AOTA’S EVIDENCE EXCHANGE
CRITICALLY APPRAISED PAPER (CAP) GUIDELINES

Annual AOTA Conference Poster Submissions

Critically Appraised Papers (CAPs) are at-a-glance summaries of the methods, findings and study limitations of selected individual articles. Critically Appraised Paper presentations are intended to provide a detailed appraisal of an individual research study to determine its value and relevance to occupational therapy.

Please submit your CAP through the AOTA Annual Conference & Expo Call for Papers. To prepare your CAP for submission, fill out the CAP Worksheet. Refer to the CAP Guidelines for submission criteria and detailed instructions on how to fill out each section of the worksheet.

When preparing the CAP, keep in mind that if accepted, you will present the final content as a poster presentation at AOTA’s Annual Conference. The CAP Worksheet should be filled out with strict adherence to the character count limit indicated on the CAP Worksheet.

The following criteria must be considered prior to selecting an article for appraisal and development of a Critically Appraised Paper (CAP):

- The article is an original quantitative research/study that describes an intervention within the scope of occupational therapy practice.
- The article was published in a peer-reviewed journal within 10 years of date of submission.
- The article is Level I (limited to randomized controlled trials), II, or III evidence. Only intervention studies will be considered. CAPs submitted on systematic reviews, assessment tools and measures, meta-analyses, descriptive studies, single-subject design, case reports, or expert opinions will not be accepted.
- The CAP cannot duplicate an article currently included in the Evidence Exchange. Additionally, more than one student from the same academic program cannot submit a CAP on the same article. To confirm that a CAP has not been previously published on your article of interest, check the Accepted CAPs list.

In addition, the following criteria must be met:

- The CAP must be formatted on the most up-to-date CAP Worksheet. CAPs submitted using an older version of the CAP Worksheet will be rejected.
- All sources must be appropriately paraphrased and cited. If plagiarism is identified, the CAP will be rejected.
- Direct quotes should be avoided. If needed, proper APA format must be used. Any CAP that contains uncited work will be rejected.
- Content should directly follow the CAP Guidelines and address the questions posted in the Guidelines.
- All acronyms must be spelled out when used for the first time.
- The CAP Worksheet must be previewed for errors in syntax and spelling.
- The DOI number for the original article must be provided in the Citation section. All submission must include an electronic version of the original article.
- Any additional articles referred to (cited) in the CAP need to be included on the
Reference list (include DOI number) using APA format (6th edition).

- If you are a student submitting a CAP, a faculty advisor must provide support throughout the CAP development process, including reviewing all CAPs prior to submission. Only exemplary CAPs that have been completed with support of a faculty advisor should be submitted.

CITATION AND DOI NUMBER
Using APA format (6th edition), include full title, all authors (last name, initials), full journal title, year, volume #, and page #s. Include the DOI number (Digital Object Identifier) for the original article. Numbers are often listed at the end of a reference in a reference list. An example citation:


CLINICAL BOTTOM LINE
This section is critical to the reader and should include a brief description of the intervention and discussion of how the evidence can be used to inform and guide occupational therapy practice (i.e., within the scope of traditional or emerging practice) AND how practitioners can use the evidence relative to the target population and practice setting. Keep these points in mind as you are developing this section:

- Are the recommendations stated in practical terms – i.e., are they ready to implement? Stating that the evidence has an “impact” or “effect” on practice is insufficient. Telling the reader how the evidence can be used is valuable.
- Do the limitations of the study affect the strength of the evidence? (i.e., type of study design, level of evidence, identified study limitations, internal and external threats to the study, to what population results may generalize)? Include implications identified by the author and any additional implications that you judge to be appropriate to clinical practice, program development, education and training, and research.
- Are the findings generalizable? Please keep in mind that you are reviewing only one article and that the implications should reflect only this single study.

RESEARCH OBJECTIVE(S), DESIGN TYPE, AND LEVEL OF EVIDENCE
List the objective(s) of the study. The objectives are usually stated briefly in the abstract of the article, and again in more detail in the introduction. They may be phrased as a statement “To evaluate or test the effectiveness of….” Please phrase in your own words. This should be no more than a brief sentence or two.

State specific type of study design from the article. While many research articles include the study design, some do not, and determining the design and level of evidence may not be clear-cut. Use the description of the study procedure to guide the decision on design and level of evidence. See Appendix for further explanation of study design types.
For purposes of the Evidence Exchange, submissions should be based on articles that are Level I (limited to randomized controlled trials), II, or III evidence. State the level of evidence using the following criteria (adapted from Sackett, D.L., Rosenberg, W.M., Muir Gray, J.A., Haynes, R.B. & Richardson, W.S. (1996). Evidence-based medicine: What it is and what it isn’t. British Medical Journal, 312, 71-72):

Level I: Randomized controlled trials
Level II: Two groups, nonrandomized studies (e.g., cohort, case-control)
Level III: One group, nonrandomized (e.g., before-and-after, pretest and posttest)

PARTICIPANT SELECTION
How were participants selected to participate? Describe sample selection procedures.
Inclusion criteria: What criteria did the researchers use to select participants (e.g., diagnosis, severity level, age)?
Exclusion criteria: What criteria did the researchers use to limit the study population (e.g., severity level, age)?

INTERVENTION(S) AND CONTROL GROUPS (Completed for each study group)
Brief description of the intervention: Provide a brief description of the intervention(s) including the location of the study. It must answer the “what?” and “how?” of the intervention.
How many participants in the group?
Where did the intervention(s) take place? Was intervention(s) received at home or in an institution? Was it the same for different groups of subjects if there was more than one intervention group?
Who delivered? Who provided the intervention(s)? Was it different for intervention and control groups?
How often? How often did the intervention(s) take place and what was the length of the individual intervention session (e.g., twice weekly for 30 minutes per session)?
For how long? How long did the intervention(s) last (e.g., 3 months)

OUTCOME MEASURES
Outcomes are the variables or issues of interest to the researcher. They represent the product or results of the intervention or exposure. Many studies include numerous outcome measures. For the purpose of the CAP, only include measures relevant to occupational therapy practice.
Complete the following for each measure:
Name of test/measure and what is measured: Example: McGill Pain Questionnaire (pain intensity and pain perception)
What outcome is measured: Name of the measure and brief description of the outcome measured
Is the measure reliable (as reported in the article)? Reliability refers to whether a measure was giving the same information over different situations. The two most common forms of reliability are (1) test–retest reliability—the same observer gets the same information on two
occasions separated by a short time interval, and (2) interrater reliability—different observers get the same information at the same time.

**Is the measure valid (as reported in the article)?** Validity asks whether the measure assesses what it is intended to measure. Consider if the measure includes all of the relevant concepts and elements of the outcome (content validity) and if the authors report that the measure has been tested in relationship to other measures to determine any relationship (criterion validity). For example, a “valid” ADL measure includes all relevant elements of self-care and has been tested with other measures of daily living activities and self-care functioning to determine that the relationship between the measures is as expected.

**When is the measure used?** It is important to note if outcomes were measured pre- and post-intervention, and whether short-term and/or long-term effects were considered.

**RESULTS**

List **key findings based on study objectives**. Key findings should include change (or lack of change) in outcomes relevant to occupational therapy practice and comparability of groups (as appropriate). Include statistical significance where appropriate ($p > .05$) and effect size if reported.

**LIMITATIONS**

As applicable, identify any of the following biases and limitations:

**Measurement biases**

When evaluating measurement biases inherent in the study, you can consider reporting on the outcome measures as a group, taking into account each outcome measure.

**Were the evaluators blind to treatment status?** Evaluators who are aware of which group a participant was allocated to or which intervention a participant received can influence the results by giving the participant, or group of participants, a more or less favorable evaluation. It is usually the intervention group that is favored. This should be considered when the evaluator was part of the research or intervention team.

**Was there recall or memory bias?** This can be a problem if outcomes were measured using self-report tools, surveys, or interviews that require the participant to recall past events. Often a person recalls fond or positive memories more than negative ones, which can favor the results of the study.

**Intervention biases**

**Contamination:** This occurs when participants of the control group inadvertently received the intervention; thus, the difference in outcomes between the two groups may be reduced. This favors the control group.

**Co-intervention:** Participants receiving another form of intervention at the same time as the study intervention can influence the results in either direction. For example, taking medication while receiving or not receiving intervention could favor the results for participants in either group. The reader must consider if the other, or additional, intervention could have positively or negatively influenced the results.
**Timing of intervention:** Consider whether the intervention was provided over a short or extended period of time. For example:

- If the intervention was provided over an extended period of time, aging can be considered a timing factor.
- In the case of children, maturation alone could be a factor in improvements seen.
- If the intervention was very short in duration, there may not have been sufficient time for a noticeable effect in the outcomes of interest. This would favor the control group.

**Site of intervention:** Where an intervention took place can influence the results. For example, an intervention program carried out in a person’s home may result in a higher level of satisfaction that favors the intervention group. The site of intervention should be consistent among all groups.

**Use of different therapists to provide intervention:** If different therapists were involved in providing the intervention(s) under study to the different groups of participants, the results could be influenced in one direction. For example, one therapist could be more motivating or positive than another; hence, the group that she worked with could demonstrate more favorable outcomes. Therapist involvement should be equal and consistent among all intervention groups.

**Baseline equality:** Intervention and control groups should be equal on important participant characteristics (e.g., age, gender, severity of disability) and outcome measures (dependent variables) at study outset. If not, differences should be dealt with appropriately in data analysis.

**Additional limitations**

**Was this study adequately powered (large enough to show a difference?)** While a \( p \)-value that is greater than .05 indicates that there is no statistically significant difference between groups, in some research studies no difference is noted because the sample size is small. In this situation, one surmises that if the groups had been larger, there would be a statistically significant difference.

**Were the analysis methods appropriate?** Did the authors justify/explain their choice of analysis methods? Do they appear to be appropriate for the study and the outcomes? You need to consider the following:

- The purpose of the study—Is it comparing two or more interventions or examining the correlation between different variables of interest? Different statistical tests are used for comparison and correlation.
- The outcomes—If there is only one outcome measured that compares two different interventions, a simple statistical test (e.g., a \( t \)-test) will probably be sufficient. However, with a larger number of outcomes involving different types of variables, more complex statistical methods, such as analysis of variance (ANOVA) or regression analysis, are usually required.

**Were statistics appropriately reported?** Refer to narrative in article as well as tables.

**Was participant dropout less than 20% in total sample and balanced between groups?** If dropout was greater than 20% and there was an unbalance between groups, attrition may be a concern. Consider reasons for attrition if reported by researchers.
CONCLUSION
The discussion section of the article should outline clear conclusions from the results. As you are developing this section, consider the following:

- Conclusions should be relevant and appropriate, given the study methods and results. For example, the investigators of a well-designed RCT study using sound outcome measures could state that the results are conclusive that intervention A is more effective than intervention B for the study population.
- Strong conclusions cannot always be made. For example, studies with methodological limitations or biases (e.g., lack of a control group, unreliable measures) make it difficult to "prove" or conclude that it was the intervention alone that influenced the outcome(s). In these situations, the authors may only conclude that the results demonstrated a difference in the specific outcomes measured in this study for the clients involved.
- The results may not be generalizable to other populations. Therefore, further study or research should be recommended.
- Determine if study results address the study objectives (or hypotheses) by means of presenting the study outcomes and corresponding statistical analyses.
- Please note that this section is not a duplicate from the clinical bottom line. Information about how the evidence can be used to inform and guide occupational therapy practice should be included in the clinical bottom line section of the CAP.


APPENDIX: STUDY DESIGN TYPES AND ISSUES

DESIGN ISSUES
There are many different types of research designs. The most common types in rehabilitation research are included. The essential features of the different types of study designs are outlined to assist in determining which were used in the study you are reviewing. Some of the advantages and disadvantages of the different types of designs are outlined to assist in determining the appropriateness of the design for the study being reported.

DESIGN TYPES
For purposes of the Evidence Exchange, CAPs are to be submitted on articles with Level I (limited to randomized controlled trials only), II, or III evidence. Only intervention studies will be considered. CAPs submitted on systematic reviews, assessment tools and measures, meta-analyses, descriptive studies, single-subject design, case reports, or expert opinions will not be accepted into the Evidence Exchange.

Level I
**Randomized Controlled Trial**

- Randomized controlled trial or randomized clinical trial (also referred to as Experimental or Type 1 study); RCTs also encompass other methods, such as crossover designs.

- The essential feature of an RCT is that a set of clients/subjects are identified and then randomly allocated (assigned) to two or more different intervention “groups.” One or more group of clients receives the intervention of interest (often a new intervention) and another group is the “control” group, which usually receives no intervention or standard practice. Random allocation to different intervention groups allows comparison of the client groups in terms of the outcomes of interest because randomization strongly increases the likelihood of similarity of clients in each group. Thus, the chance of another factor (known as a confounding variable or issue) influencing the outcomes is greatly reduced.

- The main disadvantage of RCTs is the expense involved, and in some situations it is not ethical to have “control” groups of clients who do not receive intervention. For example, if you were to study the effectiveness of a multidisciplinary inpatient program for post-surgical patients with chronic low back pain, it may be unethical to withhold intervention in order to have a control group.

- RCTs are often chosen when testing the effectiveness of an intervention, or to compare several different forms of intervention.

**Example:**
The effects of two different OT interventions (functional rehabilitation and reactivation) were evaluated using a randomized controlled trial. Forty-four patients of a long-term-care center were randomly allocated to one of the two types of intervention. Outcomes were measured using a variety of psychometric tests at 3 different points in time (Bach, Bach, Bohmer, Gruhwalk, & Grik, 1995).

**Level II**

**Nonrandomized Controlled Design (also called Cohort Design)**

A cohort is a group of people (clients) who share a common experience within a defined time period. Cohort designs are “prospective,” meaning that the direction of time is always forward. Time flows forward from the point at which the clients are identified. They are sometimes referred to as prospective studies.

- Given the nature of the cohort studies, these studies often have a comparison (“control”) group of clients/people who have not been exposed to the situation of interest (i.e., they have not received any intervention). One of the main differences between an RCT and a cohort study is that the allocation of people (clients) to the intervention and control groups is not under the control of the investigator in a cohort study. The investigator must work with the group of people who have been identified as “exposed” and then find another group of people who are similar in terms of age, gender, and other important factors.
It is difficult to know if the groups are similar in terms of all the important (confounding) factors, and therefore, the authors cannot be certain that the intervention (exposure) itself is responsible for the outcomes.

Participants → Exposure to intervention → Outcome

No exposure to intervention
Time-----------------Prospective------------------>

Example: Evaluation of a mental stimulation program used a cohort design to measure changes in mental status in 30 patients over a 2-month time period. The first 15 patients who were admitted to a day care center received intervention and composed the “exposed” group. The remaining 15 admissions did not receive intervention immediately and served as a “control” group (Koh et al., 1994).

**Case Control Design**

- Case control studies explore what make a group of individuals different. Other terms used are *case-comparison study* or *retrospective study*. Retrospective is the term used to describe how the methods look at an issue after it has happened. The essential feature of a case control study is looking backward.

- A set of clients/subjects with a defining characteristic or situation (e.g., a specific diagnosis or involvement in an intervention) is identified. The characteristic or situation of interest is compared with a control group of people who are similar in age, gender, and background but who do not have the characteristic or are not involved in the situation of interest. The purpose is to determine differences between these groups.

- It is a relatively inexpensive way to explore an issue, but there are many potential problems (flaws) that make it very difficult to conclude what factor(s) are responsible for the outcomes.

Participants with outcome → Assessment of previous exposure → Compare between of interest to intervention or casual factor groups

Participants without outcome of interest

>--------------------------Retrospective Measurement--------------------------Time

Example: If an occupational therapist wants to understand why some clients of a day care program attended an optional daily activity program on a regular basis while other clients did not attend, a case control design could be used to explore differences between the two groups of clients in relation to age, gender, interests, background, and current living situation.

**Level III**

**Nonrandomized**

**Before-After Design (also known as Pretest and Posttest)**

- Before-after design is usually used to evaluate a group of clients involved in an intervention.
- The evaluator collects information about the initial status of a group of clients in terms of the outcomes of interest and then collects information again about the outcomes after intervention is received.

- This is a useful design when you do not wish to withhold intervention from any clients. However, with no “control” group, it is impossible to judge if the intervention alone was responsible for any changes in the outcomes. Changes could be due to other factors such as disease progression, medication use, lifestyle, or environmental changes.

  Participants → Assessment → Intervention → Outcome
  Time-------------------Prospective--------------------->

Example: The level of caregiver strain following placement of an elderly family member with dementia in adult day care was evaluated using a before-after design. Outcomes of caregiver strain and burden of care were measured in 15 subjects before and after the day care placement (Graham, 1989).

**Cross-Sectional Design**

- This design involves one group of people, and evaluation of the whole group is carried out at the same time.

- This design is often used to explore what factors may have influenced a particular outcome in a group of people. It is useful when relatively little is known about an issue/outcome.

- Surveys, questionnaires, and interviews are common methods used in cross-sectional studies. They are relatively inexpensive and easy, as evaluation takes place at one point in time.

- It is impossible to know if all factors have been included in the evaluation, so it is difficult to draw cause-effect conclusions from the results beyond the group of people being studied.

  Participants → Measurement of outcomes and other factors at the same time
  Time: All done at one point in time

Example: Clients and their families who have been involved in a new activity program for seniors with dementia can be surveyed or interviewed upon discharge to evaluate the effect of the program on their quality of life, activity participation, and level of satisfaction.


For personal or educational use only. All other uses require permission from AOTA.

Contact: [www.copyright.com](http://www.copyright.com)