

# **Crowe Critical Appraisal Tool (CCAT) User Guide**

**Version 1.4 (19 November 2013)**

**Use with the CCAT Form version 1.4 only**

**Michael Crowe, PhD**

michael.crowe@my.jcu.edu.au



This work is licensed under the Creative Commons Attribution-ShareAlike 3.0 Unported License.  
To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/3.0/>

## Summary of main points

- The Crowe Critical Appraisal Tool (CCAT) consists of
  - The CCAT Form
  - The CCAT User Guide.
- Always use the CCAT Form **and** the CCAT User Guide together.
- Research designs should be appraised on their own merits, not to a ‘gold standard’.
- All categories must be scored: it does not matter which research design was used
  - The lowest score for a category is 0, the highest score is 5
  - Category scores are whole numbers only (that is 0, 1, 2, 3, 4, or 5)
  - The score for each category must be reported
  - The total score (out of 40 or as a percent) is reported in addition to each category score.
- Item descriptors may be marked  present,  absent, or  not applicable.
  - Tick marks are **not** a check list to be totalled.
  - Tick marks are simply a guide to scoring a category.
- If in doubt use your best judgement, there is no right or wrong answer.

## Contents

Introduction.....	2
Overview of scoring a paper.....	3
Guidelines for scoring categories and items.....	5
1. Preliminaries.....	5
2. Introduction.....	5
3. Design.....	6
4. Sampling.....	7
5. Data collection.....	8
6. Ethical matters.....	9
7. Results.....	10
8. Discussion.....	11
9. Total.....	11
References.....	12
Version information.....	12
My notes.....	13

## Introduction

The Crowe Critical Appraisal Tool (CCAT) consists of the CCAT Form (the Form) and the CCAT User Guide (the User Guide). The Form and the User Guide must be used together otherwise validity and reliability of the scores obtained may be severely compromised. Any changes made to the categories, items, or item descriptors, no matter how small, may also compromise the validity and reliability of the scores obtained. Changes made to the CCAT Form must be tested to verify the validity and reliability of the scores and score compatibility with other versions of the Form.

The CCAT is demanding. It assumes that you are familiar with research designs, sampling techniques, ethics, data collection methods, and statistical and non-statistical data analysis techniques. Therefore, it may be helpful to have a general research methods text book available when you appraise papers.

The information sought when appraising a paper is unlikely to be in the sequence outlined in the Form. Therefore, it is suggested that you read each paper quickly from start to finish getting an overall sense of what is being discussed.

On the first reading of a paper, these sections on the first page of the Form may be completed before you begin scoring the paper

- *Reference* Keep track of papers appraised with a unique identification.
- *Reviewer* Identify the reviewer of the paper, especially if more than one per article.
- *Citation* Match the Form with the paper appraised.
- *Research design* Indicate the research design or designs used in the paper.
  - The listed research designs are the most common ones, but other designs exist. This is why an ellipsis (...) is included at the end of each list.
  - Some descriptive, exploratory, observational (DEO) research designs may be described as a combination of items from row A and row B on the Form (e.g. prospective cohort, longitudinal survey).
- *Variables and analysis* Describe the intervention(s)/treatment(s)/exposure(s), outcome(s)/output(s)/predictor(s)/measure(s), and data analysis method(s) used. Include comments on the variables and analysis.
- *Sampling* Write down the total sample size and the sample size for each group, where applicable. Briefly describe the sample and the population the sample was selected from. Note any questions that occur about the sample.
- *Data collection* Indicate the data collection method or methods used.
- *General notes* Add thoughts and in-depth analysis during the appraisal process.

Next, re-read the paper and fill in the second page of the Form. Insert any notes or page numbers where you found relevant information as you read the paper. This will help to jog your memory if you need to go through the paper in the future or need to justify your appraisal.

Some categories have the prompt 'Is it worth continuing?' If there are serious flaws in a paper in any of these categories, you should determine if it is worth continuing to appraise the paper or whether appraisal should be abandoned and the paper rejected.

Finally, transfer the scores from the second page to the first page of the Form. By doing this, the majority of the information required for the appraisal is on the first page.

## Overview of scoring a paper

The Form is divided into eight categories and 22 items. Each item has multiple item descriptors that make it easier to appraise and score a category. Each category receives its own score on a 6 point scale from 0–5. The lowest score a category can achieve is 0, and 5 is the highest score. Categories can only be scored as a whole number or integer, i.e. 0, 1, 2, 3, 4, or 5, that is half marks are not allowed.

There are tick boxes () beside item descriptors. The tick box is useful to indicate if the item descriptor is

- Present () – For an item descriptor to be marked as present, there should be evidence of it being present rather than an assumption of presence.
- Absent () – For an item descriptor to be marked as absent, it is implied that it should be present in the first place.
- Not applicable () – For an item descriptor to be marked as not applicable, the descriptor must not be relevant given the characteristics of the paper being appraised and is, therefore, not considered when assigning a score to a category.

Whether an item descriptor is present, absent, or not applicable is further explored in the section *Guidelines for scoring categories and items*. All categories must be scored because all categories are applicable in all research designs. Only item descriptors may be marked ‘not applicable’.

While it may be tempting to add up all the present marks () and all the absent marks () in each category and to use the proportion of one to the other to calculate the score for the category, this is not appropriate. It is incorrect because not all item descriptors in a category have equal importance. For example, in the *Introduction* category there are two items (*Background* and *Objective*) and a total of five tick boxes. If a paper being appraised has all boxes marked as present () except for *Primary objective(s)*, *hypothesis(es)*, or *aim(s)*, which is marked as absent (), should the paper be scored 4/5 for that category? It could be argued that a research paper without a primary objective, hypothesis, or aim is fundamentally flawed and, as a result, should be scored 0/5 even though the other four tick boxes were marked as present.

Therefore, the tick marks for present, absent, or not applicable are to be used as a guide to scoring a category and not as a simple check list. It is up to you as the appraiser to take into consideration all aspects of each category and based on both the tick marks and judgement assign a score to a category.

Similarly, the research design used in each paper should be appraised on its own merits and not relative to some preconceived notion of a hierarchy of research designs or ‘gold standard’. What is most important is that the paper used an appropriate research design based on the research question being addressed, rather than what research design was used.

The total score given to a paper can be expressed as a percentage by dividing the *Total* by 40 (that is, eight categories multiplied by the maximum score of five) and writing the result on the first page of the Form. The *Total %* should be written to the nearest full percent (Table 1). There is no need for decimal places because they do not add anything to the accuracy of the score obtained.

Finally, the *Total* or *Total %* score a paper obtains is not the sole criterion on which an overall assessment of a paper is based. The *Total* or *Total %* score is a useful summary but may not be applicable in all cases. When reporting an appraisal using the CCAT, the score obtained in

every category must be stated along with the *Total* or *Total %* score. This prevents papers that score high overall but very poor in one or more categories being hidden amongst papers which scored high throughout all categories. Based on the reasons for the appraisal, some papers which have a low score in certain category but which have a high total score may be ranked lower than those with a lower total score but a high score in that particular category. These processes are up to you, as the appraiser, to detail before you begin appraising papers.

**Table 1** *Total* and corresponding *Total %*

Total	Total %	Total	Total %	Total	Total %	Total	Total %
0	0	10	25	20	50	30	75
1	3	11	28	21	53	31	78
2	5	12	30	22	55	32	80
3	8	13	33	23	58	33	83
4	10	14	35	24	60	34	85
5	13	15	38	25	63	35	88
6	15	16	40	26	65	36	90
7	18	17	43	27	68	37	93
8	20	18	45	28	70	38	95
9	23	19	48	29	73	39	98

## Guidelines for scoring categories and items

### 1. Preliminaries

#### Title

1. Includes study aims and design
  - Traditionally only required for reporting research.
  - It has been assumed that this does not affect the overall quality of the research but there is little evidence one way or the other.

#### Abstract

1. Contains key information
  - Traditionally only required for reporting research.
  - It has been assumed that this does not affect the overall quality of the research but there is little evidence one way or the other.
2. Balanced and informative
  - Traditionally only required for reporting research.
  - It has been assumed that this does not affect the overall quality of the research but there is little evidence one way or the other.

#### Text

**Note** This item can only be assessed when the article has been read in full.

1. Sufficient detail others could reproduce
  - This is an over-arching concept and should be present throughout the study.
2. Clear, concise writing/table(s)/diagram(s)/figure(s)
  - This is an over-arching concept and should be present throughout the study.

### 2. Introduction

#### Background

1. Summary of current knowledge
  - Current and applicable knowledge provides a context for the study.
2. Specific problem(s) addressed and reason(s) for addressing
  - Description of why the study was undertaken.
  - Links current knowledge and stated objective(s), hypothesis(es), or aim(s).

#### Objective

1. Primary objective(s), hypothesis(es), aim(s)
  - The study must have at least one stated objective, hypothesis, or aim.
2. Secondary question(s)
  - Secondary question(s) may sometimes arise based on the primary objective(s), hypothesis(es), or aim(s).
  - Since this is not always the case, a study without secondary questions should not be penalised.

### 3. Design

#### Research design

1. Research design(s) chosen and why
  - Description of the research design chosen and why it was chosen.
2. Suitability of research design(s)
  - The research design should be congruent with *Background, Objective, Intervention(s)/treatment(s)/exposure(s)*, and *Outcome(s)/output(s)/predictor(s)*.

#### Intervention, Treatment, Exposure

1. Intervention(s)/treatment(s)/exposure(s) chosen and why
  - Where a study does not normally have an intervention/treatment/exposure, it should not be penalised when none is present.
  - Statement for every intervention/treatment/exposure chosen and why it was chosen.
  - Each intervention/treatment/exposure must be congruent with *Background, Objective*, and *Research design*.
2. Precise details of the intervention(s)/treatment(s)/exposure(s) for each group
  - Full details are presented for every intervention/treatment/exposure for every participant/case/group so that other studies could duplicate.
3. Intervention(s)/treatment(s)/exposure(s) valid and reliable
  - A statement of reliability/validation or why there is no validation/reliability for each intervention/treatment/exposure.

#### Outcome, Output, Predictor, Measure

1. Outcome(s)/output(s)/predictor(s)/measure(s) chosen and why
  - All research has at least one expected outcome/output/predictor/measure.
  - Statement for each outcome/output/predictor/measure chosen and why it was chosen.
  - Each outcome/output/predictor/measure must be congruent with *Background, Objective, Research design*, and *Intervention/treatment/exposure*.
2. Clearly define outcome(s)/output(s)/predictor(s)/measure(s)
  - Full details are presented of every expected outcome/output/predictor/measure for every participant/case/group so that other studies could duplicate.
3. Outcome(s)/output(s)/predictor(s)/measure(s) valid and reliable
  - A statement of reliability/validation or why there is no validation/reliability for each outcome/output/predictor/measure.

**Note** In some cases the Outcome(s)/output(s)/predictor(s)/measure(s) may be similar to or the same as the Objective(s), hypothesis(es), aim(s). However, in most cases to achieve the Objective(s), hypothesis(es), aim(s) a series of Outcome(s)/output(s)/predictor(s)/measure(s) are required.

#### Bias, etc.

1. Potential sources of bias, confounding variables, effect modifiers, interactions
  - Identification of potential sources of:
    - Bias – e.g. attrition, detection, experimental, information, interview, observation, performance, rater, recall, selection.

- Confounding variables or factors – A variable which interferes between the intervention/treatment/exposure and the outcome/output/predictor/measure.
  - Effect modification – A variable which modifies the association between the intervention/treatment/exposure and the outcome/output/predictor/measure.
  - Interaction effects – When various combinations of intervention(s)/treatment(s)/exposure(s) cause different outcome(s)/output(s)/predictor(s)/measure(s).
  - Should be identified, as far as possible, within the *Research design* before data collection begins in order to minimise their effect.
  - See also *Sampling* and *Data collection*.
2. Sequence generation, group allocation, group balance, and by whom
    - In studies where participants/cases are allocated to groups, the methods used should be stated and procedures established before recruitment or data collection begins (e.g. blinding, method used to randomise, allocate to or balance groups).
  3. Equivalent treatment of participants/cases/groups
    - Each participant/case/group must be treated equivalently apart from any intervention/treatment/exposure.
    - If participants/cases/groups are not treated equivalently a statement regarding why this was not possible, how this may affect results, and procedures in place for managing participants/cases/groups.
    - See also *Sampling protocol*, *Collection protocol*, and *Participant ethics*.

## 4. Sampling

### Sampling method

1. Sampling method(s) chosen and why
  - Description of the sampling method chosen and why it was chosen.
  - Sampling methods are normally probability or non-probability based.
  - Examples include: Simple random, systematic, stratified, cluster, convenience, representative, purposive, snowball, and theoretical.
  - Also included here is the search strategy used for a systematic review (e.g. databases searched, search terms).
2. Suitability of sampling method
  - The sampling method should be decided and in place before recruitment or data collection begins.
  - The sampling method should be congruent with *Objective*, *Research design*, *Intervention/treatment/exposure*, *Outcome/output/predictor/measure*, and *Bias etc.*

### Sample size

1. Sample size, how chosen, and why
  - Description of the sample size, the method of sample size calculation, and why that method was chosen.
  - Sample size calculations are normally probability or non-probability based.
  - Examples of how calculations can be made include: Accuracy [e.g. confidence interval ( $\alpha$ ), population or sample variance ( $s^2$ ,  $\sigma^2$ ), effect size or index (ES, d), power ( $1-\beta$ )], analysis, population, redundancy, saturation, and budget.



## 2. Suitability of sample size

- The sample size or estimate of sample size, with contingencies, should be described and calculated before recruitment/data collection begins.
- The sample size should be congruent with *Objective, Research design, Intervention/treatment/exposure, Outcome/output/predictor/measure, and Bias etc.*

**Note** Sample size calculations are not required for systematic reviews, because it is not possible to know the number of papers that will meet the selection criteria, or for some single system designs.

### Sampling protocol

#### 1. Description and suitability of target/actual/sample population(s)

- The target/actual/sample population(s) should be described.
- The target/actual/sample population(s) should be congruent with *Objective, Research design, Intervention/treatment/exposure, Outcome/output/predictor/measure, and Bias etc.*

#### 2. Inclusion and exclusion criteria for participants/cases/groups

- Inclusion and exclusion criteria should be explicitly stated and established before recruitment/data collection begins.
- The use of inclusion and exclusion criteria (especially exclusion criteria) should not be used in such a way as to bias the sample.

#### 3. Recruitment of participants/cases/groups

- Description of procedures for recruitment and contingencies put in place.
- Recruitment should be congruent with *Objective, Research design, Intervention/treatment/exposure, Bias etc.*, and other aspects of *Sampling*.
- See also *Participant ethics, Researcher ethics, and Collection protocol*.

**Note** For systematic reviews inclusion and exclusion criteria only need to be appraised, because they refer to the parameters used to select papers.

## 5. Data collection

### Collection method

#### 1. Collection method(s) chosen and why

- Description of the method(s) used to collect data and why each was chosen.
- In systematic reviews, this refers to how information was extracted from papers, because these are the data collected.

#### 2. Suitability of collection method(s)

- The data collection method(s) should be congruent with *Objective, Research design, Intervention/treatment/exposure, Outcome/output/predictor/measure, Bias etc.*, and *Sampling*.

### Collection protocol

#### 1. Include date(s), location(s), setting(s), personnel, materials, processes

- Description of and details regarding exactly how data were collected, especially any factor(s) which may affect *Outcome/output/predictor/measure* or *Bias etc.*

## 2. Method(s) to ensure/enhance quality of measurement/instrumentation

- Description of any method(s) used to enhance or ensure the quality of data collected (e.g. pilot study, instrument calibration, standardised test(s), independent/multiple measurement, valid/reliable tools).
- Also includes any method(s) which reduce or eliminate bias, confounding variables, effect modifiers, interactions which are not an integral part of the *Design* category (e.g. blinding of participants, intervention(s), outcome(s), analysis; protocols and procedures implemented).
- In qualitative studies, this relates to concepts such as trustworthiness, authenticity, and credibility.
- See also *Bias etc.*

## 3. Manage non-participation, withdrawal, incomplete/lost data

- Description of any method(s) used to manage or prevent non-participation, withdrawal, or incomplete/lost data.
- These include but are not limited to: Intention to treat analysis (ITT); last observation carried forward (LOCF); follow up (FU), e.g. equal length, adequate, or complete; and, completer analysis, e.g. on-treatment, on-protocol.

## 6. Ethical matters

**Note** Some studies may have been conducted before Ethical matters were a major point of consideration. The research ethics standards of the time may need to be taken into account rather than the current standards.

**Note** All research requires Ethical matters consideration even if formal ethics committee or ethics board approval is not required. This includes systematic reviews.

### Participant ethics

#### 1. Informed consent, equity

- All participants must have provided their informed consent.
- Equity includes, but is not limited to, cultural respect, just and equitable actions, no harm to participants, debriefing, and consideration for vulnerable individuals or groups.

#### 2. Privacy, confidentiality/anonymity

- The privacy, confidentiality, or anonymity of participants must be catered for.
- If this is not possible, the informed and written consent of individuals affected must be obtained.

### Researcher ethics

#### 1. Ethical approval, funding, conflict(s) of interest

- A statement of ethical approval from recognised Ethics Committee(s) or Board(s) suitable for the study being undertaken.
- Any real, perceived, or potential conflict(s) of interest should be stated.
- All sources of funding should be stated.

2. Subjectivities, relationship(s) with participants/cases
  - Description of how the researcher(s) could have potentially or did affect the outcomes of the study through their presence or behaviour.
  - Includes a description of procedures used to minimise this occurring.
  - See also *Bias etc.*

## 7. Results

### Analysis, Integration, Interpretation method

1. A.I.I. (Analysis/Integration/Interpretation) method(s) for primary outcome(s)/output(s)/predictor(s) chosen and why
  - Description of statistical and non-statistical method(s) used to analyse/integrate/interpret Outcome(s)/output(s)/predictor(s)/measure(s) and why each was chosen.
2. Additional A.I.I. methods (e.g. subgroup analysis) chosen and why
  - Description of additional statistical and non-statistical method(s) used to analyse/integrate/interpret Outcome(s)/output(s)/predictor(s)/measure(s) and why each was chosen.
3. Suitability of analysis/integration/interpretation method(s)
  - The analysis/integration/interpretation method(s) should be congruent with *Objective, Research design, Intervention/treatment/exposure, Outcome/output/predictor, Bias etc., Sampling, and Data collection.*

### Essential analysis

1. Flow of participants/cases/groups through each stage of research
  - Description of how participants/cases/groups advanced through the study.
  - Explanation of course of intervention/treatment/exposure.
2. Demographic and other characteristics of participants/cases/groups
  - Description of baseline characteristics of participants/cases/groups so this can be integrated into the analysis.
3. Analyse raw data, response rate, non-participation, withdrawal, incomplete/lost data
  - Unadjusted data should be analysed.
  - There may be differences between those that completed and those that did not complete the study.

### Outcome, Output, Predictor analysis

1. Summary of results and precision for each outcome/output/predictor/measure
  - Results summarised with, where possible, an indicator of the precision and effect size of each result for each outcome/output/predictor/measure.
  - Where data are adjusted, make clear what was adjusted and why.
  - Where data are categorised, report of internal and external boundaries.
  - Use of quotations to illustrate themes/findings, privileging of subject meaning, adequate description of findings, evidence of reflexivity.

2. Consideration of benefits/harms, unexpected results, problems/failures
  - Description of all outcomes, not just ones being looked for.
  - Description of differences between planned and actual implementation, and the potential effect on results.
3. Description of outlying data (e.g. diverse cases, adverse effects, minor themes)
  - Exploration of outliers because they may not be anomalous.

## 8. Discussion

### Interpretation

1. Interpretation of results in the context of current evidence and objectives
  - Summarises key results in relation to *Background* and *Objective*.
  - Compare and contrast other research findings.
2. Draw inferences consistent with the strength of the data
  - Do not over or under represent data.
  - Draw inferences based on the entirety of available evidence.
  - See also *Sampling* and *Data collection*.
3. Consideration of alternative explanations for observed results
  - Exploration of reasons for differences between observed and expected.
  - Determines if other factors may lead to similar results.
4. Account for bias, confounding, interactions, effect modifiers, imprecision
  - Discussion on magnitude and direction of *Bias etc.* and how this may have affected the results.
  - See also *Essential analysis*.

### Generalisation

1. Consideration of overall practical usefulness of the study
  - Discussion on practical vs. theoretical usefulness.
2. Description of generalisability (external validity) of the study
  - Dependent on *Design*, *Sampling*, and *Data collection*.

### Concluding remarks

1. Highlight study's particular strengths
  - What did the study do well?
2. Suggest steps that may improve future results (e.g. limitations)
  - How could the study have been better?
3. Suggest further studies
  - Where should the next study begin?

## 9. Total

### Total score

1. Add all scores for categories 1–8
  - Total the scores for all categories.
  - To calculate the total percent, divide the total score by 40 (see p. 4).

## References

- Crowe, M., & Sheppard, L. (2011). A review of critical appraisal tools show they lack rigor: alternative tool structure is proposed. *Journal of Clinical Epidemiology*, 64(1), 79–89. doi:10.1016/j.jclinepi.2010.02.008
- Crowe, M., & Sheppard, L. (2011). A general critical appraisal tool: an evaluation of construct validity. *International Journal of Nursing Studies*, 48(12), 1505–1516. doi:10.1016/j.ijnurstu.2011.06.004
- Crowe, M., Sheppard, L., & Campbell, A. (2011). Comparison of the effects of using the Crowe Critical Appraisal Tool versus informal appraisal in assessing health research: a randomised trial. *International Journal of Evidence-Based Healthcare*, 9(4), 444–449. doi:10.1111/j.1744-1609.2011.00237.x
- Crowe, M., Sheppard, L., & Campbell, A. (2012). Reliability analysis for a proposed critical appraisal tool demonstrated value for diverse research designs. *Journal of Clinical Epidemiology*, 65(4), 375–383. doi:10.1016/j.jclinepi.2011.08.006

## Version information

### Changes between version 1.3 and 1.4

- Add: CCAT User Guide (p. 2) and Form (p. 1), Variables and analysis section.
- Add: CCAT Form (p. 2), scores ([/5] or [/40]) for each Category and Total.
- Update: CCAT User Guide, layout and grammar.

### Changes between version 1.2 and 1.3

- Add: CCAT User Guide (p. 1) and Form (p. 1), licensed under Creative Commons.

### Changes between version 1.1 and 1.2

- Emphasis: CCAT User Guide (p. 2) and Form (p. 1), any change to the CCAT form requires testing for validity and reliability of the scores.
- Emphasis: CCAT User Guide (p. 3), CCAT Form is not a check list and not to be used as one.
- Add: CCAT User Guide (p. 12), this version information section.
- Update: CCAT User Guide (p. 12), references for the CCAT User Guide and Form.

### Changes between version 1.0 and 1.1

- Add: CCAT Form (p. 1), letters A and B to the rows in DEO research designs.
- Add: CCAT User Guide (p. 12), references for the CCAT User Guide and Form.
- Update: CCAT Form (p. 1), layout.
- Update: CCAT User Guide, made it easier to read.

