving instruments like single-item VAS, or using standardized instruments of craving assessment and emotion or stress reactivity)	4.07 (0.95)	4.3.1. Provide the details of the validation process. Even if the validation has been done in another study, the validation study should be cited and then the validation process of the cues should be briefly introduced as well.	37 (80%)
Drug and Neutral/Control Cue Content	4.4. Content of drug cues and its relationship to the targeted drug (These include stimulus related to the drug, stimulus related to instruments of drug use, stimulus related to various stages of drug use (i.e. "beginning" or "end" stimuli (lit cigarette vs. ashtray)), stimulus related to drug intake, stimulus related to typical drug consumption environments, stimulus related to preparation	4.07 (0.93)	4.4.1. Explicitly report if they are willing to share their drug and neutral/control cue database/task in the published paper. Providing a reliable link (like GitHub or other open science repositories) to a shared database inside the paper is the ideal scenario, meanwhile, facing copyright concerns for the drug cues collected from the web or other copyright protected resources might limit this potential. All too often, links are provided in papers that are broken a few years after publication.	29 (62%)

	of drug, stimulus related to purchasing the drug, etc.)		4.4.2. Explain the nature of neutral/control cues and why they were chosen, as they might belong to several types in terms of their content.	40 (83%)
Drug to Neutral/Control Cues Matching for Physical Features	4.5. Factors for which drug and neutral/control cues have been matched (color, brightness, hue, content, complexity, scrambled drug cue, etc.)	4.09 (0.84)		
General Recommendations			4.0.1. Report the characteristics of the cue sets used when a task is repeated, if a study involves a longitudinal design.	42 (89%)
			4.0.2. Control and report being naïve to drug cue exposure or previous experiences of cue exposure before the target study. Recent evidence shows participants will respond differently to drug cues in the second exposure. However asking people to report cue exposure outside of the target study might be complex.	29 (62%)
			4.0.3. Report whether and how drug and neutral/control cues were tailored for each participant. Drug and neutral/control cue tailoring could involve asking participants to choose cues from a cue database or developing participant-specific cues based on consultation with individual participants. Details of the individualization protocol should be provided.	45 (96%)

Table 6. Items to report and recommendations in the craving assessment categories (categories 5 & 6) of the checklist. Rating for items (1-5) are reported as mean (standard deviation) and rating for recommendations (Yes or No) are reported as frequency of Yes (percent of Yes reports).

Sub-Categories	Main Items to Report	Item Importance (1 to 5)	Specific Recommendation to Report	Recommendation Inclusion (Yes/No)
		Mean (SD)		Yes (Percent)
Craving Assessment inside Scanner, Presence	5.1. Presence of any craving assessment inside the scanner	4.41 (0.72)		
Craving Assessment inside Scanner, Time Points	5.2. Time points at which craving-related assessment is performed inside the scanner (e.g. before and/or after each cue/event/block/trial/scan/run/session) (Yes/No/Not Applicable [in case when there is no assessment inside the scanner])	4.47 (0.69)	5.2.1. Report the timeframe of craving assessment (i.e., now (after cue presentation) or during cue presentation).	48 (100%)
Craving Assessment Inside Scanner, Instrument(s)	5.3. Description of the instrument(s) used to assess craving or craving-related constructs inside the scanner (Yes/No/Not Applicable [in case when there is no assessment inside the scanner])	4.53 (0.73)	5.3.1. Report the exact characteristics of the instrument(s) used to assess craving and craving constructs (i.e., urge, desire, interest, like vs. want, and etc.) inside the scanner, including number of items, range of possible responses, whether it was VAS or Likert, internal consistency and whether any transformations were applied to the instrument and its scores prior to the data collection and analysis.	48 (100%)

			5.3.2. Cite any relevant sources of instruments, and whenever possible provide the exact wording of the craving question/s.	42 (88%)
			5.3.3. Provide information on the start position of the slider, when using VAS or other continuous scales with a slider (e.g. in the middle or lateral ends of the scales).	34 (71%)
			5.3.4. Report information on the reliability of the instrument if the instrument(s) administered repeatedly before/during/after scanning.	37 (77%)
Craving Assessment Inside Scanner, Technology	5.4. Description of the hardware used to obtain participant responses, with specifications of models and brands of devices, if necessary (e.g., response box, fiber-optic pad) (Yes/No/Not Applicable [in case when there is no assessment inside the scanner])	4.53 (0.73)		
General Recommendations			5.0.1. Report analyses related to the craving measurements, i.e., whether they differed between the main group and control/s or from pre to post scan.	45 (94%)
			5.0.2. Probe and report physiological correlates of craving (i.e., skin conductance, heart rate, temperature, respiration and	28 (58%)

			blood volume pulse amplitude) before/during/after cue presentation.	
Craving Assessment Outside Scanner, Presence	6.1. Presence of any craving-related assessment outside the scanner	4.24 (0.95)	6.1.1. Probe and report craving assessment outside or inside the scanner in FDCR tasks. The assessment of cue-induced craving is of great relevance to the validity of the FDCR task. Thus, the authors should at least clarify whether they have considered including a craving assessment inside/outside the scanner, even if they have finally decided not to report the results.	40 (89%)
Craving Assessment Outside Scanner, Time Points	6.2. Time points at which craving-related assessment is performed outside the scanner (e.g., immediate before entering the scanner, immediately after exiting the scanner, etc.) (Yes/No/Not Applicable [in case when there is no assessment outside the scanner])	4.21 (0.95)		
Craving Assessment Outside Scanner, Instrument(s)	6.3. Description of the instrument(s) used to assess craving or craving-related constructs outside the scanner (Yes/No/Not Applicable [in case when there is no assessment outside the scanner])	4.3 (0.93)	6.3.1. Report the exact characteristics of the instrument(s) to assess craving and craving constructs (i.e., urge, desire, interest, like vs. want, etc.) outside the scanner, including number of items, range of responses, internal consistency, and whether it was VAS or Likert, and whether any	45 (96%)

	transformations were applied to the instrument and its scores prior to the data collection and analysis.	
	6.3.2. Cite any relevant sources of instruments, and whenever possible provide exact wording of the craving question/s.	44 (94%)
	6.3.3. Report that the instrument is self-assessed or experimenter assessed.	42 (89%)
	6.3.4. Report the timeframe of craving assessment (i.e., now or during the scan).	45 (96%)
General Recommendations	6.0.1. Report analyses related to the craving measurements, i.e., whether they differed between the main group and control/s or from pre to post scan.	44 (94%)
	6.0.2. Probe and report physiological correlates of craving (i.e., skin conductance, heart rate, temperature, respiration and blood volume pulse amplitude) before/during/after cue presentation.	28 (61%)
	6.0.3. Probe cue-provoked behaviors (e.g., drug seeking or using behaviors) after scanning, whenever possible.	32 (71%)

Table 7. Items to report and recommendations in the pre- and post-scanning considerations category (category 7) of the checklist. Rating for items (1-5) are reported as mean (standard deviation) and rating for recommendations (Yes or No) are reported as frequency of Yes (percent of Yes reports).

Sub-Categories	Main Items to Report	Item Importance (1 to 5)	Specific Recommendation to Report	Recommendation Inclusion (Yes/No)
		Mean (SD)		Yes (Percent)
Pre-scanning Training and Familiarization	7.1. Procedure to train/familiarize participants with the task/scanner before the scanning	3.72 (1.14)	7.1.1. Report both task training and scanner familiarization procedures before scanning. Familiarization can be done using variety methods e.g., describing the situation for the participants, entering the subjects to mock scanners. Training can be done by letting the participants do the task outside the scanner.	39 (81%)
Pre-scanning Drug and Smoking Consumption	7.2. Whether participants were allowed to smoke or use other drugs prior to scanning	4.72 (0.58)	7.2.1. Report the time interval between the last use of nicotine and other drugs and scanning.	45 (96%)
			7.2.2. Consider and report the consumption of caffeine, prescribed medications or food eaten based on the context of the study (e.g., controlling the time and the quantity of consumption). For instance, for the assessment of alcohol craving it is essential to control for liquid intake prior scanning, since drinking high amounts of water can blunt alcohol craving.	32 (70%)
Other Tasks and Procedures in the Imaging Session	7.3. Presence and order of other tasks and procedures (e.g. resting fMRI or DTI before drug cue reactivity, etc.) in the imaging session	3.13 (1.15)		

Post-scanning Craving Management	7.4. Steps taken to reduce participant craving after performing the task	3.72 (1.14)		
General Recommendations			7.0.1. Report the elements that might change the fMRI drug cue reactivity as a potentially partially state-dependent data, across the study days (i.e., time of scanning during the day considering the diurnal variation in responding to cues) or between studies (i.e. sequence of imaging tasks/protocols) to make sure that the result is representing a data unconfounded by procedural differences.	36 (77%)
			7.0.2. Explicitly report the participants' drug use expectancy, even though the potentials for having access or expectation of drug use after the cue exposure process is usually implicit in the study setting and inclusion/exclusion criteria. It has been shown that the participants' expectancy for drug use might influence cue reactivity. Participants who expect that they might have access to drugs after cue exposure will respond differently to cues compared to those who are sure that there is no access to drugs after cue exposure (e.g. being back in an in-patient or residential setting). Also explicitly discuss how they considered the influence of expectancy, and whether they attempted to modulate or control for it in the study.	29 (63%)
			7.0.3. Report a measure of sleepiness or alertness before fMRI drug cue reactivity.	12 (26%)

References

Albrecht, J., Burke, M., Haegler, K., Schopf, V., Kleemann, A. M., Paolini, M., . . . Linn, J. (2010). Potential impact of a 32-channel receiving head coil technology on the results of a functional MRI paradigm. *Clin Neuroradiol, 20*(4), 223-229. doi:10.1007/s00062-010-0029-2

Bach, P., Kirsch, M., Hoffmann, S., Jorde, A., Mann, K., Frank, J., . . . Vollstädt-Klein, S. (2015). The effects of single nucleotide polymorphisms in glutamatergic neurotransmission genes on neural response to alcohol cues and craving. *Addict Biol, 20*(6), 1022-1032. doi:10.1111/adb.12291

Bailey, S. R., Goedeker, K. C., & Tiffany, S. T. (2010). The impact of cigarette deprivation and cigarette availability on cue-reactivity in smokers. *Addiction, 105*(2), 364-372. doi:10.1111/j.1360-0443.2009.02760.x

Billieux, J., Khazaal, Y., Oliveira, S., de Timary, P., Edel, Y., Zebouni, F., . . . Van der Linden, M. (2011). The Geneva Appetitive Alcohol Pictures (GAAP): development and preliminary validation. *Eur Addict Res, 17*(5), 225-230. doi:10.1159/000328046

Blaine, S., Claus, E., Harlaar, N., & Hutchison, K. (2013). TACR1 genotypes predict fMRI response to alcohol cues and level of alcohol dependence. *Alcohol Clin Exp Res, 37 Suppl 1*, E125-130. doi:10.1111/j.1530-0277.2012.01923.x

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L., . . . Group, S. (2015). STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ, 351*, h5527. doi:10.1136/bmj.h5527

Carmichael, O., Schwarz, A. J., Chatham, C. H., Scott, D., Turner, J. A., Upadhyay, J., . . . Hawkins, K. R. (2018). The role of fMRI in drug development. *Drug Discov Today, 23*(2), 333-348. doi:10.1016/j.drudis.2017.11.012

Carter, B. L., & Tiffany, S. T. (1999). Meta-analysis of cue-reactivity in addiction research. *Addiction, 94*(3), 327-340.

Casey, B., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., . . . Garavan, H. (2018). The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. *Developmental cognitive neuroscience, 32*, 43-54.

Casey, B. J., Jones, R. M., & Hare, T. A. (2008). The adolescent brain. *Annals of the New York Academy of Sciences, 1124*, 111-126. doi:10.1196/annals.1440.010

Chang, C., Cunningham, J. P., & Glover, G. H. (2009). Influence of heart rate on the BOLD signal: the cardiac response function. *Neuroimage, 44*(3), 857-869.

Chase, H. W., Eickhoff, S. B., Laird, A. R., & Hogarth, L. (2011). The neural basis of drug stimulus processing and craving: an activation likelihood estimation meta-analysis. *Biological psychiatry, 70*(8), 785-793. doi:10.1016/j.biopsych.2011.05.025

Chen, J., Hutchison, K. E., Calhoun, V. D., Claus, E. D., Turner, J. A., Sui, J., & Liu, J. (2015). CREB-BDNF pathway influences alcohol cue-elicited activation in drinkers. *Human brain mapping, 36*(8), 3007-3019. doi:10.1002/hbm.22824

Cheng, G. L. F., Zeng, H., Leung, M. K., Zhang, H. J., Lau, B. W. M., Liu, Y. P., . . . Lee, T. M. C. (2013). Heroin abuse accelerates biological aging: a novel insight from telomerase and brain imaging interaction. *Translational Psychiatry, 3*(5), e260-e260. doi:10.1038/tp.2013.36

Chipchase, L., Schabrun, S., Cohen, L., Hodges, P., Ridding, M., Rothwell, J., . . . Ziemann, U. (2012). A checklist for assessing the methodological quality of studies using transcranial magnetic stimulation to study the motor system: an international consensus study. *Clin Neurophysiol, 123*(9), 1698-1704. doi:10.1016/j.clinph.2012.05.003

Chua, H. F., Ho, S. S., Jasinska, A. J., Polk, T. A., Welsh, R. C., Liberzon, I., & Strecher, V. J. (2011). Self-related neural response to tailored smoking-cessation messages predicts quitting. *Nat Neurosci, 14*(4), 426-427. doi:10.1038/nn.2761

Claus, E. D., Blaine, S. K., Filbey, F. M., Mayer, A. R., & Hutchison, K. E. (2013). Association Between Nicotine Dependence Severity, BOLD Response to Smoking Cues, and Functional Connectivity. *Neuropsychopharmacology, 38*(12), 2363-2372. doi:10.1038/npp.2013.134

Clayton, R. B., Bailey, R. L., & Liu, J. (2019). Conditioned "Cross Fading": The Incentive Motivational Effects of Mediated-Polysubstance Pairings on Alcohol, Marijuana, and Junk Food Craving. *Journal of Health Communication, 24*(3), 319-327. doi:10.1080/10810730.2019.1601304

Coffey, S. F., Schumacher, J. A., Stasiewicz, P. R., Henslee, A. M., Baillie, L. E., & Landy, N. (2010). Craving and physiological reactivity to trauma and alcohol cues in posttraumatic stress disorder and alcohol dependence. *Experimental and clinical psychopharmacology, 18*(4), 340-349. doi:10.1037/a0019790

Colizoli, O., De Gee, J. W., Van Der Zwaag, W., & Donner, T. H. (2020). Comparing fMRI responses measured at 3 versus 7 Tesla across human cortex, striatum, and brainstem. *BioRxiv*.

Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. (2015). Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) The TRIPOD Statement. *Circulation, 131*(2), 211-219.

Cuzzocreo, J. L., Yassa, M. A., Verduzco, G., Honeycutt, N. A., Scott, D. J., & Bassett, S. S. (2009). Effect of handedness on fMRI activation in the medial temporal lobe during an auditory verbal memory task. *Human brain mapping, 30*(4), 1271-1278. doi:10.1002/hbm.20596

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human brain mapping, 8*(2-3), 109-114.

Degenhardt, L., Charlson, F., Ferrari, A., Santomauro, D., Erskine, H., Mantilla-Herrara, A., . . . Griswold, M. (2018). The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Psychiatry, 5*(12), 987-1012.

Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M., & Wales, P. W. (2014). Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol, 67*(4), 401-409. doi:10.1016/j.jclinepi.2013.12.002

Dong, G., Wang, L., Du, X., & Potenza, M. N. (2018). Gender-related differences in neural responses to gaming cues before and after gaming: implications for gender-specific vulnerabilities to Internet gaming disorder. *Social cognitive and affective neuroscience, 13*(11), 1203-1214.

Dong, G., Wang, Z., Wang, Y., Du, X., & Potenza, M. N. (2019). Gender-related functional connectivity and craving during gaming and immediate abstinence during a mandatory break: Implications for development and progression of internet gaming disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 88*, 1-10.

Dong, G., Zheng, H., Liu, X., Wang, Y., Du, X., & Potenza, M. N. (2018). Gender-related differences in cue-elicited cravings in Internet gaming disorder: The effects of deprivation. *Journal of behavioral addictions, 7*(4), 953-964.

Ekhtiari, H. (2020a). Methodological Checklist for fMRI Drug Cue Reactivity Studies: Development and Consensus. Retrieved from https://osf.io/gwrh6/?view_only=2a0005afd5da470ebe30ef79fcc8e62f

Ekhtiari, H. (2020b). A Systematic Review on Methodological Dimensions of fMRI Drug Cue Reactivity. Retrieved from https://osf.io/eb972/?view_only=856eb3533fc64096bbcfaff6cbf99aee

Ekhtiari, H., Faghiri, A., Oghabian, M.-A., & Paulus, M. P. (2016). Functional neuroimaging for addiction medicine: From mechanisms to practical considerations *Progress in brain research* (Vol. 224, pp. 129-153): Elsevier.

Ekhtiari, H., Kuplicki, R., Aupperle, R., & Paulus, M. P. (2020). It is Never as Good the Second Time Around: Brain Areas Involved in Salience Processing Habituate During Repeated Drug Cue Exposure in Methamphetamine and Opioid Users. *bioRxiv*, 2020.2004.2018.036368. doi:10.1101/2020.04.18.036368

Ekhtiari, H., Kuplicki, R., Aupperle, R., & Paulus, M. P. (2020). It is Never as Good the Second Time Around: Brain Areas Involved in Salience Processing Habituate During Repeated Drug Cue Exposure in Methamphetamine and Opioid Users. *BioRxiv*.

Ekhtiari, H., Kuplicki, R., Pruthi, A., & Paulus, M. (2020). Methamphetamine and Opioid Cue Database (MOCD): Development and Validation. *Drug Alcohol Depend, 209*, 107941. doi:10.1016/j.drugalcdep.2020.107941

Ekhtiari, H., Nasseri, P., Yavari, F., Mokri, A., & Monterosso, J. (2016). Neuroscience of drug craving for addiction medicine: From circuits to therapies *Progress in brain research* (Vol. 223, pp. 115-141): Elsevier.

Engelmann, J. M., Versace, F., Robinson, J. D., Minnix, J. A., Lam, C. Y., Cui, Y., . . . Cinciripini, P. M. (2012). Neural substrates of smoking cue reactivity: a meta-analysis of fMRI studies. *Neuroimage, 60*(1), 252-262. doi:10.1016/j.neuroimage.2011.12.024

Esteban, O., Ciric, R., Finc, K., Blair, R. W., Markiewicz, C. J., Moodie, C. A., . . . Gorgolewski, K. J. (2020). Analysis of task-based functional MRI data preprocessed with fMRIPrep. *Nat Protoc, 15*(7), 2186-2202. doi:10.1038/s41596-020-0327-3

Filbey, F. M., Ray, L., Smolen, A., Claus, E. D., Audette, A., & Hutchison, K. E. (2008). Differential neural response to alcohol priming and alcohol taste cues is associated with DRD4 VNTR and OPRM1 genotypes. *Alcohol Clin Exp Res, 32*(7), 1113-1123. doi:10.1111/j.1530-0277.2008.00692.x

Filbey, F. M., Schacht, J. P., Myers, U. S., Chavez, R. S., & Hutchison, K. E. (2009). Marijuana craving in the brain. *Proceedings of the National Academy of Sciences, 106*(31), 13016. doi:10.1073/pnas.0903863106

Fischer, H., Wright, C. I., Whalen, P. J., McInerney, S. C., Shin, L. M., & Rauch, S. L. (2003). Brain habituation during repeated exposure to fearful and neutral faces: a functional MRI study. *Brain research bulletin, 59*(5), 387-392.

Forsman, A. K., Wahlbeck, K., Aaro, L. E., Alonso, J., Barry, M. M., Brunn, M., . . . Consortium, R. (2015). Research priorities for public mental health in Europe: recommendations of the ROAMER project. *Eur J Public Health, 25*(2), 249-254. doi:10.1093/eurpub/cku232

Franken, I. H., Hendriks, V. M., & van den Brink, W. (2002). Initial validation of two opiate craving questionnaires: the Obsessive Compulsive Drug Use Scale and the Desires for Drug Questionnaire. *Addict Behav, 27*(5), 675-685.

Franklin, T., Wang, Z., Suh, J. J., Hazan, R., Cruz, J., Li, Y., . . . Childress, A. R. (2011). Effects of varenicline on smoking cue-triggered neural and craving responses. *Arch Gen Psychiatry, 68*(5), 516-526. doi:10.1001/archgenpsychiatry.2010.190

Franklin, T. R., Jagannathan, K., Ketcherside, A., Spilka, N., & Wetherill, R. R. (2019). Menstrual cycle phase modulates responses to smoking cues in the putamen: Preliminary evidence for a novel target. *Drug and Alcohol Dependence, 198*, 100-104. doi:https://doi.org/10.1016/j.drugalcdep.2019.01.039

Gagnier, J. J., Kienle, G., Altman, D. G., Moher, D., Sox, H., Riley, D., & Group, C. (2013). The CARE guidelines: consensus-based clinical case reporting guideline development. *Journal of medical case reports, 7*(1), 223.

Gloria, R., Angelos, L., Schaefer, H. S., Davis, J. M., Majeskie, M., Richmond, B. S., . . . Baker, T. B. (2009). An fMRI investigation of the impact of withdrawal on regional brain activity during nicotine anticipation. *Psychophysiology, 46*(4), 681-693. doi:10.1111/j.1469-8986.2009.00823.x

Goldstein, R. Z., Tomasi, D., Alia-Klein, N., Honorio Carrillo, J., Maloney, T., Woicik, P. A., . . . Volkow, N. D. (2009). Dopaminergic Response to Drug Words in Cocaine Addiction. *The Journal of Neuroscience, 29*(18), 6001-6006. doi:10.1523/jneurosci.4247-08.2009

Goudriaan, A. E., Veltman, D. J., van den Brink, W., Dom, G., & Schmaal, L. (2013). Neurophysiological effects of modafinil on cue-exposure in cocaine dependence: A randomized placebo-controlled cross-over study using pharmacological fMRI. *Addictive Behaviors, 38*(2), 1509-1517. doi:https://doi.org/10.1016/j.addbeh.2012.04.006

Gu, X., Lohrenz, T., Salas, R., Baldwin, P. R., Soltani, A., Kirk, U., . . . Montague, P. R. (2016). Belief about Nicotine Modulates Subjective Craving and Insula Activity in Deprived Smokers. *Front Psychiatry, 7*, 126. doi:10.3389/fpsyt.2016.00126

Gusnard, D. A., Raichle, M. E., & Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nat Rev Neurosci, 2*(10), 685-694. doi:10.1038/35094500

Hanlon, C. A., Dowdle, L. T., Gibson, N. B., Li, X., Hamilton, S., Canterberry, M., & Hoffman, M. (2018). Cortical substrates of cue-reactivity in multiple substance dependent populations: transdiagnostic relevance of the medial prefrontal cortex. *Transl Psychiatry, 8*(1), 186. doi:10.1038/s41398-018-0220-9

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *J Adv Nurs, 32*(4), 1008-1015.

Heishman, S. J., Lee, D. C., Taylor, R. C., & Singleton, E. G. (2010). Prolonged duration of craving, mood, and autonomic responses elicited by cues and imagery in smokers: Effects of tobacco deprivation and sex. *Experimental and clinical psychopharmacology, 18*(3), 245.

Heishman, S. J., Saha, S., & Singleton, E. G. (2004). Imagery-induced tobacco craving: duration and lack of assessment reactivity bias. *Psychology of Addictive Behaviors, 18*(3), 284.

Heishman, S. J., Singleton, E. G., & Liguori, A. (2001). Marijuana Craving Questionnaire: Development and initial validation of a self-report instrument. *Addiction, 96*(7), 1023-1034.

Hsu, C.-C., & Sandford, B. A. (2007). The Delphi technique: making sense of consensus. *Practical Assessment, Research, and Evaluation, 12*(1), 10.

Janes, A. C., Frederick, B., Richardt, S., Burbridge, C., Merlo-Pich, E., Renshaw, P. F., . . . Kaufman, M. J. (2009). Brain fMRI reactivity to smoking-related images before and during extended smoking abstinence. *Exp Clin Psychopharmacol, 17*(6), 365-373. doi:10.1037/a0017797

Janes, A. C., Smoller, J. W., David, S. P., Frederick, B. D., Haddad, S., Basu, A., . . . Kaufman, M. J. (2012). Association between CHRNA5 genetic variation at rs16969968 and brain reactivity to smoking images in nicotine dependent women. *Drug Alcohol Depend, 120*(1-3), 7-13. doi:10.1016/j.drugalcdep.2011.06.009

Jasinska, A. J., Stein, E. A., Kaiser, J., Naumer, M. J., & Yalachkov, Y. (2014). Factors modulating neural reactivity to drug cues in addiction: a survey of human neuroimaging studies. *Neurosci Biobehav Rev, 38*, 1-16. doi:10.1016/j.neubiorev.2013.10.013

Jasinska, A. J., Stein, E. A., Kaiser, J., Naumer, M. J., & Yalachkov, Y. (2014). Factors modulating neural reactivity to drug cues in addiction: a survey of human neuroimaging studies. *Neuroscience and biobehavioral reviews, 38*, 1-16. doi:10.1016/j.neubiorev.2013.10.013

Jorde, A., Bach, P., Witt, S. H., Becker, K., Reinhard, I., Vollstädt-Klein, S., . . . Kiefer, F. (2014). Genetic variation in the atrial natriuretic peptide transcription factor GATA4 modulates amygdala responsiveness in alcohol dependence. *Biological psychiatry, 75*(10), 790-797. doi:10.1016/j.biopsych.2013.10.020

Jorm, A. F. (2015). Using the Delphi expert consensus method in mental health research. *Aust N Z J Psychiatry, 49*(10), 887-897. doi:10.1177/0004867415600891

Joseph, J. E., McRae-Clark, A., Sherman, B. J., Baker, N. L., Moran-Santa Maria, M., & Brady, K. T. (2019). Neural correlates of oxytocin and cue reactivity in cocaine-dependent men and women with and without childhood trauma. *Psychopharmacology*. doi:10.1007/s00213-019-05360-7

Josephs, O., & Henson, R. N. (1999). Event-related functional magnetic resonance imaging: modelling, inference and optimization. *Philos Trans R Soc Lond B Biol Sci, 354*(1387), 1215-1228. doi:10.1098/rstb.1999.0475

Kareken, D. A., Liang, T., Wetherill, L., Dzemidzic, M., Bragulat, V., Cox, C., . . . Foroud, T. (2010). A polymorphism in GABRA2 is associated with the medial frontal response to alcohol cues in an fMRI study. *Alcohol Clin Exp Res, 34*(12), 2169-2178. doi:10.1111/j.1530-0277.2010.01293.x

Karoly, H. C., Schacht, J. P., Meredith, L. R., Jacobus, J., Tapert, S. F., Gray, K. M., & Squeglia, L. M. (2019). Investigating a novel fMRI cannabis cue reactivity task in youth. *Addict Behav, 89*, 20-28. doi:10.1016/j.addbeh.2018.09.015

Kleykamp, B. A., De Santis, M., Dworkin, R. H., Huhn, A. S., Kampman, K. M., Montoya, I. D., . . . Strain, E. C. (2019). Craving and opioid use disorder: A scoping review. *Drug Alcohol Depend, 205*, 107639. doi:10.1016/j.drugalcdep.2019.107639

Kober, H., Lacadie, C. M., Wexler, B. E., Malison, R. T., Sinha, R., & Potenza, M. N. (2016). Brain activity during cocaine craving and gambling urges: an fMRI study. *Neuropsychopharmacology, 41*(2), 628-637.

Kosten, T. R., Scanley, B. E., Tucker, K. A., Oliveto, A., Prince, C., Sinha, R., . . . Wexler, B. E. (2006). Cue-Induced Brain Activity Changes and Relapse in Cocaine-Dependent Patients. *Neuropsychopharmacology, 31*(3), 644-650. doi:10.1038/sj.npp.1300851

Kozlowski, L. T., Mann, R. E., Wilkinson, D., & Poulos, C. X. (1989). “Cravings” are amibiguous: Ask about urges or desires. *Addictive Behaviors, 14*(4), 443-445.

Kozlowski, L. T., & Wilkinson, D. A. (1987). Use and misuse of the concept of craving by alcohol, tobacco, and drug researchers. *British Journal of addiction, 82*(1), 31-36.

Kuhn, A. B., Feis, D. L., Schilbach, L., Kracht, L., Hess, M. E., Mauer, J., . . . Tittgemeyer, M. (2016). FTO gene variant modulates the neural correlates of visual food perception. *Neuroimage, 128*, 21-31. doi:10.1016/j.neuroimage.2015.12.049

Lang, P. J. (2005). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. *Technical report*.

LaRowe, S. D., Saladin, M. E., Carpenter, M. J., & Upadhyaya, H. P. (2007). Reactivity to nicotine cues over repeated cue reactivity sessions. *Addict Behav, 32*(12), 2888-2899. doi:10.1016/j.addbeh.2007.04.025

Li, Q., Wang, Y., Zhang, Y., Li, W., Yang, W., Zhu, J., . . . Qin, W. (2012). Craving correlates with mesolimbic responses to heroin-related cues in short-term abstinence from heroin: an event-related fMRI study. *Brain research, 1469*, 63-72.

Lindquist, M. (2020). Neuroimaging results altered by varying analysis pipelines: Nature Publishing Group.

Lindsey, K. P., Bracken, B. K., Maclean, R. R., Ryan, E. T., Lukas, S. E., & Frederick, B. D. (2013). Nicotine content and abstinence state have different effects on subjective ratings of positive versus negative reinforcement from smoking. *Pharmacol Biochem Behav, 103*(4), 710-716. doi:10.1016/j.pbb.2012.11.012

Lorenz, R. C., Kruger, J. K., Neumann, B., Schott, B. H., Kaufmann, C., Heinz, A., & Wustenberg, T. (2013). Cue reactivity and its inhibition in pathological computer game players. *Addict Biol, 18*(1), 134-146. doi:10.1111/j.1369-1600.2012.00491.x

Lou, M., Wang, E., Shen, Y., & Wang, J. (2012). Cue-elicited craving in heroin addicts at different abstinent time: an fMRI pilot study. *Subst Use Misuse, 47*(6), 631-639. doi:10.3109/10826084.2011.646381

Lu, L., Grimm, J. W., Hope, B. T., & Shaham, Y. (2004). Incubation of cocaine craving after withdrawal: a review of preclinical data. *Neuropharmacology, 47 Suppl 1*, 214-226. doi:10.1016/j.neuropharm.2004.06.027

Macatee, R. J., Carr, M., Afshar, K., & Preston, T. J. (2020). Development and Validation of a Cannabis Cue Stimulus Set. *Addict Behav*, 106643.

Manoliu, A., Haugg, A., Sladky, R., Hulka, L., Kirschner, M., Bruhl, A. B., . . . Scharnowski, F. (2020). SmoCuDa: A Validated Smoking Cue Database to Reliably Induce Craving in Tobacco Use Disorder. *Eur Addict Res*, 1-8. doi:10.1159/000509758

Mayer, A. R., Ling, J. M., Dodd, A. B., Shaff, N. A., Wertz, C. J., & Hanlon, F. M. (2019). A comparison of denoising pipelines in high temporal resolution task-based functional magnetic resonance imaging data. *Human brain mapping, 40*(13), 3843-3859. doi:10.1002/hbm.24635

McBride, D., Barrett, S. P., Kelly, J. T., Aw, A., & Dagher, A. (2006). Effects of expectancy and abstinence on the neural response to smoking cues in cigarette smokers: an fMRI study. *Neuropsychopharmacology, 31*(12), 2728-2738. doi:10.1038/sj.npp.1301075

McClernon, F. J., Conklin, C. A., Kozink, R. V., Adcock, R. A., Sweitzer, M. M., Addicott, M. A., . . . DeVito, A. M. (2016). Hippocampal and insular response to smoking-related environments: neuroimaging evidence for drug-context effects in nicotine dependence. *Neuropsychopharmacology, 41*(3), 877-885.

McClernon, F. J., Hiott, F. B., Huettel, S. A., & Rose, J. E. (2005). Abstinence-induced changes in self-report craving correlate with event-related FMRI responses to smoking cues. *Neuropsychopharmacology, 30*(10), 1940-1947.

McClernon, F. J., Hutchison, K. E., Rose, J. E., & Kozink, R. V. (2007). DRD4 VNTR polymorphism is associated with transient fMRI-BOLD responses to smoking cues. *Psychopharmacology (Berl), 194*(4), 433-441. doi:10.1007/s00213-007-0860-6

McClernon, F. J., Kozink, R. V., Lutz, A. M., & Rose, J. E. (2009). 24-h smoking abstinence potentiates fMRI-BOLD activation to smoking cues in cerebral cortex and dorsal striatum. *Psychopharmacology (Berl), 204*(1), 25-35. doi:10.1007/s00213-008-1436-9

McClernon, F. J., Kozink, R. V., & Rose, J. E. (2008). Individual Differences in Nicotine Dependence, Withdrawal Symptoms, and Sex Predict Transient fMRI-BOLD Responses to Smoking Cues. *Neuropsychopharmacology, 33*(9), 2148-2157. doi:10.1038/sj.npp.1301618

McHugh, R. K., Fulciniti, F., Mashhoon, Y., & Weiss, R. D. (2016). Cue-induced craving to paraphernalia and drug images in opioid dependence. *The American journal on addictions, 25*(2), 105-109. doi:10.1111/ajad.12344

Moeller, S. J., Parvaz, M. A., Shumay, E., Beebe-Wang, N., Konova, A. B., Alia-Klein, N., . . . Goldstein, R. Z. (2013). Gene x abstinence effects on drug cue reactivity in addiction: multimodal evidence. *J Neurosci, 33*(24), 10027-10036. doi:10.1523/JNEUROSCI.0695-13.2013

Moeller, S. J., & Paulus, M. P. (2018). Toward biomarkers of the addicted human brain: Using neuroimaging to predict relapse and sustained abstinence in substance use disorder. *Prog Neuropsychopharmacol Biol Psychiatry, 80*(Pt B), 143-154. doi:10.1016/j.pnpbp.2017.03.003

Moeller, S. J., Zilverstand, A., Konova, A. B., Kundu, P., Parvaz, M. A., Preston-Campbell, R., . . . Goldstein, R. Z. (2018). Neural correlates of drug-biased choice in currently using and abstinent individuals with cocaine use disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 3*(5), 485-494.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS med, 6*(7), e1000097.

Nichols, T. E., Das, S., Eickhoff, S. B., Evans, A. C., Glatard, T., Hanke, M., . . . Poline, J.-B. (2017). Best practices in data analysis and sharing in neuroimaging using MRI. *Nature neuroscience, 20*(3), 299-303.

Nichols, T. E., Das, S., Eickhoff, S. B., Evans, A. C., Glatard, T., Hanke, M., . . . Yeo, B. T. (2017). Best practices in data analysis and sharing in neuroimaging using MRI. *Nat Neurosci, 20*(3), 299-303. doi:10.1038/nn.4500

O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med, 89*(9), 1245-1251. doi:10.1097/ACM.0000000000000388

Panman, J. L., To, Y. Y., van der Ende, E. L., Poos, J. M., Jiskoot, L. C., Meeter, L. H. H., . . . Hafkemeijer, A. (2019). Bias Introduced by Multiple Head Coils in MRI Research: An 8 Channel and 32 Channel Coil Comparison. *Front Neurosci, 13*, 729. doi:10.3389/fnins.2019.00729

Parvaz, M. A., Moeller, S. J., & Goldstein, R. Z. (2016). Incubation of Cue-Induced Craving in Adults Addicted to Cocaine Measured by Electroencephalography. *JAMA Psychiatry, 73*(11), 1127-1134. doi:10.1001/jamapsychiatry.2016.2181

Perry, R. N., Schlagintweit, H. E., Darredeau, C., Helmick, C., Newman, A. J., Good, K. P., & Barrett, S. P. (2019). The impacts of actual and perceived nicotine administration on insula functional connectivity with the anterior cingulate cortex and nucleus accumbens. *J Psychopharmacol, 33*(12), 1600-1609. doi:10.1177/0269881119872205

Plant, R. R., Hammond, N., & Whitehouse, T. (2003). How choice of mouse may affect response timing in psychological studies. *Behavior Research Methods, Instruments, & Computers, 35*(2), 276-284.

Plant, R. R., & Turner, G. (2009). Millisecond precision psychological research in a world of commodity computers: New hardware, new problems? *Behavior Research Methods, 41*(3), 598-614.

Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafo, M. R., . . . Yarkoni, T. (2017). Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nat Rev Neurosci, 18*(2), 115-126. doi:10.1038/nrn.2016.167

Poldrack, R. A., Fletcher, P. C., Henson, R. N., Worsley, K. J., Brett, M., & Nichols, T. E. (2008). Guidelines for reporting an fMRI study. *Neuroimage, 40*(2), 409-414. doi:10.1016/j.neuroimage.2007.11.048

Potenza, M. N., Hong, K.-i. A., Lacadie, C. M., Fulbright, R. K., Tuit, K. L., & Sinha, R. (2012). Neural correlates of stress-induced and cue-induced drug craving: influences of sex and cocaine dependence. *American journal of psychiatry, 169*(4), 406-414.

Prisciandaro, J. J., Joseph, J. E., Myrick, H., McRae-Clark, A. L., Henderson, S., Pfeifer, J., & Brady, K. T. (2014). The relationship between years of cocaine use and brain activation to cocaine and response inhibition cues. *Addiction, 109*(12), 2062-2070. doi:10.1111/add.12666

Prisciandaro, J. J., Myrick, H., Henderson, S., McRae-Clark, A. L., & Brady, K. T. (2013). Prospective associations between brain activation to cocaine and no-go cues and cocaine relapse. *Drug Alcohol Depend, 131*(1-2), 44-49. doi:10.1016/j.drugalcdep.2013.04.008

Rubinstein, M. L., Luks, T. L., Moscicki, A.-B., Dryden, W., Rait, M. A., & Simpson, G. V. (2011). Smoking-related cue-induced brain activation in adolescent light smokers. *The Journal of adolescent health : official publication of the Society for Adolescent Medicine, 48*(1), 7-12. doi:10.1016/j.jadohealth.2010.09.016

Sacchet, M. D., & Knutson, B. (2013). Spatial smoothing systematically biases the localization of reward-related brain activity. *Neuroimage, 66*, 270-277. doi:10.1016/j.neuroimage.2012.10.056

Schacht, J. P., Anton, R. F., & Myrick, H. (2013). Functional neuroimaging studies of alcohol cue reactivity: a quantitative meta-analysis and systematic review. *Addict Biol, 18*(1), 121-133. doi:10.1111/j.1369-1600.2012.00464.x

Schacht, J. P., Anton, R. F., Randall, P. K., Li, X., Henderson, S., & Myrick, H. (2011). Stability of fMRI striatal response to alcohol cues: a hierarchical linear modeling approach. *Neuroimage, 56*(1), 61-68.

Schacht, J. P., Randall, P. K., Latham, P. K., Voronin, K. E., Book, S. W., Myrick, H., & Anton, R. F. (2017). Predictors of Naltrexone Response in a Randomized Trial: Reward-Related Brain Activation, OPRM1 Genotype, and Smoking Status. *Neuropsychopharmacology, 42*(13), 2640-2653. doi:10.1038/npp.2017.74

Schacht, J. P., Voronin, K. E., Randall, P. K., & Anton, R. F. (2018). Dopaminergic Genetic Variation Influences Aripiprazole Effects on Alcohol Self-Administration and the Neural Response to Alcohol Cues in a Randomized Trial. *Neuropsychopharmacology, 43*(6), 1247-1256. doi:10.1038/npp.2017.298

Schulz, K. F., Altman, D. G., Moher, D., & Group, C. (2010). CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *Trials, 11*(1), 32.

Segalowitz, S. J., & Graves, R. E. (1990). Suitability of the IBM XT, AT, and PS/2 keyboard, mouse, and game port as response devices in reaction time paradigms. *Behavior Research Methods, Instruments, & Computers, 22*(3), 283-289.

Seo, D., Lacadie, C. M., Tuit, K., Hong, K. I., Constable, R. T., & Sinha, R. (2013). Disrupted ventromedial prefrontal function, alcohol craving, and subsequent relapse risk. *JAMA Psychiatry, 70*(7), 727-739. doi:10.1001/jamapsychiatry.2013.762

Siegle, G. J., Steinhauer, S. R., Thase, M. E., Stenger, V. A., & Carter, C. S. (2002). Can't shake that feeling: event-related fMRI assessment of sustained amygdala activity in response to emotional information in depressed individuals. *Biol Psychiatry, 51*(9), 693-707.

Stritzke, W. G., Breiner, M. J., Curtin, J. J., & Lang, A. R. (2004). Assessment of substance cue reactivity: advances in reliability, specificity, and validity. *Psychol Addict Behav, 18*(2), 148-159. doi:10.1037/0893-164X.18.2.148

Tiffany, S. T., & Drobes, D. J. (1991). The development and initial validation of a questionnaire on smoking urges. *British Journal of addiction, 86*(11), 1467-1476.

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International journal for quality in health care, 19*(6), 349-357.

Verdejo-Garcia, A., Lorenzetti, V., Manning, V., Piercy, H., Bruno, R., Hester, R., . . . Bates, M. E. (2019). A roadmap for integrating neuroscience into addiction treatment: a consensus of the neuroscience interest Group of the international society of addiction medicine. *Front Psychiatry, 10*, 877.

Volkow, N. D., Wang, G.-J., Telang, F., Fowler, J. S., Logan, J., Childress, A.-R., . . . Wong, C. (2006). Cocaine Cues and Dopamine in Dorsal Striatum: Mechanism of Craving in Cocaine Addiction. *The Journal of Neuroscience, 26*(24), 6583. doi:10.1523/JNEUROSCI.1544-06.2006

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gotzsche, P. C., Vandenbroucke, J. P., & Initiative, S. (2008). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol, 61*(4), 344-349. doi:10.1016/j.jclinepi.2007.11.008

Wall, M. B., Mentink, A., Lyons, G., Kowalczyk, O. S., Demetriou, L., & Newbould, R. D. (2017). Investigating the neural correlates of smoking: feasibility and results of combining electronic cigarettes with fMRI. *Scientific reports, 7*(1), 1-8.

Wertz, J. M., & Sayette, M. A. (2001). Effects of smoking opportunity on attentional bias in smokers. *Psychol Addict Behav, 15*(3), 268-271.

Wetherill, R. R., Young, K. A., Jagannathan, K., Shin, J., O'Brien, C. P., Childress, A. R., & Franklin, T. R. (2013). The impact of sex on brain responses to smoking cues: a perfusion fMRI study. *Biology of sex differences, 4*(1), 9-9. doi:10.1186/2042-6410-4-9

Wilson, S. J., & Sayette, M. A. (2015). Neuroimaging craving: urge intensity matters. *Addiction, 110*(2), 195-203.

Wilson, S. J., Sayette, M. A., & Fiez, J. A. (2004). Prefrontal responses to drug cues: a neurocognitive analysis. *Nature neuroscience, 7*(3), 211-214.

Xu, K., Seo, D., Hodgkinson, C., Hu, Y., Goldman, D., & Sinha, R. (2013). A variant on the kappa opioid receptor gene (OPRK1) is associated with stress response and related drug craving, limbic brain activation and cocaine relapse risk. *Transl Psychiatry, 3*, e292. doi:10.1038/tp.2013.62

Yalachkov, Y., Kaiser, J., Gorres, A., Seehaus, A., & Naumer, M. J. (2013). Sensory modality of smoking cues modulates neural cue reactivity. *Psychopharmacology (Berl), 225*(2), 461-471. doi:10.1007/s00213-012-2830-x

Yang, B. Z., Balodis, I. M., Lacadie, C. M., Xu, J., & Potenza, M. N. (2016). A Preliminary Study of DBH (Encoding Dopamine Beta-Hydroxylase) Genetic Variation and Neural Correlates of Emotional and Motivational Processing in Individuals With and Without Pathological Gambling. *J Behav Addict, 5*(2), 282-292. doi:10.1556/2006.5.2016.026

Yucel, M., Oldenhof, E., Ahmed, S. H., Belin, D., Billieux, J., Bowden-Jones, H., . . . Verdejo-Garcia, A. (2019). A transdiagnostic dimensional approach towards a neuropsychological assessment for addiction: an international Delphi consensus study. *Addiction, 114*(6), 1095-1109. doi:10.1111/add.14424

Zeng, H., Su, D., Wang, P., Wang, M., Vollstädt-Klein, S., Chen, Q., & Ye, H. (2018). The Action Representation Elicited by Different Types of Drug-Related Cues in Heroin-Abstinent Individuals. *Front Behav Neurosci, 12*, 123. doi:10.3389/fnbeh.2018.00123