



Beer Judge Certification Program

5115 Excelsior Blvd, # 326
St. Louis Park, MN 55416

www.bjcp.org

BJCP Scoresheet Guide

I. Objective

The BJCP Scoresheet Guide was initially developed to serve as a rubric for grading the BJCP Beer Judging Exam. However, after assembling a comprehensive list of what constitutes a great score sheet, we realized that this document can also be used as a training tool for current and aspiring judges.

II. Background

The information in this document is intended for use by current judges, by new judges preparing for the BJCP Beer Judging Exam, and by BJCP Beer Judging Exam graders. It was drafted by the BJCP Exam Directorate in response to suggestions¹ for a more transparent and consistent process for grading BJCP Beer Judging Exams. The quantitative metrics were created after evaluation of many (more than 100) scoresheets completed by National, Master and Grand Master judges who proctored BJCP Beer Judging exams in 2012 and 2013. Additional information was incorporated after review by the BJCP Associate Exam Directors and by active BJCP exam graders. The BJCP Scoresheet Guide was created using a data-driven approach that should minimize any subjectivity (or perceived subjectivity) with the exam grading process. It should be noted that as with grading a written exam on any subject, there is generally a learning curve needed to make an accurate and expedient evaluation of the quality of a beer scoresheet. This document should make that learning curve less steep for graders and make the grading process more transparent to our BJCP membership. The BJCP Scoresheet Guide is intended to supplement the Exam Scoring Guide used by BJCP exam graders but should not be used as a recipe for finding faults when evaluating a scoresheet. In other words, the intent is not to incorporate the algorithm into a spreadsheet for graders to enumerate every positive and negative aspect of a score sheet, but to serve as a self-consistent check that the assigned score correlates with the judging level in which the score lies.

III. BJCP Scoresheet Components

The five components of beer scoresheets evaluated on the BJCP Beer Judging Exam are Scoring Accuracy, Perception, Descriptive Ability, Completeness and Communications, and Feedback. These competencies are discussed below, and guidelines are given for achieving proficiency in each. While these can be quantified to some extent, the scoresheet should be evaluated as a whole and not as the sum

¹ Special thanks to Andrew Luberto for drafting a document outlining his concerns with the grading process and providing an impetus to move forward with this guide

of its parts. It should also be noted that the criteria for a Master level scoresheet is 90% level rather than the 100% that denotes perfection, and that was taken into account in designing the grading rubric.

A. Scoring Accuracy

Scoring Accuracy is evaluated on the individual exam beers based on the differences between the total scores of the examinee and the consensus score from the proctors, according to the chart below. A variance of less than seven points is required to earn at least 60% of the possible scoring accuracy points for an exam beer, i.e. the expected scoring deviation between BJCP judges at a competition. The consensus score is typically the value assigned by the proctors, but this may be adjusted by the Exam Director if their consensus score is inconsistent with the average score of the exam participants and the beer descriptions supplied by the exam administrator. While scoring adjustments are not common due to the requirement that the proctors are National, Master or Grand Master judges, this system provides checks and balances to ensure that the baseline for Scoring Accuracy is not solely dependent on the input of the proctors.

1. Scoring Accuracy Rubric

Variance from Consensus	Points/Beer
0	20
1.5	19
2.5	18
3.5	17
4.5	16
5.5	15
6	14
6.5	13
7	12
8	11
9	10
10	9

B. Perception

Perception is arguably the most difficult competency to evaluate because it requires condensing data from several sources into a concise representation of the characteristics of each exam beer. These data sources include not only the proctors' scoresheets, but the scoresheets completed by the exam participants, the exam beer descriptions supplied by the exam administrator, and the BJCP style guidelines.

The process that works for many graders is to first create a preliminary list of primary and secondary characteristics from the proctors' scoresheets and the exam beer descriptions, and then reinforcing or de-emphasizing characteristics based on observations from the exam participants. A process that works for some graders for each exam beer is:

1. Review the BJCP Beer Style Guidelines for that beer style.
2. Review the information provided by the exam administrator on the background of the exam beers and make note of any property that might impact the judging. For example, was it a classic example of the style, was it doctored, or was it perhaps an aged sample?

3. Begin with the scoresheet of the more experienced proctor and write down the key descriptors that were noted for the malt, hop and other components of the aroma. Try to capture three or four primary and secondary characteristics that are at least low to moderate in intensity as well as anything that is out of place or missing in the style being judged.
4. Write down key descriptors that were noted for the appearance, particularly those which may not be appropriate in the style, such as haze or being too dark or light in color.
5. Write down the key descriptors that were noted for the malt, hop character, balance and other components of the flavor. Again, try to capture three or four primary and secondary characteristics that are at least low to moderate in intensity as well as anything that is out of place or missing in the style being judged.
6. Write down the key descriptors that were noted for the mouthfeel.
7. Read through the scoresheets of the other proctor(s) and circle any descriptors which are already captured on the list you created from the scoresheet from the lead proctor. Add any new descriptors that capture the perception of the other proctors(s).
8. Narrow down the list of descriptors from the proctors to a subset of six to eight consensus characteristics that capture the essence of the beer. These are the primary and secondary characteristics that are expected to be noted on exemplary participant scoresheets.
9. On the participant scoresheets, use a red or colored pen to circle any descriptors that are in agreement with the consensus ones assembled from the proctors scoresheets and underline or highlight the ones that diverge. This information is easily transferred to the list of characteristics correctly and incorrectly perceived on the Report to Participant form.
10. Score the perception accuracy according to the rubric below.

While grading the scoresheets, you may find that the examinees may have correctly identified a characteristic that was not noted by the proctors. For example, if the proctors did not perceive diacetyl in an exam beer, but over one-half of the examinees did, then the examinees who noted that flavor component should not be penalized. Note that a perception error can also be an error of omission if an attribute is not mentioned on the scoresheet.

Each exam beer results in a unique scenario, so it is impossible to outline every possibility, but with some practice, most graders are able to develop a process for distilling the beer profile down into a finite number of essential components which should be identified on a Master-level scoresheet. As with Scoring Accuracy, it is important for BJCP Beer Judge Exam participants and graders to recognize that this competency is not evaluated solely on the information supplied by the exam proctors. By the same token, BJCP Beer Judge Exam graders should be careful to not create a lengthy list of every characteristic independently noted by each proctor with the unrealistic expectation that all of them should be mentioned by the examinee. Judges and graders need to be cognizant that the same characteristic can be perceived differently by different tasters. For example, if one proctor perceives a grainy character, another proctor comments on husky notes, then an examinee who mentions either descriptor or even a closely related one such as grassiness should be credited with accurate perception for this characteristic. These flavors are closely related and are adjacent on the Beer Flavor Wheel (versions of this sensory training tool are available at <http://www.beerflavorwheel.com/> and <http://morebeer.com/brewingtechniques/library/backissues/issue5.6/beerflavorwheel.html>), so describing the same characteristic differently should not be regarded as a perception error.

There will be cases where the proctors may note an array of flavors such as “rich, toasty melanoidins with caramel notes” in a classic example of a Bock while the examinee may only note the presence of toasted malt. While the examinee should receive some credit for identifying one of the melanoidin components, not recording others which were present and in accordance with the style guidelines is both

a perception error by omission as well as an incomplete description. This situation often occurs when the examinee is not well-versed in the style being judged.

1. Perception Competencies

- a. Master: At least three components of the aroma described on the examinee's scoresheet are consistent with the consensus of the proctors and/or over one-half of the examinees.
- b. Master: There are no aroma components perceived at low or higher levels that are not included in the consensus of the proctors and/or over one-half of the examinees.
- c. Master: The descriptions of the color, clarity and head are consistent with the consensus of the proctors and/or over one-half of the examinees. It is also desirable to describe other characteristics of the head, but this is not essential information for the brewer.
- d. Master: At least three components of the flavor described on the scoresheet are consistent with the consensus of the proctors and/or over one-half of the examinees.
- e. Master: There are no flavor components perceived at low or higher levels that are not included in the consensus of the proctors and/or over one-half of the examinees.
- f. Master: At least three components of the mouthfeel described on the scoresheet are consistent with the consensus of the proctors and/or over one-half of the examinees.

2. Perception Rubric

When evaluating this competency, errors relating to secondary characteristics should be given 50% of the weight given to errors in perceiving primary characteristics.

- a. Master: No more than two elements of the aroma, appearance, flavor and mouthfeel differ from the consensus of the proctors and/or over one-half of the examinees.
- b. National: No more than three elements of the aroma, appearance, flavor and mouthfeel differ from the consensus of the proctors and/or over one-half of the examinees.
- c. Certified: No more than four elements of the aroma, appearance, flavor and mouthfeel differ from the consensus of the proctors and/or over one-half of the examinees.
- d. Recognized: No more than five elements of the aroma, appearance, flavor and mouthfeel differ from the consensus of the proctors and/or over one-half of the examinees.
- e. Apprentice: More than five elements of the aroma, appearance, flavor and mouthfeel differ from the consensus of the proctors and/or over one-half of the examinees.

C. Descriptive Ability

Descriptive Ability is a measure of the competency to use adjectives and phrases to identify the aroma, flavor and other characteristics of the beer. Beer judges should be able to describe "how much" as well as "what kind of" aroma and flavor components are present. Scoresheets from Master and Grand Master judges typically have a proliferation of descriptive information beyond just "malty" or "hoppy," and this is important because the map between the perceived characteristics and the BJCP guidelines for any style in part determines the stylistic accuracy and technical merits of the beer. As an example of adding layers or descriptive information, consider an American Pale Ale which an Apprentice judge might describe as having a "hoppy aroma." A Recognized judge might add a descriptive adjective, stating that it has an "American hop aroma," while a Certified judge may also note the intensity and write that it has a "moderate American hop aroma." National judges typically add an additional layer of descriptive information, such as "moderate citrusy American hop aroma," while a Master judges could take that a step further and write "moderate citrusy American hop aroma, with notes of tangerine and pine needles."

1. Descriptive Ability Competencies

- a. Master level: At least six descriptive adjectives² or phrases³ are used to describe the aroma of the beer.
- b. Master: At least four descriptive adjectives or phrases are used to describe the intensity of the aromatic components of the beer.
- c. Master level: At least four descriptive adjectives or phrases are used to describe the appearance of the beer.
- d. Master level: At least six descriptive adjectives or phrases are used to describe the flavor of the beer.
- e. Master level: At least four descriptive adjectives or phrases are used to describe the intensity of the flavor components of the beer.
- f. Master level: At least four descriptive adjectives or phrases are used to describe the mouthfeel of the beer.

2. Descriptive Ability Rubric

- a. There are a total of twenty eight (28) opportunities to use descriptive adjectives or phrases, and the targets for Master/National/Certified/Recognized judge (based on 90/80/70/60 percent of that total) are: Total number of adjectives and phrases: Master: 25 or more, National=22-24, Certified=19-21, Recognized=16-18, Apprentice=15 or fewer.

D. Feedback

The BJCP Exam Scoring Guide states that “the brewer should receive useful and constructive feedback explaining how to adjust the recipe or brewing procedure in order to produce a beer that is closer to style. The comments should be constructive and consistent with the characteristics perceived by the examinee as well as with the score assigned to the beer.” This is a great synopsis on the purpose and desired elements of feedback on a beer scoresheet. Note that feedback is not restricted to the Overall Impression section, so careful reading of the entire scoresheet is required to accurately assess this competency. The feedback should be specific in terms of how the process or recipe should be corrected rather than just pointing out that something needs to be addressed. Feedback also includes calling out when characteristics are permissible for the style, such as “low levels of yeast character” in an American Premium Lager.

1. Feedback Competencies

1. (1 point) The feedback is constructive, polite and professional, generally including at least one supportive or positive comment about the beer.
2. (1 point) The feedback is consistent with the score assigned to the beer.
3. (2 points) The judge provides an accurate diagnosis of any stylistic flaws which impacted the perceived beer quality and/or score. A good rule of thumb is incorporating one suggestion for improvement for each scoring level below Outstanding (45-50).
4. (2 points) The judge provides an accurate diagnosis of any technical flaws which impacted the perceived beer quality and/or score. A good rule of thumb is

² In this context, descriptive means adjectives more specific than generic words such as "malty," "hoppy," and "estery." Note that "No" or "None" are valid adjectives for characteristics which may be appropriate in the beer style being judged but are not present in that particular sample.

³ In this context, the comment, "A medium spicy noble hop flavor emerges mid-palate," has two descriptive adjectives for the hop flavor, one for the intensity and one descriptive phrase.

incorporating one suggestion for improvement for each scoring level below Outstanding (45-50).

5. (2 points) The feedback given to the brewer is accurate with respect to the characteristics perceived by the judge. This is independent of whether the perception was accurate.
6. (1 point) The feedback does not make any assumptions about the process or ingredients.
7. (1 point) The feedback given in the Overall Impression section is consistent with comments in other sections of the scoresheet.
8. (2 points) Observations of characteristics of the aroma, appearance, flavor and mouthfeel are noted in the appropriate section of the scoresheet, with feedback either given in that section or as part of the Overall Impression. For example, astringency should be discussed in the Mouthfeel section, not Flavor.

2. Feedback Rubric

- a. Number of above feedback elements on scoresheet: Master: 11-12, National: 10, Certified: 8-9, Recognized: 7, Apprentice: 6 or fewer.

E. Completeness and Communication

Completeness and Communication are measures of the ability of a judge to produce a high level of content in the scoresheet while clearly and effectively communicating information about the beer to the brewer. This is, in theory, the easiest category in which to achieve a high score on the Beer Judging Exam, but it does require some attention to detail. A complete master-level scoresheet not only requires addressing the components of the aroma, appearance, flavor and mouthfeel that are appropriate in a given beer style, but includes book-keeping details such as identifying the beer style, the beer or entry number, the participant number or name of the judge, filling in the scores for the beer, and marking the appropriate checkboxes at the bottom and left side of the scoresheet.

1. Completeness and Communication Competencies

1. (2 points) All applicable components⁴ of the aroma listed on the scoresheet are addressed.
2. (2 points) All applicable components of the appearance listed on the scoresheet are addressed.
3. (2 points) All applicable components of the flavor listed on the scoresheet are addressed.
4. (2 points) All applicable components of the mouthfeel listed on the scoresheet are addressed.
5. (2 points) Overall Impression section includes a comment on overall drinking pleasure associated with entry (1 point) and if the total score is less than 45, offers at least one suggestion for improvement (1 point).
6. (4 points) Efficient use of vertical space: For perfect score, fewer than two blank lines remain on the completed scoresheet (these are typically only in the Appearance or Mouthfeel sections on a Master level scoresheet). Deduct 0.5 point for every blank line beyond 2, up to a maximum of 4 points deducted. For example, 5 blank lines would be a $(5-2)/2 = 1.5$ point deduction.

⁴ In this context, the phrase "applicable components" means characteristics which are expected in the beer style being judged according to the BJCP Style Guidelines and/or to perceived characteristics which would be regarded as flaws in that particular style.

7. (2 points) Numerical values are assigned for all component scores (1 point) and also for the total score (1 point).
8. (1 point) The stylistic accuracy, technical merit, and intangibles boxes are checked. No partial credit here.
9. (1 point) Descriptor definitions are checked when applicable (flavors are either perceived at low or higher levels or are flaws in the style being judged). Partial credit may be awarded.
10. (1 point) Comments are well organized and legible.
11. (1 point) Efficient use of horizontal space: The scoresheet typically has six to seven words per line with a font size and spacing that balances content and legibility. The objective here is to discourage judges from writing in an extremely large font to fill up the space on the scoresheet without conveying much information.

2. Completeness and Communication Rubric

- a. There are a total of twenty points available on a complete scoresheet. The targets for Master/National/Certified/Recognized judge (approximately 90/80/70/60 percent are: Master: 18 or more, National=16-17, Certified=14-15, Recognized=12-13, Apprentice=fewer than 12.

IV. Summary

This BJCP Scoresheet Guide is intended to provide guidance for the creation of high quality beer scoresheets exemplified by those produced by most Master and Grand Master judges. As with other BJCP documents, it is not intended to be a stand-alone document, but should supplement other material on the BJCP website related to beer styles, beer judging and exam grading. If you have any comments or suggestions, please communicate them to the BJCP Exam Directors at exam_director@bjcp.org.