



Beer Judge Certification Program

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

www.bjcp.org

2015 EXAM SCORING GUIDE

Revised April, 2015

INTRODUCTION

One of the primary purposes of the BJCP is to recognize beer tasting and evaluation skills. The BJCP exam is the heart of the program and is a unique process for peer review of prospective judges and continuing education for both the examinees and the graders. Prior to 2012, the BJCP exam had a combined essay and tasting format, with 70% and 30% weights, respectively. The score on this exam, along with experience points, determined the maximum rank attainable for each judge.

In 2012, the BJCP moved to a three tier exam system consisting of an online entrance exam, a judging exam and a written proficiency exam. The entrance exam, when passed, qualifies a participant to take a judging exam during which six beers are evaluated as if one were at a competition. This exam qualifies a judge for the Apprentice, Recognized or Certified levels, depending on the score and the number of experience points. Those judges seeking to advance to the National and Master levels must still take an written exam, and this option is only available to those who have scored at least 80% on the judging exam and amassed at least 10 judging points. The scores of the judging and essay exam in this new system are combined with equal 50% weights to give a comprehensive exam score.

The majority of the exams that the BJCP graders score are the judging exams taken by new prospective judges, and our goal is to evaluate these exams as objectively as possible. This is admittedly difficult in an area as complex as beer evaluation, but most graders are able to accomplish this task through a combination of experience and mentoring by senior graders.

The grader's priority is to determine the proper level of the judge, i.e. fail, 60, 70, 80 or 90. The second priority is the position within the level, i.e. 70 or 75. Remember than a 70 may turn a person off from the program, while a 75 may inspire further study and achievement. Graders should avoid assigning scores that end in 9 since that can make examinees feel they should argue for an additional point to move up to the next level. As in judging a beer, one should be careful about becoming a fault finder. Higher rankings should be attainable, and the grader should be careful about being too critical. In particular, they should be careful about relying on a "bottom-up" approach since this generally leads to underscoring.

EXAM GRADING GUIDELINES

In scoring a test, the scorer should be comfortable that the examinee has demonstrated skills that relate to the judge level for which the score qualifies, using the following:

<60: On the written proficiency exam, little knowledge of brewing and/or styles is conveyed, and major gaps are evident. On the judging exam, the examinee displays weak tasting skills,

and the scoresheets will generally have unacceptably low levels of completeness, descriptive information and/or feedback. This examinee will be an Apprentice judge.

60s: The examinee demonstrates a basic grasp of fundamentals on the written proficiency exam, but there may be some significant knowledge gaps. The judging exam demonstrates the minimum acceptable communication and judging skills expected of a Recognized judge.

70s: There can be errors and small gaps in the answers on the written proficiency exam, but depth in answers is not necessary. On the judging exam, at least three of the six exam beers are accurately evaluated. The scoresheets should have reasonably good completeness, descriptive information and feedback, appropriate to the Certified judging level.

80s: The written proficiency exam indicates good knowledge of all subjects. Some errors are allowable, but there are no significant gaps and most of the answers demonstrate depth. On the judging exam, at least four of the six exam beers are accurately evaluated with the high quality scoresheets expected of a National judge.

90s: The written proficiency exam demonstrates excellent knowledge level. There are no significant errors, no knowledge gaps, good depth to answers, and evidence of independent thought. On the judging exam, it should be obvious that the examinee is an experienced beer taster. At least five of the six exam beers are accurately evaluated and the scoresheets have Master levels of completeness, descriptive information and feedback.

SCORING MECHANICS

JUDGING EXAM: There are six beers which are evaluated in a 90 minute time period. Each beer is scored on a 100 point scale, with 20 points allocated to scoring accuracy, as follows:

1) **SCORING ACCURACY:** (20 points/beer) The judges' scores and the consensus scores of the proctors for each beer are entered by the exam director into the Exam Grading Form (EGF). The scoring accuracy is calculated using the following variance table:

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

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. Scoring by the proctors is sometimes variable, so consensus scores may be

adjusted if there are biases that adversely affect most of the examinees. For example, suppose that the majority of the participants observe a flaw such as diacetyl that was not noted by the proctors, and as a result, they give the beer a lower average score. Without tasting the beer, the graders cannot determine if the flaw was actually present, but they should check the exam administrator's comments on the exam beer for more clues. In some cases, the proctors' consensus scores may need a small adjustment to more accurately reflect the quality of the beer. Comparing the average participant scores with the proctors' scores is another way to evaluate the accuracy of the consensus scores. *These adjustments to the consensus score are rare and should only be done with the approval of the Exam Director and/or Associate Director.*

2) **SCORESHEET COMMENTS** (80 points/beer):

The remaining points per beer are equally divided between Perception, Descriptive Ability, Feedback and Completeness. In 2014, the BJCP exam directorship created the [BJCP Scoresheet Guide](#), which is essentially a detailed rubric to help both judges and graders understand the criteria that determine the quality of a beer scoresheet. Please review that document prior to beginning the grading assignment.

The 80 points for the evaluation of each beer are assigned as follows:

- a) Perception (20 points/beer): Points should be deducted for missed flaws and errors in aroma, appearance, flavor, and mouthfeel perception. The rubric formed by the proctors' scoresheets enables the graders to make a correlation between the characteristics identified by the examinees and those noted by the proctors.
- b) Descriptive Ability (20 points/beer): A beer judge should be able to describe the intensity and characteristics of the aroma, appearance, flavor and mouthfeel using the proper terminology. The BJCP Style Guidelines serve as somewhat of an answer key for this aspect of the scoresheet.
- c) Feedback (20 points/beer): 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.
- d) Completeness/Communication (20 points/beer): A complete scoresheet should have well-organized, legible and have informative comments that fill all of the available space. The checkboxes for stylistic accuracy, technical merit and intangibles should also be marked. This aspect of the scoresheet is generally consistent with the level of descriptive information and feedback conveyed by the examinee.

The points awarded for each aspect of the beer should be correlated with the experience levels; i.e., 12-13 would be expected from a Recognized judge, 14-15 from a Certified judge, 16-17 from a National judge and 18-20 from a Master judge. Scoresheets which are indicative of a subpar judging performance generally fall in the 9-11 point range. Record the score for each beer on the EGF.

WRITTEN EXAM: There are six questions to be answered in 90 minutes, with the first question comprising twenty true-false (TF) questions on the BJCP levels, the judging process and judging ethics. These TF questions only impact the exam score if they are answered incorrectly, in which case a one half point (0.5) deduction is made for each error or omission. The other five questions are essay format, worth 100 points each, and covering beer styles, beer characteristics and the brewing process. The primary reference for grading any aspects of beer styles is the 2008 BJCP Style Guidelines. Before grading the written proficiency exam, read each question, review relevant

references and make a checklist of the key information that should be included in a complete answer. When evaluating the answer provided by the examinee, consider the accuracy, the depth of knowledge demonstrated, the completeness and communication skills, including neatness and organization. Understanding positions of various authorities on controversial subjects is desirable, as is knowledge of commercial and classic examples of the styles. Omissions and incorrect or contradictory information should detract from score; however some of these deductions may be compensated by inclusion of greater depth in other aspects of the description.

The points awarded for each answer should be correlated with the experience levels; i.e., 60-70 points would be expected from a Recognized judge, 70-80 from a Certified judge, 80-90 from a National judge and 90-100 from a Master judge. Answers which are indicative of a subpar judging performance generally fall in the 40-60 point range. The score for each answer should be entered on the EGF, and comments for the feedback portion of the Report to Participant Form noted.

REPORT TO PARTICIPANT (RTP) FORMS

The lead scorer is responsible for completing the RTP that will be returned to each examinee. The format for the RTP consists of a cover page summarizing and explaining the results followed by additional pages giving tabular feedback on specific beers (Beer Judging Exam) and essay questions (Written Proficiency Exam). The RTP form is setup so that the header for the second page and any possible additional pages will have the participant number as part of the page header. The sections of the RTP should be completed as follows:

SCORES: This section will be completed by the Exam Director after the exams have been reviewed, so please leave this section blank.

RECOMMENDED STUDY: Indicate which references should be read or reviewed to correct deficiencies on the written proficiency and tasting portions.

JUDGING EXAM SUMMARIES: The EGF tabulates the scoring accuracy result and the average scores for each aspect of beer evaluation: perception, descriptive ability, completeness and feedback. These generate tables for each exam in the tabs named “Grids 1-12” and “Grids 13-24.” These tables are to be copied and pasted into the RTP, and then additional feedback can be given using the checkboxes or in prose format. However this additional feedback is optional and should be kept brief since the summary tables already provide detailed information about the performance on each scoresheet.

WRITTEN EXAM SUMMARIES: Three questions on the BJCP written proficiency exam focus on beer styles, while two are more technical in nature. There are also TF questions relating to the levels of the BJCP, the judging process and ethics. The correct answers for the TF questions and those submitted by the examinees have already been entered into the EGF by the Exam Director, so these do not need to be graded. The EGF also calculates the average scores for the style and technical questions on the exam, and these generate tables that are to be copied and pasted into the RTP.

SCORING CONSENSUS

The exam graders reconcile their scores by e-mail and agree on a final result. Both graders should be comfortable with the location of this consensus score not only with respect to the judging levels, but also within a given level, i.e. low, mid or high end of the range. If this is not the case, one or

both of the graders should adjust his/her score. If the deviations are more than seven points, then the graders should discuss the exam in detail and adjust their scores until they reach a consensus. If there is still a problem, request further scoring by the Associate Director to break the deadlock or to determine what final scores should be assigned to borderline exams. **When a consensus score has been reached for each exam, the lead grader should also e-mail the completed EGF (including scores from the second grader), the consensus scores and completed RTPs to the Associate and Exam Director.**

ADDITIONAL GRADING TIPS

1. Take advantage of the statistics when scores are entered into the EGF. For example, on a Beer Judging exam, the two graders and six beers generate twelve data points for each of the four scoresheet characteristics (Perception, Descriptive Ability, Feedback and Completeness). Small differences between the graders will be averaged out. For example, if the graders arrive at scores of 15 (75%) and 17 (85%) for the Perception on one of the exam beers, this corresponds to $2 \times 20\% \times (1/6) = 0.07$ points of the total score for that examinee. This small difference should not warrant much debate or discussion.
2. When scores from both graders have been entered into the EGF, first compare the scores in the “Summary” tab of the EGF. The exams for which the two scores are within five points and within the same judging level, then the consensus score should be the average score – even if there are variations in the scores for the individual beers or essay questions. Let the statistics smooth out these differences. The scoring part of the assignment is finished for these exams, and the lead grader can work on the corresponding RTPs. For exams where the total scores diverge, look at the scores for the individual beers or essay questions and see if you can pinpoint the source of the difference. It may be a matter of discussing the rubric and making sure you are on the same page. The next layer of investigation would be to look at the components of each exam beer or essay question, but this is often unnecessary after the graders identify which exams need further review.
3. If there is ever an impasse in assigning a consensus score after each grader has taken another pass through the exams on which there are scoring divergences, remember that the Associate Director is available to provide a third opinion. There are sometimes systematic biases between the graders, and the AD will help identify if one or both of the graders is being too harsh or lenient. Note that the AD will also automatically review exams that end up close to the threshold between judging levels.
4. If there are circumstances in your professional or personal life that will result in a significant delay in completion of the grading assignment, please communicate that information to the AD and ED as soon as possible. It is much better to reassign the exams early in the process rather than to let them languish for six months or longer. Communication is the key!

TIMETABLE and EXPENSES

Our target is to turn tasting exams around in eight weeks and written proficiency exams in twelve weeks. This requires that graders complete the scoring in no more than four and six weeks, respectively. The BJCP is a nonprofit organization, so the graders and the directors are not expected to profit from the grading of exams. Reasonable expenses may be tabulated and submitted to the BJCP treasurer with receipts for reimbursement; however this is rare.