

SHAREWORKS

BUDDYCHECK

Calculations - Adjustment Factor

Made by: Lorenzo Presilli
4th October 2021

Shareworks Solutions BV
Schiekade 105
3033 BH Rotterdam
The Netherlands

Contact person: Flip van Haaren
Email: flip@shareworks.nl

Abstract

Determining students' performance in group projects is necessary since members can contribute in different degrees towards the final group result. To take the contribution of each member into account, Buddycheck uses a peer-to-peer evaluation system based on the Catme peer evaluation method [2]. The evaluations are used to better understand the dynamics within a group; this can be used by the instructors of these groups to adjust the individual grades.

The calculation methods used by Buddycheck are divided into several sections. First, there are standard calculation methods such as calculating the average and the mean of student results. There is the calculation of the adjustment factor. Instructors can customize the outcome of the adjustment factor by choosing to use with or without self-scores. Instructors can also use two methods for calculating the group average. Buddycheck also determines the reliability of evaluations where some surveys are not filled in, where the total number of people who filled in the survey, and how much their given scores differ both impact this reliability score.

Based on the Buddycheck calculations, feedback labels are added to individuals. A student can be a high performer, low performer, overconfident, underconfident, manipulator, or in conflict. When multiple students receive evaluation scores that significantly differ depending on who gave the scores, this will result in the group being labeled as 'clique'. Finally, this document will show what object is returned to the front-end after the calculations are finished, the content of which is used to provide peers and instructors with clear reports of the group performance and feedback labels.

List of symbols

Abbreviations

Abbreviation	Description
Avg	Average
d	Length of a scale
FP	Fixed point on a scale, fraction between 0 and 1
max	Maximum number on a scale
min	Minimum number on a scale
n	Amount/length of array
N	Number
R	Reliability
SS	Sum of squares
t	t-score
T	Threshold value
δP	difference from the mean in points
μ	average
σ	standard deviation

Subscript

Abbreviation	Description
<i>2dec</i>	Two decimals
<i>scale,i</i>	Context is a scale of dimension 'i' (5,7,9...)

Contents

Introduction	1
Mapping	2
Basic Methods	4
Two decimals	4
Sum	4
Average	4
Median	5
Buddycheck Methods	6
With and without self	6
Group average calculation method	6
Scaling	7
Deviations	8
Adjustment factor	9
Reliability	10
Feedback labels	12
SKEW	12
CONF	13
UNDER	13
OVER	13
HIGH	14
LOW	14
TENS	14
BYPASS	15
Results	16
Epilogue	17

Introduction

The flow of calculations for Buddycheck evaluations is displayed in Figure 1. First, a map is created that contains all the relevant empty arrays, such as what students received or what they gave their peers. Then the scores are added to this map based on the surveys. The reliability is calculated with these scores. After determining whether the evaluation is reliable or not, averages are calculated, and the group standard deviation will be determined based on the scores and the averages. The adjustment factor is determined next. Finally, the group and each peer's feedback label are determined based on the averages, deviations, and scores obtained from previous steps. An object is returned to the front end that contains all the relevant mapped scores, averages, Buddycheck parameters (factors), and feedback labels.

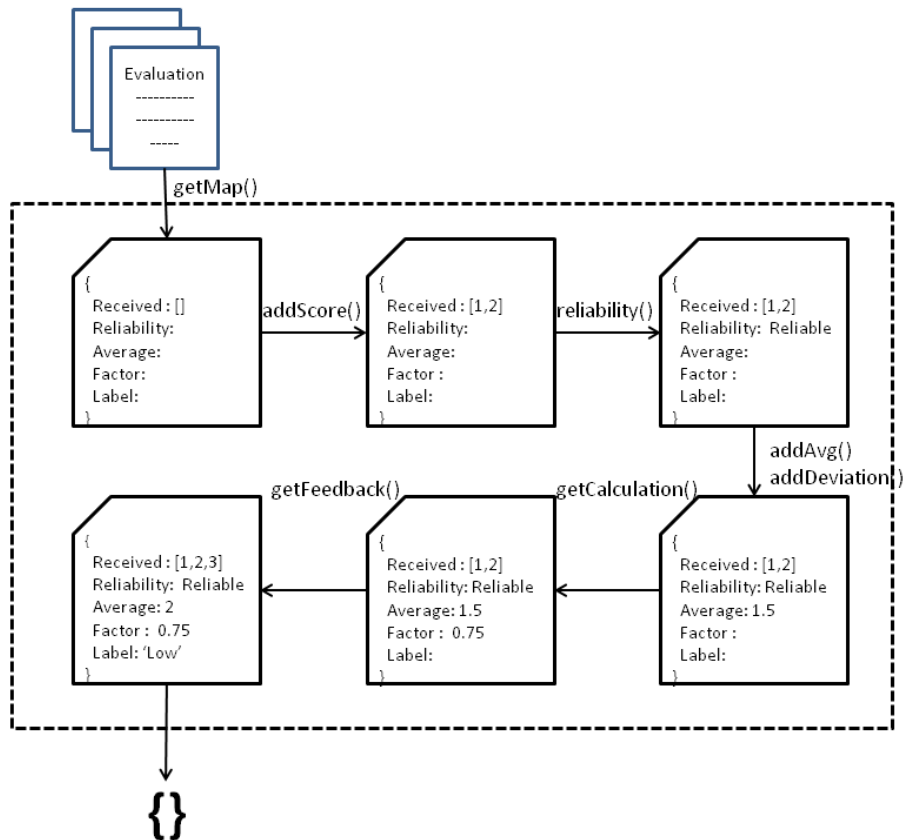


Figure 1: The flow of calculations for Buddycheck evaluations. Arrows represent the functions that are called, the images represent the objects that the functions operate on..

Mapping

Before discussing the calculation methods of Buddycheck, it is essential to understand what data the evaluations contain and know the structure of how data is stored. An evaluation has to be filled in by peers that work in a group. Evaluations can include a custom amount of questions, which need to be filled in by a peer for each of her peer members and herself. When the peer fills in the evaluation, this is called a survey. A survey for a peer could look as follows:

peer	Teamwork	Contribution	Responsibility
John	2	3	2
Anna	3	2	2
Klaas	4	4	5

Table 1: Example of a filled in survey

When all members hand in the surveys, the obtained results for the group as a whole is then obtained as shown in example table 2

	Teamwork			Contribution			Responsibility		
	John	Anna	Klaas	John	Anna	Klaas	John	Anna	Klaas
John	2	3	3	3	3	2	2	3	3
Anna	3	2	2	2	2	3	2	3	2
Klaas	4	4	5	4	4	5	5	3	4

Table 2: Example of a filled in evaluation, the thick numbers represent the scores given by John from the survey in table 1

From the filled-in evaluation, data is extracted and stored in the form of a map to make it easier to make calculations later on. An example of a map that stores arrays for a peer member is shown in the following code snippet; note that not the entire map used in Buddycheck is shown here since it includes many arrays. However, the example shown here is, in essence, the same for all the other mapped data.

```
1 map[Peer] = {  
2     scoresGivenToSelf: [],  
3     scoresReceivedWithSelf: [],  
4     scoresReceivedWithoutSelf: [],  
5 };
```

The map can be filled with scores from the evaluation results. To give an example for John from table 2, the scores given to self are obtained by adding the scores where the row of 'John' intersects with the columns of 'John'. The received scores without self are filled by taking the scores in the row of 'John' but excluding the cells that intersect with the column of 'John'. The filled-in map for John is shown in the snippet below:

```

1 map[John] = {
2     scoresGivenToSelf: [2,3,2],
3     scoresReceivedWithSelf: [2,3,3,3,3,2,2,3,3],
4     scoresReceivedWithoutSelf: [3,3,3,2,3,3],
5 };

```

Mapping is used on three levels, the group as a whole, for each peer, and for each question. The scores of a question with self can, for example, be obtained by adding all scores in the column under a question. Without self would be obtained by adding all scores under the column of a question, but excluding the diagonal where the row and column of a peer intersect. Mapping enables to preselect sets of data to make it easier and more clear to do calculations. If we would now, for example, determine the average of John without self scores, we simply take the sum of the array `scoresReceivedWithoutSelf` and divide it by the length of the array:

$$Avg_{John} = \frac{\sum scoresReceivedWithoutSelf}{scoresReceivedWithoutSelf.length} = \frac{\sum [3, 3, 3, 2, 3, 3]}{6} = 2.83 \quad (1)$$

Basic Methods

This chapter will go over the basic methods which are used throughout all the calculations performed in Buddycheck.

Two decimals

The function `twoDecimals()` Takes a number as input and turns it into two decimals by using the `Math.Round()` function in JavaScript. Where N is a number and N_{2dec} is a number with two decimals. `twoDecimals()` is used after all calculations are completed and before the results are sent to the front end. The moment that results are rounded is of importance since there should be no rounding in between calculations.

```
1 return (Math.round(number * 1e2) / 1e2);
```

Sum

The function `sum()` simply adds all numbers in a series.

$$sum = \sum_i (N_i) \quad (2)$$

```
1 let sum = 0;
2 for (let i = 0; i < series.length; i++) {sum += series[i];}
3 return sum;
```

Average

The average Avg over a series of numbers is calculated by taking the sum of these numbers and dividing the sum by the amount of numbers 'n'.

$$Avg = \frac{\sum_i^n N_i}{n} \quad (3)$$

```
1 return sum(series) / series.length;
```

Median

The `median()` function takes the number in the middle of a sorted series of numbers. The first step is to sort the series, as can be seen in line 1 of the code snippet below. Then the index position in the middle of the sorted series is found in line 2. If the sorted series is uneven in length, the number in the middle is simply returned (line 3). If it is even in length, it returns the average between the two numbers in the middle of the series.

```
1 sorted = series.sort(function (a, b) { return a - b; });
2 const half = Math.floor(sorted.length / 2);
3 if (sorted.length % 2) return sorted[half];
4 else return (sorted[half - 1] + sorted[half]) / 2.0;
```

Buddycheck Methods

The Buddycheck calculation methods contain the calculation of the adjustment factor. The adjustment factor is a fraction obtained by dividing the peer average by the group average. The average is determined by using the received scores of the peers in the evaluation. Whether the self-given scores of peers are taken into account or not is an option that the instructor can enable to customize the outcome. The two methods to calculate the group average also have an impact on the outcome of the adjustment factor and can be chosen by the instructor. The deviation inside the group is determined based on the received scores and averages, which will be used for determining if the group is a 'tension'. The other feedback labels are calculated with the mapped scores and averages as well but will be discussed in more detail in the following chapter. The used scale of the survey (standard from 1 to 5) has to be taken into account when calculating the adjustment factor and feedback labels. This chapter will look into more detail to using with and without self scores, average calculation method, scaling, calculating deviations, and finally, how the adjustment factor is calculated.

With and without self

The withself scores option influences the way how the peer averages are calculated. Withself directly influences the mapped group scores, the mapped peer scores, the mapped peer scores per question, and the mapped group scores per question. Consequently, the averages for these mapped scores are also dependent on the withself option. The adjustment factor is calculated by comparing the peer average with the group average. Therefore the withSelf option influences the outcome of the adjustment factor. Below is a code snippet that shows how each survey's 'score' is added to the map when withSelf is enabled.

```
1  if (recipientId !== respondentId) {map.scores.push(score);}
2  if (recipientId === respondentId) {if (withSelf) map.scores.push(score);}
```

Average calculation method

The calculation method for the group average is of importance when a peer handed in no survey. This situation is shown in table 3, where Anna did not give scores to her peers. Buddycheck supplies two methods for calculating averages to deal with evaluations where surveys are missing, namely the peer based average and the score based average.

	Teamwork			Contribution			Responsibility		
	John	Anna	Klaas	John	Anna	Klaas	John	Anna	Klaas
John	2		3	3		2	2		3
Anna	3		2	2		3	2		2
Klaas	4		5	4		5	5		4

Table 3: Example of an evaluation where a peer did not hand in his survey.

Peer based average

The peer based average calculation model first calculates the average of each peer with the available scores. For example, the score for John without self becomes as follows:

$$Avg_{John} = \frac{\sum[3, 2, 3]}{3} = 2.67 \quad (4)$$

We can do the same for Anna and Klaas, which received a 2.33 and 4.33, respectively. Now we can sum up all their averages and determine the group average as follows:

$$Avg_{group} = \frac{\sum[2.67, 2.33, 4.33]}{3} = 3.11 \quad (5)$$

The factor can be determined by dividing each peer's average with the group average, and the average adjustment factor will be 1. However, the peer 'Anna' that did not fill in her survey received more scores than John and Klaas. Meaning that when John handed in his survey, the 4 he gave for teamwork to Klaas has a much greater impact on the group average than the 3 he gave to Anna. Note that this problem does not arise when using with self scores since then each peer has an equal amount of scores. To sum up the advantage and disadvantage of a peer based average: the average adjustment factor will be 1, but certain scores that peers give can disproportionately influence the group average.

Score based average

The score based average calculates the group average by adding all the scores in an evaluation. In case of using without self and the scores of table 3, this would result in a group average of:

$$Avg_{group} = \frac{\sum[3, 2, 3, 3, 2, 2, 3, 2, 2, 4, 4, 5]}{12} = 2.92 \quad (6)$$

Notice that the group average is different from the one calculated in equation 5 from the peer based average. When using with self scores, the outcome of a score based average would be the same as that of a peer based average. To sum up the advantage of a score based average: Each given score is equally important for obtaining the group average. However, the average adjustment factor no longer has to be equal to 1.

As can be seen by comparing the advantages and disadvantages, each method is a double-edged sword. The ideal solution would be to have all the surveys filled in. When this is not the case, the school can consider what is most important in its evaluations and choose between these two methods.

Scaling

The standard scale used in calculating the adjustment factor is from 1 to 5. This scale is based on the Catme peer evaluation method [2]. When instructors require another scale, the criteria that Buddycheck has is that the adjustment factor is calculated in the same way. Therefore, different scales are converted back to a scale between 1 and 5. To convert the scale, the length of scale i ' d_i ' is determined by subtracting the start point from the endpoint of the scale as shown in equation 7. Any number ' N ' that needs to be scaled back to a scale between 1 and 5 is then converted according to equation 8. Here the max_{scale} and min_{scale} represent the beginning and end of the scale.

$$d_i = max_{scale,i} - min_{scale,i} \quad (7)$$

$$N_{scale,5} = (N_{scale,i} - min_{scale,i}) \cdot \frac{d_5}{d_i} + 1 \quad (8)$$

```

1  const scaleDistance = scaleMax - scaleMin;
2  AvgScale5 = (map.Avg - scaleMin) * (4 / scaleDistance) + 1

```

Different scales are converted because although the relative position on a scale of a peer and group average can remain the same, the starting score of 1 has different impacts on different sized scales. When a peer would receive 5 out of 5, and the average is 3, her factor will be 1.66. However, if she received 9 out of 9 and the average is 5, her factor would be 1.80. The difference is caused by the fact that the start of the scale at 1 has a smaller impact on larger scales. Thus extreme differences in the average score result in more extreme outcomes on a bigger scale than on smaller scales. Finally, the scaling is also used for feedback labels since the criteria should always be at the same relative position on any scale, as shown in Figure 2.

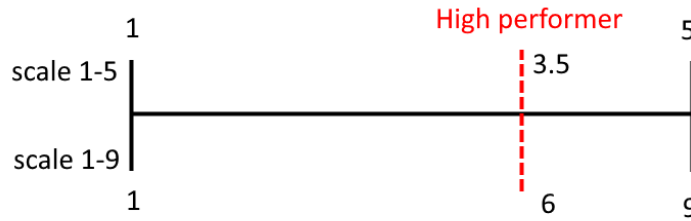


Figure 2: Scaling is used to calculate feedback label criteria.

Using the figure above as example, the high performer is at a fixed position $FP = 0.625$ on the scale. When the scale goes from 1 to 5, the criteria is met above a score of 3.5. The equation for calculating the criteria score $N_{criteria}$ is shown in equation 9. The impact of scaling on the criteria is also shown in the code snippet below.

$$N_{criteria} = d_i * FP + min_{scale,i} \quad (9)$$

```

1  if (score >= (d*FP + MinScale)) result.indvFeedback = 'Label'

```

Deviations

The deviation is determined based on the received score of each peer per question with respect to the average of each question. The formula of the standard deviation is displayed in equation 10. After all deviations for all peer members and each question are obtained, the group deviation is determined by taking the sum of all deviations and dividing it by the number of deviations. This is also according to the Catme peer evaluation method [3]. The deviation is then used to determine whether the group is labeled as a tension or not.

$$\sigma = \sqrt{\frac{\sum_i (N_i - \mu)^2}{n - 1}} \quad (10)$$

The group deviation is calculated by summing all the deviations for each questions 'i' and peers 'j' combination and dividing it by the product of the number of questions and peers.

$$\sigma_{group} = \frac{\sum_j \sum_i (\sigma_{i,j})}{n_i \cdot n_j} \quad (11)$$

Adjustment factor

The calculation for the adjustment factor is very straightforward, as can be seen in equation 12. It is determined by dividing the peer average by the group average on a scale from 1 to 5. It can be chosen to keep the original factor or cap it at a minimum and maximum threshold that the instructor can set. Furthermore, scores between 0.95 and 1 are rounded up to 1. Furthermore, a weight can also be given by the instructor to adjust further the impact of the 'adjustment' on the final grade as shown in equation 13. The weight will adjust the difference between the group grade and the peer grade created with the adjustment factor.

$$factor = \frac{Avg_{peer}}{Avg_{group}} \quad (12)$$

$$grade_{peer} = (1 - (1 - factor) \cdot weight) \cdot grade_{group} \quad (13)$$

The code snippet below shows the factor calculation, the maxMean and minMean are the thresholds given by the instructor.

```
1  const mean = (AvgPeer / AvgGroup);  
2  const factor = mean > maxMean ? maxMean : mean > 0.95 && mean < 1 ? 1 : mean;  
3  return minMean > factor ? minMean : factor;
```

Reliability

The reliability of an evaluation is determined to provide the instructors more insight into the reliability of the feedback labels and results. The Buddycheck reliability model should satisfy the following criteria:

- The deviation of received scores should be taken into account when determining the reliability. When the received scores have a small deviation from the average, it is more reliable than when the received scores differ widely.
- A minimum of two peers should fill in an evaluation. One peer is statistically meaningless.
- Bigger groups are more reliable than smaller groups with the same percentage of handed in surveys because each peer received more scores, making the standard deviation more reliable.

Criteria 1: standard deviation

To take the first criteria into account, it is assumed that received peer scores behave like a normal distribution. The deviation is determined per question per peer, and the group deviation is calculated by taking the sum of deviations divided by the number of deviations. This is analogue to the deviation equations 10 and 11 from the previous chapter. The criteria for reliability can be chosen as a percentage of the confidence interval where the deviation should fall in. The equation for the reliability score can be seen in equation 14. Where R is the reliability, T is the threshold value determined with the confidence interval, and σ_{group} is the group standard deviation. If the confidence interval is set to 80%, this corresponds to an allowed standard deviation ' σ_T ' of 1.28. The standard deviation is evaluated at a difference ' δP ' of one point from the average according to equation 15, meaning that a standard deviation of 1.28 would also correspond to a deviation of the average grade of 1.28 point.

$$R = \frac{T - \sigma_{group}}{T} \quad (14)$$

$$T = \sigma_T \cdot \delta P \quad (15)$$

The reliability score ' R ' is a number between 1 and -1, where 1 can be achieved if there is no deviation, and a score of 0 is achieved if the deviation is equal to the threshold value. Buddycheck defined three regions based on the reliability score. Between 0,2 and 1, the score is labeled as reliable, between -1 and -0,2, it is labeled as unreliable, between -0.2 and 0.2, the evaluation is labeled as questionable.

```
1 R = (T - sigma)/T;  
2 if (R < -0.2) reliability = 'Unreliable'  
3 if (R >= -0.2 && R < 0.2) reliability = 'Questionable'  
4 if (R >= 0.2) reliability = 'Reliable'
```

Criteria 2: minimum of two surveys

To consider the second criteria, a boundary condition is added to the reliability model that states that if the number of people filled in the survey is less than 2, the reliability score becomes -1, which is unreliable.

```
1 if (sampleSize < 2) reliabilityScore = -1;
```

Criteria 3: group size and missing surveys

Finally, the third criterion is implemented to consider both the group size and the number of people that filled in the survey. Therefore the t-distribution is used. The t-distribution converges to the normal distribution when the number of peers that fill in an evaluation is large (bigger than 30), and it is assumed that the calculated standard deviation is correct. However, when the group size is small, there is uncertainty about the actual standard deviation. The t-distribution has a property that the tails become thicker at smaller sample sizes, which takes into account the uncertainty of the standard deviation. The t-distribution compared to the normal distribution can be seen in Figure 3.

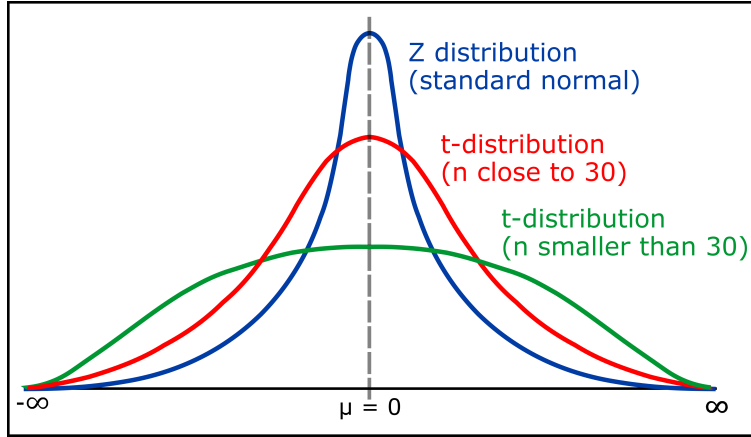


Figure 3: t-distribution compared to the normal distribution at different sample sizes. Image source[1]

A model's difficulty in calculating reliability is that it is impossible to know what a peer that did not fill in the survey would have given. Therefore, the Buddycheck reliability model assumes that the deviation that corresponds to the missing peer survey is equal to the t-score at the chosen confidence interval. The standard deviation calculation is thus altered based on these missing peer t-test scores, as shown in the equations below. Where 'SS' is the sum of squares, and t-score 't' is determined at the chosen confidence interval.

$$SS = \sum_i (N_i - \mu)^2 \quad (16)$$

$$SS_{adjusted} = SS + n_{missing} \cdot t^2 \quad (17)$$

$$\sigma_{adjusted} = \sqrt{\frac{SS_{adjusted}}{n - 1}} \quad (18)$$

This assumption will have three results: the first is that the more people don't fill in the survey, the bigger the adjusted group standard deviation will become. The second result is that the group size determines how severe the impact is of a student not filling in the survey. In case the number of peers that handed in the survey is very large, this would result in the t-score converging to the threshold value. Hence if the standard deviation calculated for peers that did fill in the survey is smaller than the threshold value, then the entire group would still be labeled as reliable independent of the total group size. The third result is that because the standard deviation for peers that did not fill in the survey is proportional to the threshold value, result two will always be true independent of the chosen threshold.

The Buddycheck reliability score can be used as an indicator by instructors to quickly check if the evaluation is reliable or if it might be better to send a reminder to those who did not fill in the survey to get a more trustworthy evaluation result. Although predicting what a student would fill in exactly is impossible, the assumption made to satisfy criteria three of this model provides a valuable tool to quickly estimate the impact of group size and missing scores on the reliability of an evaluation.

Feedback labels

The feedback labels are used to provide the peers and the instructor with insights into the work dynamics within a group. Table 4 shows all the feedback labels and their criteria and priority. Where Self is the average score given to self, Other is the average given to other peers, Other₁ is the given score to a single peer, REC is the average received score, GROUP is the average group score, and MEDIAN is the median score that a peer received.

label	description	criteria	priority
'SKEW'	skewed outcome of other peers	Self > 4 & Δ (Self - Other) > 2	1
'CONF'	is in conflict with another peer	Other ₁ < 2 & MEDIAN > 3	2
'UNDER'	is underconfident	REC > 3 & Δ (REC - Self) > 1	3
'OVER'	is overconfident	REC < 3 & Δ (Self - REC) > 1	4
'HIGH'	is a high performer	REC > 3.5 & Δ (REC - GROUP) > 0.5	5
'LOW'	is a low performer	REC > 2.5	6
'TENS'	group has formed tensions	$\sigma_{group} > 1$	7
'BYPASS'	group is not taking it serious	ALL SCORES EQUAL	8

Table 4: feedback labels with their criteria and priority, the highest priority label is valid and returned to the users.

SKEWED

Full: Skewed

Short value: SKEW

Description:

One of the students on the team appears to be trying to 'skew the curve' by giving themselves high ratings while rating the other team members poorly (it's possible that the student honestly believes that they were a star performer and the other team members failed to contribute). Strictly speaking, the student has to have given themselves an overall rating of 4 or higher while rating all of the other members on their team at least two points below this rating. The 'Skewed' condition is one that can cause erroneous 'Adjustment Factor' values and will result in the 'Adjustment Factor' column being highlighted.

Feedback to student:

Your self-evaluation indicates you made the primary contribution to the project with little value added by your teammates. The ratings from your teammates did not concur with your assessment. Your instructor may require additional information to clarify what happened in your team.

When (ALL must be true):

- If the user gives self a AVG higher than 4
- If the user gives others a lower AVG than self
- If the difference of the AVG the user gave self and the AVG gave to others is equal or bigger than 2

CONFLICT

Full: Somebody didn't get along

Short value: CONF

Description:

This condition is best summarized as 'somebody didn't get along'. Specifically, this student rated another team member at 2 or less, while the median rating for that student from the other team members is 3 or more. Note that the system checks for the 'Tension' condition before evaluating the data for potential personality conflicts. For the 'Tension' condition not to be flagged there must be significant agreement among the other students' scores, so this would appear to be an individual personality conflict. This is another condition that can result in erroneous 'Adjustment Factor' values, and so the 'Adjustment Factor' column will be highlighted in this case.

Feedback to student:

Your evaluation indicates that another student contributed very little to the project. This is not consistent with the assessment of the rest of the team. Your instructor may require additional information to clarify what happened in your team.

When (ALL must be true):

- If a user gives another student an AVG of 2 or less
- If that student received a 3 or higher as a median of the AVG scores given by others

UNDER

Full: Underconfident

Short value: UNDER

Description:

The overall team rating for this student is greater than 3, but the student rated themselves at least a point lower than this value. This would indicate that the student is 'underconfident' or too critical of their own contributions.

Feedback to student:

Your self-ratings were significantly lower than your teammates. The members of your team have indicated that you were a highly effective team member. Please try not to minimize the value of your contributions to the team.

When (ALL must be true):

- If the user received an AVG of 3 or higher from others
- If the user gives self a lower AVG than received from others
- If the difference of the AVG the user gave self and the AVG received from others is equal or bigger than 1

OVER

Full: Overconfident

Short value: OVER

Description:

This is essentially the opposite of the 'Underconfident' situation described above. The overall rating for the student is less than 3, but the student rated themselves a full point or more higher than their average rating. Possibly the student contributed but managed to alienate the other team members, or more likely they tend to overstate their own contributions.

Feedback to student:

Your self-ratings were significantly higher than your teammates. The members of your team indicated that your contributions to the team were below expectations. This report gives you details about how the

members of your team perceived your team contributions in one or more areas. Please use this information to identify problem areas in order to contribute effectively in future teamwork situations. Please contact your course instructor if you need assistance or if you believe that your ratings were inappropriate.

When (ALL must be true):

- If the user received an AVG of 3 or lower from others
- If the user gives self a higher AVG than received from others
- If the difference of the AVG the user gave self and the AVG received from others is equal or bigger than 1

HIGH

Full: High performer

Short value: HIGH

Description:

The student received an average rating better than 3.5 and their rating was more than half a point higher than the overall average for the team as a whole. Clearly this student was an exceptional contributor to this team's work.

Feedback to student:

Congratulations! The members of your team have indicated that you were a highly effective team member. Keep up the good work!

When (ALL must be true):

- If the user received an AVG of 3.5 or higher from others
- If the user received an higher AVG than the group AVG
- If the difference of the group AVG and the AVG received from others is equal or bigger than 0.5

LOW

Full: Low performer

Short value: LOW

Description:

This student did not contribute greatly to the team's success– their average overall rating is less than 2.5. This is difficult to achieve without significant lack of effort.

Feedback to student:

The members of your team indicated that your contributions to the team were below expectations. This report gives you details about how the members of your team perceived your team contributions in one or more areas. Please use this information to identify problem areas in order to contribute effectively in future teamwork situations. Please contact your course instructor if you need assistance or if you believe that your ratings were inappropriate.

When (ALL must be true):

- If the user received an AVG of 2.5 or lower from others

TENSION

Full: Considerable disagreements in group

Short value: TENS

Description:

None of the other conditions described above apply, but there is still significant disagreement between the ratings from various team members. 'Significant disagreement' here is defined as the sum of the standard

deviations for the ratings in each behavioral category being greater than the total number of behaviors surveyed (typically 5). When this occurs, it usually means that the team has split into multiple non-cooperating groups– hence 'Tension'. This is another situation that can lead to erroneous 'Adjustment Factor' values, and so the 'Adjustment Factor' column will be highlighted in this case.

Feedback to student:

There was considerable disagreement among your teammates as to which team members were most effective during team assignments. Your instructor may require additional information to clarify what happened in your team.

When (ALL must be true):

- If the sum of the standard deviations is higher than the amount of group users

BYPASS

Full: Zero difference in group

Short value: BYPASS

Description:

It seems this group is bypassing this peer evaluation. This is indicated when there are zero differences in the scoring and there are two or more peer questions in the evaluation. It is unlikely that each score of each peer is exactly the same. As an instructor, you could ask this group to fill in the evaluation again and to take it more seriously.

Feedback to student:

It looks like your group is bypassing this peer evaluation by giving everybody the exact same results for all questions. Your instructor may require additional information to clarify what happened in your team.

When (ALL must be true):

- All submitted scores of all peer questions are the same
- There are two or more peer questions

Results

After the calculation method is finished, the format of the object that is sent to the front end looks as follows in the next code snippet. Based on this object, each peer receives information about how he performed compared to the group. All peer and group reports are constructed with the data of this object.

```
1 map.[peer] = {
2   Model: 'string', // ('factorPeer'|'factorScore')
3   withSelf: boolean,
4   feedback: 'Label',
5   scaling: {min: number ,max: number, stepsize: number}},
6   receivedAvg: number,
7   factorOriginal: number,
8   factorCapped: number,
9
10  \\ each question gets his own object
11  question_id: {
12    givenAvgToSelf: number,
13    groupAvg: number,
14    receivedAvgWithoutSelf: number
15  }
16 };
```

Epilogue

Buddycheck aims to keep optimizing its calculation models and expand them with new features that can further improve the user experience. If you have any questions regarding the Buddycheck calculation model or if you have suggestions for how we can make Buddycheck even better, please don't hesitate to contact us:

Shareworks
Schiekade 105
3033 BH Rotterdam

Email: info@shareworks.nl

Bibliography

- [1] *geeksforgeeks Student's t-distribution in Statistics* (n.d.). URL: <https://www.geeksforgeeks.org/students-t-distribution-in-statistics/>. (accessed: 22.04.2021).
- [2] *Purdue University: Catme* (n.d.). URL: <https://info.catme.org/>. (accessed: 22.04.2021).
- [3] *Purdue University: Catme - Clique* (n.d.). URL: <https://info.catme.org/instructor-term/cliq-qliq/>. (accessed: 22.04.2021).