Supplement A: Additional Types of Opposites

Kimberly A. Barchard[1], James M. Carroll[2], Shawn Reynolds[2], and James A. Russell[2,3]
[1] Department of Psychology, University of Nevada, Las Vegas
[2] Department of Psychology, University of British Columbia
[3] Department of Psychology and Neuroscience, Boston College

This document is a supplement to the manuscript *Testing Bipolarity* by Barchard, Carroll, Reynolds, and Russell, published in *Psychological Methods*. This supplement provides expanded discussion of the possible relations between opposites and the resulting correlations.

Opposites That Are Free of Error

Imagine that B is a bipolar continuum, which has a neutral tipping point at 0, and which extends in both the positive and negative directions. Further imagine that x, y, and B are collinear: as x increases, B increases, and as y increases, B decreases. Thus, x and y are opposites. For convenience, let x and y each have the same scale as each other and as B, such that a one unit change in x or y is associated with a one unit change in B. In this section, x, y, and B are error free.

When x and y cover arbitrary portions of B, they can result in floor and ceiling effects. Floor effects occur when a large number of people get the minimum score. If x and y cover opposite halves of B, for example, half the scores on each variable are 0, and so both variables have strong floor effects. In contrast, ceiling effects occur when a large number of people get the maximum score. Ceiling effects can occur on any psychological measure where the response scale has a finite maximum (e.g., a five-point rating scale with a maximum score of 5). A generic model of opposites allows each variable to have any degree of floor and ceiling effects.

Because ceiling effects can occur on their own or in combination with floor effects, opposites have a wide variety of possible relationships beyond the ones shown in the main manuscript in Figure 1. Some of these additional possibilities are shown in Figures A1 and A2, below. For example, variables with gaps and overlaps may have the relationships shown in Panels A and B of Figures A1 and A2 (rather than the relationships shown in Figures 4 and 5). Additionally, two variables could be coextensive (as shown in Panel C of Figures A1 and A2), even if they do not both cover the whole of B (as was shown in Figures 1A and 2). If one variable has a floor effect and the other has a ceiling effect, they could cover the same end of B (as shown in Panel D of Figures A1 and A2), rather than opposite ends (as was shown in Figures 1 and 2). A special case of this occurs if they cover the same half of B (as shown in Panel E of Figures A1 and A2), rather than opposite halves (as shown in Panel B of Figures 1 and 2).

Correlations Between Opposites that are Free of Error

As shown in the main manuscript, if x and y cover opposite halves of B and B has a normal distribution, then x and y will correlate -.467. This supplement shows what happens to the correlation between x and y, ρ_{xy}, when those variables cover other portions of B. In many cases, the correlation is even further from –1.

To estimate the correlation between x and y when they cover arbitrary portions of B, we used an Excel file called CensorCorr (Barchard, 2023). CensorCorr randomly generates normally distributed[1] data for B, calculates scores for x and y based upon the portions that the user specifies (e.g., x covers the 20% of B and y covers 40%), and then calculates the correlation between x and y. We used CensorCorr to explore what happens to ρ_{xy} when x and y do not divide B evenly.

There are four ways x and y can fail to divide B evenly. First, even when together they cover all of B and are mutually exclusive, they may cover unequal proportions of B (see Figure 4), rather than dividing it in half. Indeed, with real data from human participants, a 50-50 split is unlikely. As the split moves away from 50-50, ρ_{xy} reduces in magnitude, approaching (but never quite reaching) 0. See Figure A3. For example, if the split is 25-75, their correlation is estimated as –.43. If the split is 10-90, their correlation is estimated as –.36.

Second, there might be a gap between x and y, as shown in Figure 5. The smaller the portions of B covered by x and y, the larger the gap and the more scores at (0,0). Because those (0,0) scores do not fall on the diagonal, they decrease the magnitude of the correlation (as shown in Figure A4). For example, if x and y each cover 40% of B (as shown near the middle of the figure), 20% of the scores are at (0,0), and ρ_{xy} is estimated as –.33.

Third, x and y might overlap each other. As the overlap increases, Figure 6 shows there are more scores on the diagonal section (and fewer on the vertical and horizontal sections); hence the magnitude of the correlation increases. For example, if x and y each cover 80% of B (as shown near the right-hand side of Figure A4), 60% of the scores overlap, and ρ_{xy} is estimated as –.88. When x and y overlap completely, they are coextensive (as seen in Figure 2) and correlate –1.

Fourth, x and y might cover the same end of B (as shown in Figures 1F and 7), rather than opposite ends (as shown in Figures 1C and 4). When x and y cover the same end of B, the correlation between them is larger in magnitude. For example, first imagine x covers 40% of B and y covers 60% (near the right-hand side of Figure A5). When x and y cover *opposite* ends (i.e., x covers the top of B and y covers the bottom, or vice versa), there is no overlap and the solid blue line shows ρ_{xy} is estimated as –.46. However, when x and y cover the *same* end of B (they both cover the top of B or both cover the bottom), there is 40% overlap and the red dashed line shows ρ_{xy} is estimated as –.96, which is .50 larger. The correlation is much larger when x and y cover the same end of B.

As the proportions become more uneven, the difference in the correlation decreases, but there is still a difference. For example, imagine x covers 10% of B and y covers 90% (near the

[1] Normally distributed data were simulated because normality is assumed by many statistical procedures (including significance tests on correlations) and may distort correlations less than other distributions.

left-hand side of the Figure A6). When x and y cover *opposite* ends, there is no overlap and the solid blue line shows ρ_{xy} is estimated as –.36. However, when x and y cover the *same* end of B, there is 40% overlap and the dashed red line shows ρ_{xy} is estimated as –.55, which is .19 larger. Thus, other things being equal, if x or y (or both) cover less than all of B, then the correlation between x and y will always be larger if they cover the *same* end of B than if they cover *opposite* ends.

In the past, researchers have often assumed that the correlation between error-free opposites is –1. In the main document, we have shown that this intuition is correct when the variables cover the same part of their dimension, whether or not they cover the whole of that dimension. However, our calculations and simulations have shown that the correlation between two error-free variables moves further and further from –1 when they cover smaller proportions of the dimension and when they cover more unequal proportions, especially if they cover opposite ends of the dimension – the case that is most often of interest when researchers want to test if two variables are opposites.

Opposites That May Be Influenced by Error

All empirical data are subject to measurement error. Such errors can distort the apparent relationships between opposites, possibly leading researchers to erroneously conclude that two variables are not opposites when they are – or are opposites when they are not.

To examine the effect of random and systematic errors on opposites, we turn to variables not necessarily ideal. We define two such variables, X and Y. Let x cover part or all of X and let y cover part or all of Y, so that these variables, too, might be influenced by measurement error. For convenience, let X and Y be zero-centered.

When x and y cover the whole of X and Y, they have the relationship shown in Figure 8. When x and y each cover the top of their dimension, they have the relationship shown in Figure 9. However, because x and y can cover arbitrary portions of X and Y (and have arbitrary levels of floor and ceiling effects), a wide variety of other relationships are possible. For example, x and y might cover more than half of their dimensions, as shown in Figure A7, or less than half, as shown by Figure A8. Or they might cover unequal parts, as shown in Figure 11. Critically, it would be possible for x and y to cover different ends of their dimensions. Perhaps x covers the top half of X and y covers the bottom half of Y, as shown in Figure A9. Or perhaps x and y cover unequal portions of their dimensions, as shown in Figure 12.

Correlations Between Opposites That May Be Influenced by Error

The correlation between x and y, ρ_{xy}, depends on the correlation between X and Y, ρ_{XY}, which is a function of their angle of separation (θ). For ideal (error free) scores, θ is 180°, and so $X = -Y$ and the correlation between X and Y is –1. For empirical scores influenced by measurement error, θ is less than 180° and the correlation between X and Y is smaller. We estimated the correlation ρ_{xy} using CensorCorr. We explored four patterns of coverage.

First, we explored the effect of uneven proportions, like those shown in Figure 11. Figure A6 shows that uneven proportions affect ρ_{xy} more. Both the solid blue line in Figure A6 (error-free data with $\rho_{XY} = -1$) and the dashed red line (data with some measurement error, such that $\rho_{XY} = -.707$) show that ρ_{xy} approaches 0 (getting closer to the top of the graph) as the split becomes more and more uneven (moving to the left). However, the lines also get closer together. Imagine x covers 40% of X and y covers 60% of Y (near the right-hand side of the figure where the horizontal axis reads .40). As ρ_{XY} varies from –1 to –.707, estimates of ρ_{xy} vary from –.462 to –.392, a difference of .070. Now imagine a more extreme example, one in which x covers 10% of X and y covers 90% of Y (near the left-hand side of the graph): As ρ_{XY} varies from –1 to –.707, estimates of ρ_{xy} vary from –.356 to –.298, a difference of just .058. Thus, ρ_{XY} has less and less effect on ρ_{xy} as the proportions become more uneven. In other words, the angle between X and Y, θ, which we assume here is caused entirely by measurement error, has less and less effect.

Second, we explored the effect of smaller coverage, as shown in Figure A8. Figure A10 shows that ρ_{xy} is affected more when x and y cover smaller proportions of X and Y. Both the solid blue line (where $\rho_{XY} = -1$) and the dashed red line (where $\rho_{XY} = -.707$) show that ρ_{xy} approaches 0 (getting closer to the top of the figure) as x and y cover less of X and Y (moving to the left). However, as x and y cover less of X and Y, the two lines also get closer together. Imagine x and y cover 90% of X and Y (near the right-hand side of the figure where the horizontal axis reads .90). As ρ_{XY} varies from –1 to –.707, estimates of ρ_{xy} vary from –.958 to –.678, a difference of .280. Now imagine x and y cover 60% of X and Y (near the middle of the figure). As ρ_{XY} varies from –1 to –.707, estimates of ρ_{xy} vary from as –.772 to –.564, a difference of .208. Finally imagine x and y cover just 20% of X and Y (near the left-hand side). As ρ_{XY} varies from –1 to –.707, estimates of ρ_{xy} vary from –.133 to –.132, a difference of only .001. In short, as x and y cover less and less of X and Y, ρ_{XY} (and θ, and measurement error) have less and less effect on the correlation ρ_{xy}.

Third, we checked what happens when the two variables cover different ends of their dimensions, as shown in Figure 12. If data are error-free and $\rho_{XY} = –1$, we previously showed that $\rho_{xy} = \rho_{XY}$ any time x and y cover the same sections of B (as in Figure A1, Panels C and E). Figure 8 shows this equality by the solid blue line at the bottom. However, if measurement error reduces the size of ρ_{XY}, then ρ_{xy} usually differs from ρ_{XY}. The dashed red line in Figure A11 shows the case where $\rho_{XY} = –.707$. For example, if x covers the top 30% of X and y covers the bottom 30% of Y (near the middle of the figure where the horizonal axis reads .30), ρ_{xy} is estimated as –.592 (so that $\hat{\rho}_{xy}$ and ρ_{XY} differ by .115). If x instead covers the top 10% of X and y covers the bottom 10% of Y (near the left-hand side of the figure), ρ_{xy} is estimated as –.491 (a difference of .216). Thus, as x and y cover smaller proportions of X and Y, ρ_{xy} moves further from ρ_{XY} and closer to 0. Similarly, if x and y cover larger proportions of X and Y (moving towards the right-hand side of the figure), ρ_{xy} moves towards ρ_{XY}. For example, if x covers the top 80% of X and y covers the bottom 80% of Y (near the right-hand side of the figure), ρ_{xy} is estimated as –.687, which is close to (but not identical to) ρ_{XY} of –.707. In short, when measurement error moves ρ_{XY} away from –1 and θ away from 180°, ρ_{xy} only equals ρ_{XY} when x and y cover *all* of X and Y.

Fourth, we compared the case where x and y cover the same parts of their dimensions (each covering the top of its dimension or each covering the bottom, as in Figure A7) to the case where they cover different parts (one covering the top of its dimension and the other covering the bottom, as in Figure 12). When correlations are *negative*, the differences between ρ_{xy} and ρ_{XY} are largest when x and y cover the *same* parts of their dimensions. Imagine ρ_{XY} = –.90 (the bottom panel of Figure 13), and consider the case where x covers 30% of X and y covers 70% of Y (near the middle of the figure). If x and y cover the same parts of their dimensions (the dashed red line), ρ_{xy} is estimated as –.43 (so that $\hat{\rho}_{xy}$ and ρ_{XY} differ by .47, as shown by the long arrow), but if x and y cover different parts (the solid blue line), ρ_{xy} is estimated as –.77 (a difference of just .13, as shown by the short arrow). Thus, for negative correlations, the difference between ρ_{xy} and ρ_{XY} are largest when x and y each cover the top of their dimensions, as would usually be the case when testing if two measures are opposite.

The reverse occurs when correlations are *positive*: The differences between ρ_{xy} and ρ_{XY} are largest when x and y cover *different* parts. For example, imagine ρ_{XY} = +.90 (the top panel of Figure 13). Consider again the case where x covers 30% of X and y covers 70% of Y (near the middle of the figure). If x and y cover different parts of their dimensions (the solid blue line), ρ_{xy} is estimated as +.43 (so that $\hat{\rho}_{xy}$ and ρ_{XY} differ by .47, as shown by the long arrow), but if x and y cover the same parts (the dashed red line), ρ_{xy} is estimated as +.77 (a difference of just .13, as shown by the short arrow). Thus, for positive correlations, the difference between ρ_{xy} and ρ_{XY} is *smallest* when x and y cover the *same* parts of their dimensions, as would usually be the case when two variables are intended to measure the same concept.

Because of this reversal, statistical methods that work well for positive correlations may not work well for negative correlations.

Summary

If researchers want to test if x and y are opposites, they must measure them separately and study their relations. If data are error free, x and y are opposites if they are collinear with each other. If instead data are subject to measurement error, then two variables are opposites if (a) x covers part or all of X, (b) y covers part of all of Y, and (c) after correcting for measurement error, the correlation between X and Y is -1. If data are error free, these two definitions are identical.

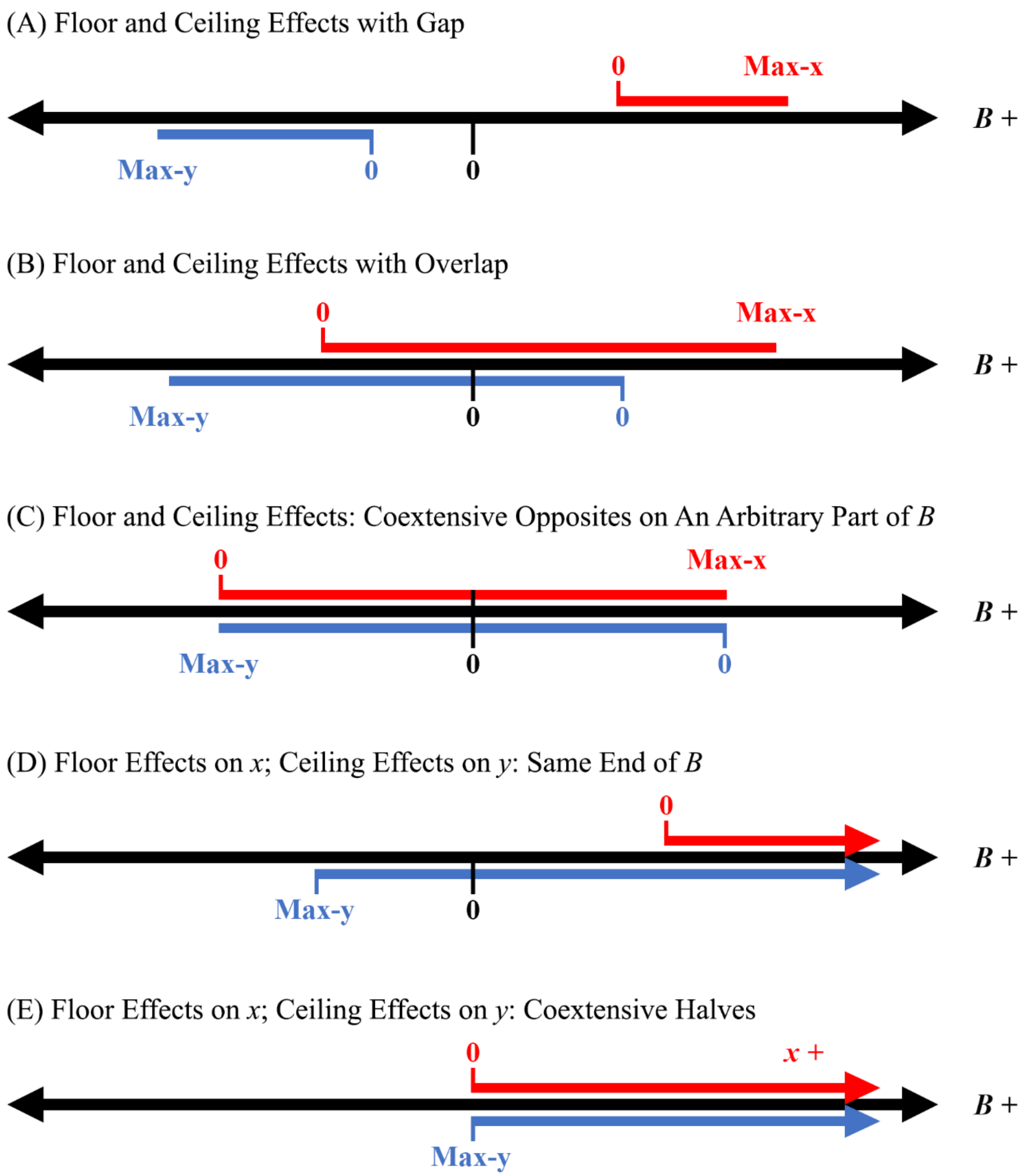
Testing if variables x and y are opposites is complicated because they (1) probably do not cover the whole of their dimensions, (2) probably do not cover identical proportions of their dimensions, and (3) probably cover the same ends of their dimensions. These three factors all reduce the magnitude of the correlation, so that ρ_{xy} can be far from -1. Moreover, this supplement has demonstrated that measurement error (in addition to changing the correlation directly) differentially affects negative correlations under those same three circumstances. Given that negative correlations are expected when testing if two variables are opposites, these differential effects have serious consequences for testing bipolarity, as discussed in the main manuscript.

Reference

Barchard, K. A. (2023). *CensorCorr: Estimating the effect of censoring on correlations* (Version 28) [Computer software]. https://osf.io/pfqy2/

Figure A1

Some Additional Types of Opposites



Note. B is a bipolar continuum, and x, y, and B are collinear. As x increases, B increases, and as y increases, B decreases. Thus, x and y are opposites. For convenience, let x and y each have the same scale as each other and as B, such that a one unit change in x or y is associated with a one unit change in B.

In Panel A, x and y both have ceiling effects and floor effects. In terms of the ceiling effects, the maximum value of x does not correspond to the maximum positive value on B, and so there are a lot of data points at the maximum value of x. Similarly, the maximum value of y does not correspond to the maximum negative value on B, and so there are a lot of data points at the maximum value of y. In terms of the floor effects, the minimum value of x does not correspond to the maximum negative value on B, and so there are many data points at the minimum value of x. Similarly, the minimum value of y does not correspond to the maximum positive value on B, and so there are many data points at the minimum value of y.

Here, x and y do not overlap. Instead, there are some values of B for which both x and y are 0. If x is any value besides 0, y is 0; similarly, if y is any value besides 0, x is 0.

In Panel B, x and y both have ceiling effects and floor effects, as was the case in the panel above. Here, x and y cover overlapping parts of B. Thus, there are some values of B for which both x and y are both non-zero.

In Panel C, x and y are coextensive; they cover the exact same part of B. This occurs because x and y both have ceiling effects and floor effects, and the ceiling effect of one exactly match the floor effect of the other (e.g., if 12% of the data are at the lowest value of x, then precisely 12% of the data are at the highest value of y; and if 37% of the data are at the highest value of x, then precisely 37% of the data are at the lowest value of y). Because x and y have the same scale, the maximum values of x and y are identical. For all values of x and y, x = max-x $- y$. $\rho_{xy} = -1$.

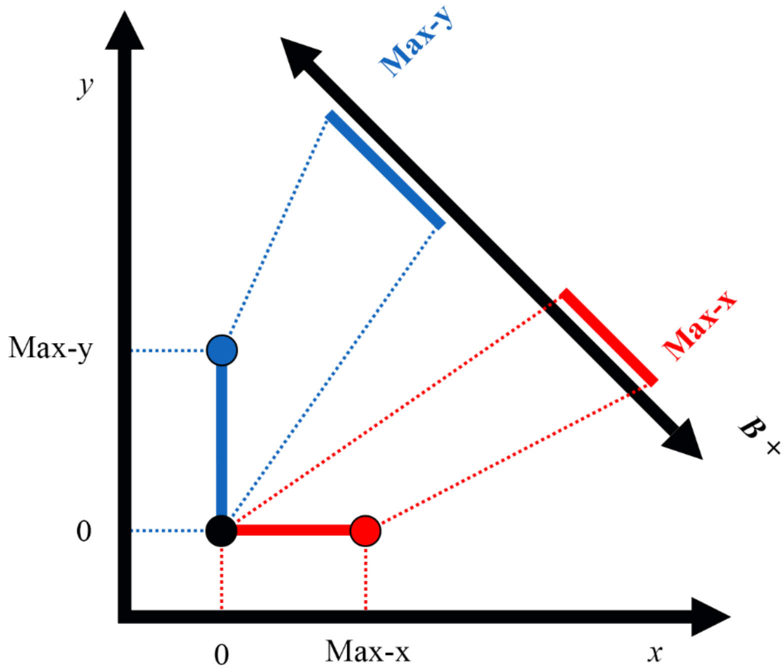
In Panel D, x and y each covers the top of B. In other words, x has a floor effect and y has a ceiling effect. However, x and y might cover different proportions of B, and so the maximum values of x, y, and B might not be identical.

In Panel E, x and y are coextensive. Specifically, each covers the top half of B. Thus, when $B > 0$, $x = B$, $y = -B$, and $y = -x$. When $B \leq 0$, $x = 0$, and y is at its maximum value. In other words, x has a floor effect and y has a ceiling effect. Because x, y, and B have the same scale, the maximum values of x, y, and B are identical. $\rho_{xy} = -1$.

Figure A2

Relationships between Additional Types of Opposites

(A) Floor and Ceiling Effects with Gap



Note. In all panels of Figure A2, x and y are opposites. The one-dimensional graphs in the upper right show the relations of x and y to B: These are reproductions of the panels in Figure A1. For example, the one-dimensional graph in Figure A2 Panel A is the same as the one-dimensional graph in Figure A1 Panel A.

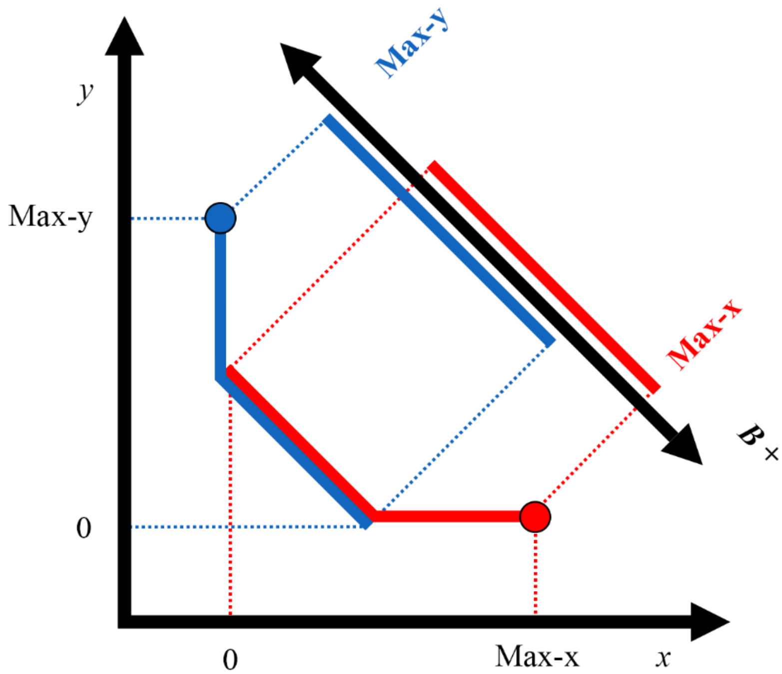
The two-dimensional graphs in the bottom left show the relation of x to y. Vertical lines represent sections of B in which x at its minimum value, corresponding to a floor effect on x. Horizontal lines represent sections in which y is at its minimum value, corresponding to a floor effect on y. Diagonal lines represent sections in which x and y overlap: As one increases the other decreases. Circles indicate that many data points occur at that location because of floor or ceiling effects.

The relations between the one-dimensional and two-dimensional graphs are shown by dotted lines.

Panel A. When x and y both have floor effects and ceiling effects (as shown in the top right of the panel), the two-dimensional graph (in the bottom left of the panel) includes a horizontal section where x is at its minimum and a vertical section where y is at its minimum. In fact, the horizontal section will occur whenever the floor effect on y is larger than the ceiling effect on x, and the vertical section will occur whenever the floor effect on x is larger than the ceiling effect on y. If the floor effects are large (as shown here), the two-dimensional graph will include some data points where x and y are both at their minimum (in the black circle), corresponding to the gap between x and y in the one-dimensional graph. This concentration of data points in the bottom left-hand corner will occur whenever sum(floor on x, floor on y) > 1. Finally, the two-dimensional graph includes concentrations of data points at the top of the vertical line (because of a ceiling effect on y) and at the right of the horizontal line (because of a ceiling effect on x). This combination of floor and ceiling effects creates an L-shaped graph.

This L-shaped relationship might occur any time researchers are attempting to determine if two variables are opposites: Ceiling effects may occur any time a response scale has a finite maximum value, but floor effects are likely to be larger than ceiling effects if researchers have attempted to divide a bipolar dimension into two. Depending upon the size of the floor effects, the relation between x and y will be either this one or the one in the figure below, Figure A2 panel B.

(B) Floor and Ceiling Effects with Overlap



Note. In all panels of Figure A2, x and y are opposites. The one-dimensional graphs in the upper right show the relations of x and y to B: These are reproductions of the panels in Figure A1. For example, the one-dimensional graph in Figure A2 Panel A is the same as the one-dimensional graph in Figure A1 Panel A.

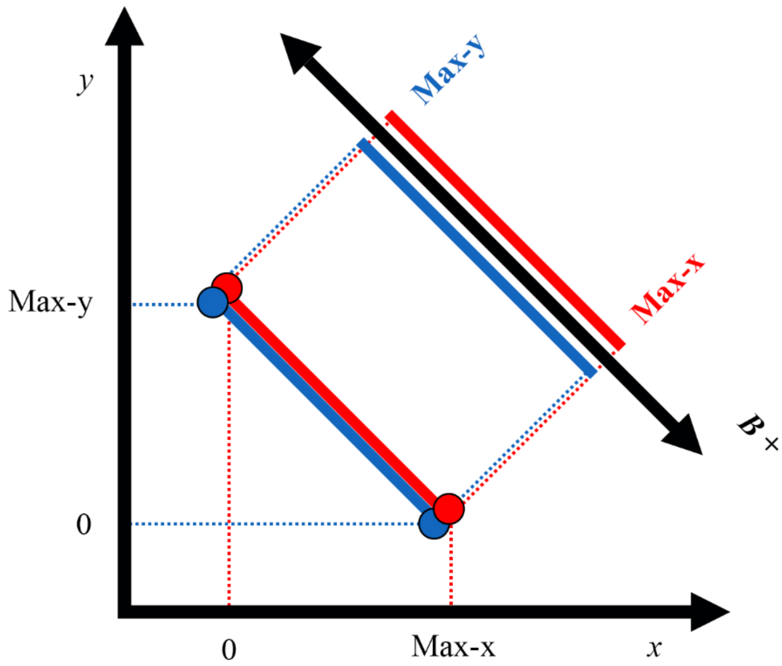
The two-dimensional graphs in the bottom left show the relation of x to y. Vertical lines represent sections of B in which x at its minimum value, corresponding to a floor effect on x. Horizontal lines represent sections in which y is at its minimum value, corresponding to a floor effect on y. Diagonal lines represent sections in which x and y overlap: As one increases the other decreases. Circles indicate that many data points occur at that location because of floor or ceiling effects.

The relations between the one-dimensional and two-dimensional graphs are shown by dotted lines.

Panel B: When x and y both have floor effects and ceiling effects (as shown in the top right of the panel), the two-dimensional graph (in the bottom left of the panel) includes a horizontal section where x is at its minimum and vertical section where y is at its minimum. In fact, the horizontal section will occur whenever the floor effect on x is larger than the ceiling effect on y, and the vertical section will occur whenever the floor effect on y is larger than the ceiling effect on x. If the floor effects are small (as shown here), the two-dimensional graph will include a diagonal section where x and y are both non-zero, corresponding to the overlap between x and y in the one-dimensional graph. This diagonal will occur whenever max(floor on x, ceiling on y) + max(ceiling on x, floor on y) < 1. Finally, the two-dimensional graph includes concentrations of data points at the top of the vertical line (because of a ceiling effect on y) and at the right of the horizontal line (because of a ceiling effect on x). This combination of floor and ceiling effects creates a bevel-edge shaped graph.

This beveled edge relationship might occur any time researchers are attempting to determine if two variables are opposites: Ceiling effects may occur any time a response scale has a finite maximum value, but floor effects are likely to be larger than ceiling effects if researchers have attempted to divide a bipolar dimension into two. Depending upon the size of the floor effects, the relationship will be either this one or the one in the figure above, Figure A2 Panel A

(C) Floor and Ceiling Effects: Coextensive Opposites on an Arbitrary Part of B



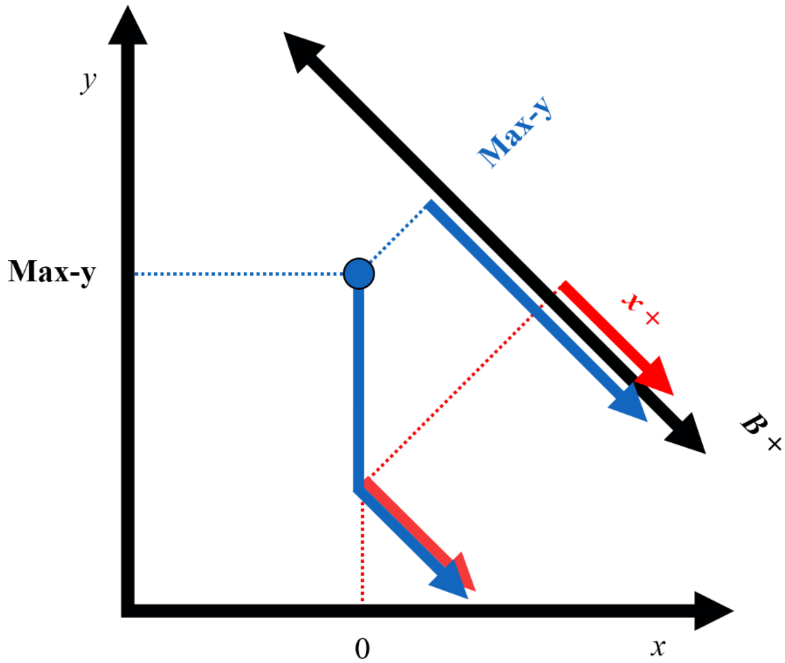
Note. In all panels of Figure A2, x and y are opposites. The one-dimensional graphs in the upper right show the relations of x and y to B: These are reproductions of the panels in Figure A1. For example, the one-dimensional graph in Figure A2 Panel A is the same as the one-dimensional graph in Figure A1 Panel A.

The two-dimensional graphs in the bottom left show the relation of x to y. Vertical lines represent sections of B in which x at its minimum value, corresponding to a floor effect on x. Horizontal lines represent sections in which y is at its minimum value, corresponding to a floor effect on y. Diagonal lines represent sections in which x and y overlap: As one increases the other decreases. Circles indicate that many data points occur at that location because of floor or ceiling effects.

The relations between the one-dimensional and two-dimensional graphs are shown by dotted lines.

Panel C: Whenever x and y are coextensive opposites (covering identical sections of B), they have an inverse relationship: As one variable increases, the other decreases. Most commonly, this relationship will occur when there are no ceiling or floor effects so that x and y each cover the whole of B. Theoretically, though, this relationship occurs (as shown in this panel) whenever (a) the floor effect on x is identical to the ceiling effect on y, and (b) the ceiling effect on x is identical to the floor effect on y.

(D) Floor Effects on x; Ceiling Effects on y: Same End



Note. In all panels of Figure A2, x and y are opposites. The one-dimensional graphs in the upper right show the relations of x and y to B: These are reproductions of the panels in Figure A1. For example, the one-dimensional graph in Figure A2 Panel A is the same as the one-dimensional graph in Figure A1 Panel A.

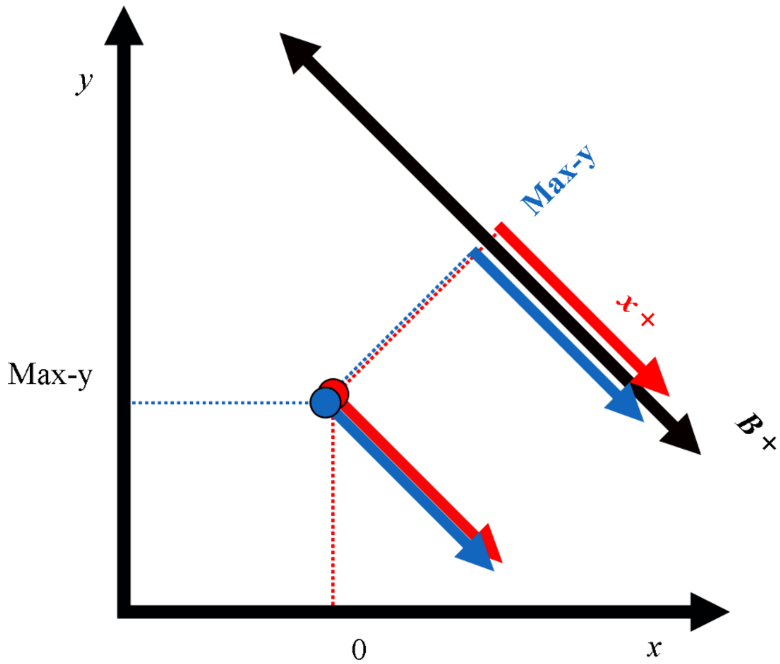
The two-dimensional graphs in the bottom left show the relation of x to y. Vertical lines represent sections of B in which x at its minimum value, corresponding to a floor effect on x. Horizontal lines represent sections in which y is at its minimum value, corresponding to a floor effect on y. Diagonal lines represent sections in which x and y overlap: As one increases the other decreases. Circles indicate that many data points occur at that location because of floor or ceiling effects.

The relations between the one-dimensional and two-dimensional graphs are shown by dotted lines.

Panel D: Here x and y each covers the top of B. In other words, x has a floor effect and y has a ceiling effect. However, x and y might cover different proportions of B, and so the maximum values of x, y, and B might not be identical.

This boomerang-shaped relationship might occur any time a researcher has attempted to divide a bipolar dimension into two parts and has created positively and negatively keyed items to measure one part (e.g., two measures of heat being used to measure half of the hot—cold dimension; or two measures of extraversion being used to measure half of the extraversion—introversion dimension).

(E) Floor Effects on x; Ceiling Effects on y: Coextensive Halves



Note. In all panels of Figure A2, x and y are opposites. The one-dimensional graphs in the upper right show the relations of x and y to B: These are reproductions of the panels in Figure A1. For example, the one-dimensional graph in Figure A2 Panel A is the same as the one-dimensional graph in Figure A1 Panel A.

The two-dimensional graphs in the bottom left show the relation of x to y. Vertical lines represent sections of B in which x at its minimum value, corresponding to a floor effect on x. Horizontal lines represent sections in which y is at its minimum value, corresponding to a floor effect on y. Diagonal lines represent sections in which x and y overlap: As one increases the other decreases. Circles indicate that many data points occur at that location because of floor or ceiling effects.

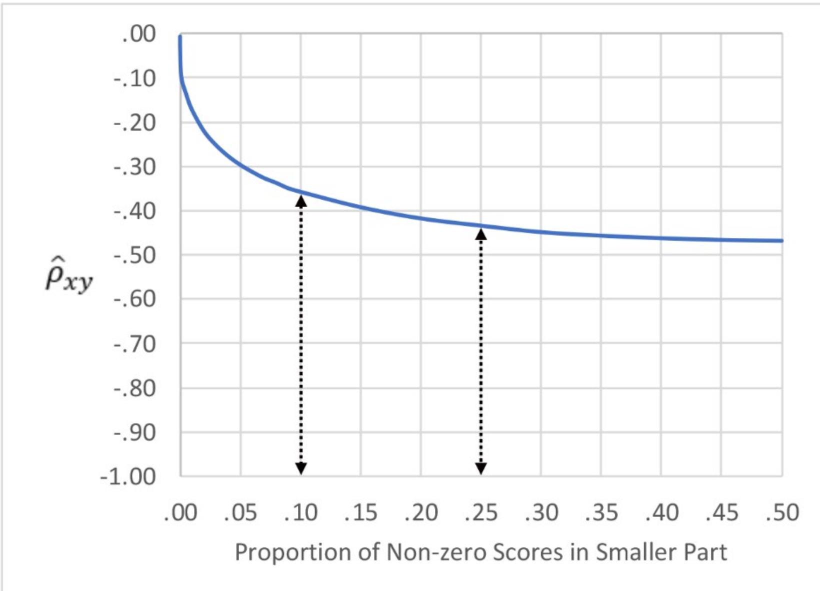
The relations between the one-dimensional and two-dimensional graphs are shown by dotted lines.

Panel E: Whenever x and y are coextensive opposites (covering identical sections of B), they have an inverse relationship: As one variable increases, the other decreases. Here, x and y both cover the top of B: When $B > 0$, $x = B$ and $y = -B$; when $B < 0$, $x = 0$ and y = its maximum value. Thus, the floor effect on x exactly matches the ceiling effect on y.

Theoretically, this relationship might occur any time a researcher has attempted to divide a bipolar dimension in half and has created positively and negatively keyed items to measure one part (e.g., two measures of heat being used to measure half of the hot—cold dimension; or two measures of extraversion being used to measure half of the extraversion—introversion dimension). However, in actual research, it is unlikely that two items will cover identical proportions of B (without covering the whole of B) or that either will cover exactly 50%. Thus, the relation shown in Figure A2 Panel D is more likely.

Figure A3

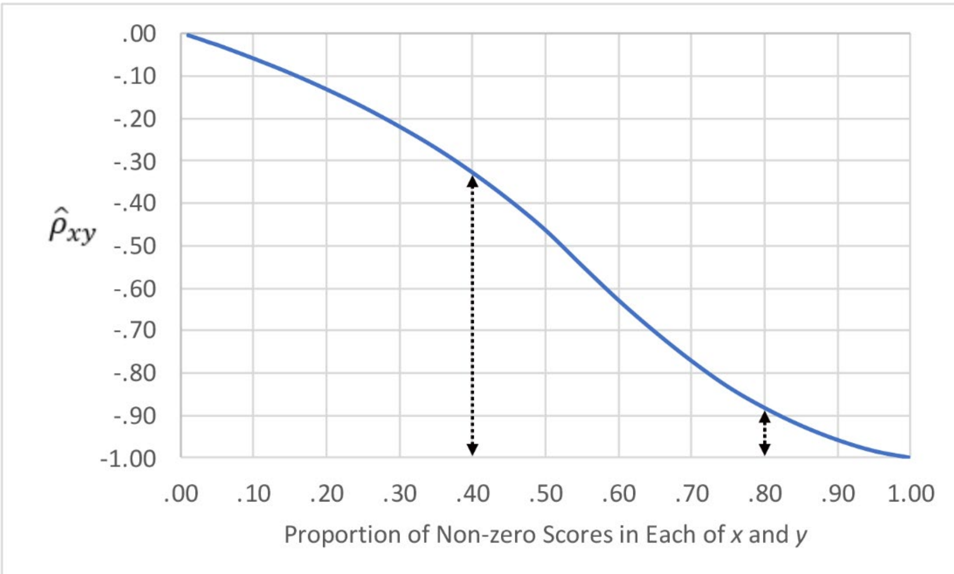
Correlation Between Error-free x and y as a Function of the Proportion of Non-Zero Values for the Smaller Part



Note. Here, x and y are complementary parts of B (i.e., they are mutually exclusive and exhaustive), as shown in Figures 1 and 2, Panel C. Thus, the proportion that is covered by x is p, and the proportion that is covered by y is 1-p. The horizontal axis shows the smaller part, which is the minimum of 1 and 1-p. The vertical axis shows the correlations between x and y, which were estimated using 500,000 cases that were randomly generated using CensorCorr (Barchard, 2023), which assumes that B has a normal distribution. For example, if x and y divide B .25/.75, $\hat{\rho}_{xy}$ = -.43; if x and y divide B .10/.90, $\hat{\rho}_{xy}$ = -.36.

Figure A4

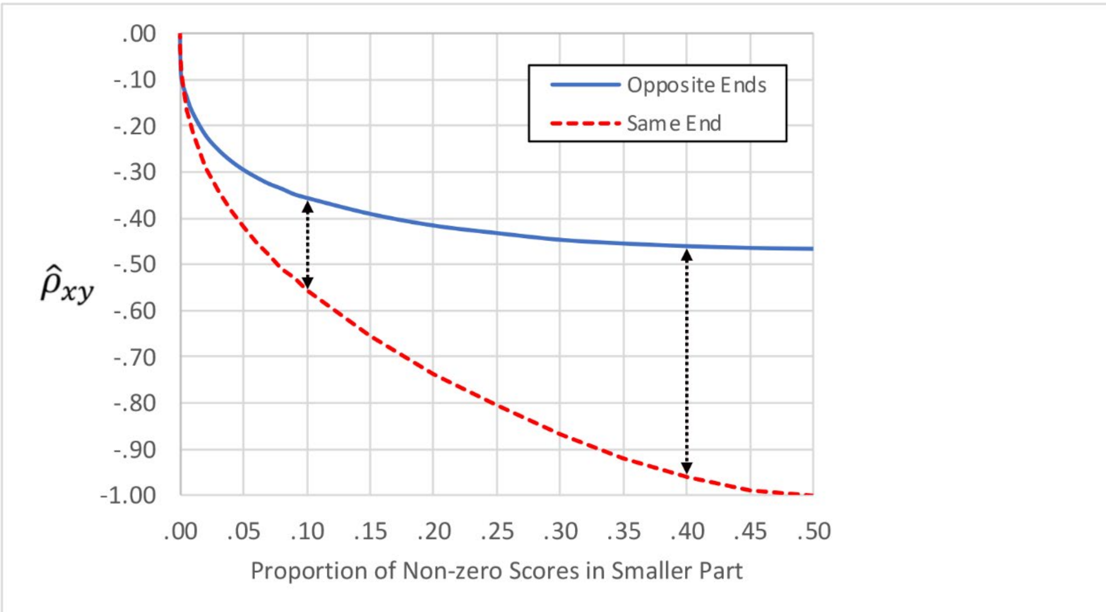
Correlation Between Error-free x and y as a Function of the Proportion of Non-Zero Values for Each



Note. Here, x and y cover opposite ends of B, and each covers the same proportion of B. If the proportion covered by each is less than .50, there is a gap, as shown in Figures 1D and 5. For example, if x and y each cover .40 of B, there is .20 gap. If the proportion covered by each is more than .50, there is an overlap, as shown in Figures 1E and 6. For example, if x and y each cover .80 of B, there is .60 overlap. The vertical axis shows the correlations between x and y, which were estimated using 500,000 cases that were randomly generated using CensorCorr (Barchard, 2023), which assumes that B has a normal distribution. For example, if x and y each cover .40 of B, $\hat{\rho}_{xy}$ = -.33; but if x and y each cover .80 of B, $\hat{\rho}_{xy}$ = -.88.

Figure A5

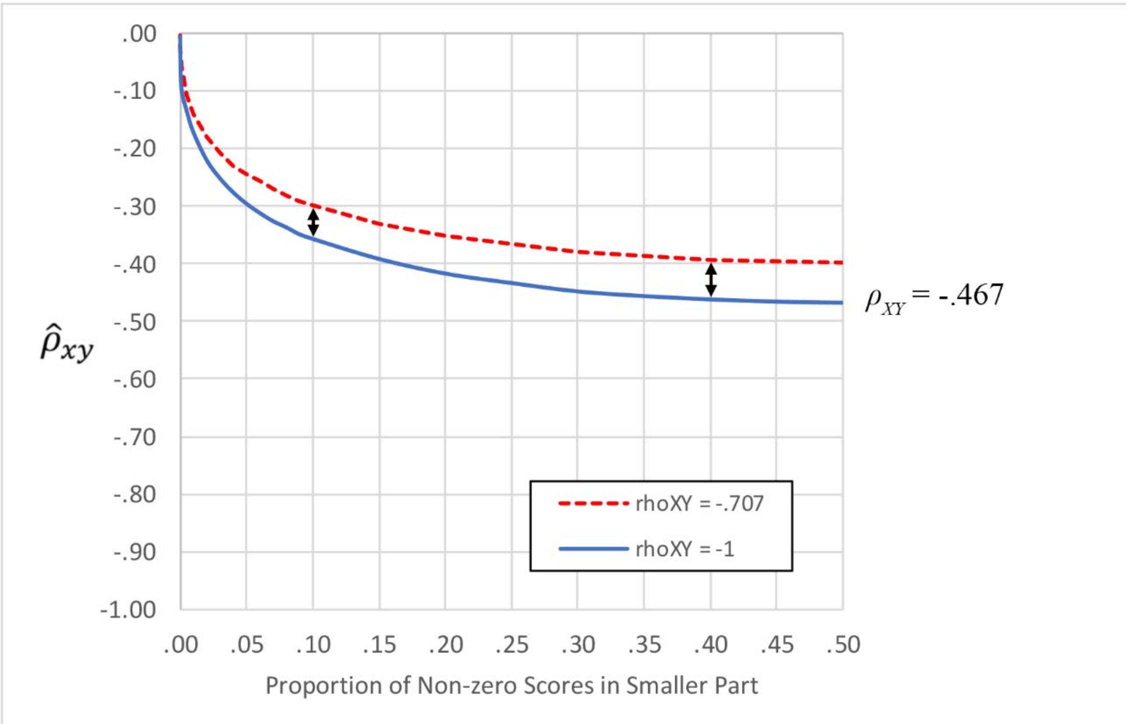
Correlation Between Error-free x and y as a Function of the Proportion of Non-Zero Values for the Smaller Part: When x and y Cover the Same End of B or Opposite Ends



Note. Here, x and y are opposites that cover unequal proportions of B: The proportion that is covered by x is p, and the proportion that is covered by y is 1-p, as shown in Figures 1 and 2, Panel C. The solid blue line assumes x and y cover opposite ends of B so that they are mutually exclusive and exhaustive. The dashed red line assumes that x and y cover the same end of B and hence overlap. The horizontal axis shows the smaller part, which is the minimum of 1 and 1-p. The vertical axis shows the correlations between x and y, which were estimated using 500,000 cases that were randomly generated using CensorCorr (Barchard, 2023), which assumes that B has a normal distribution.

Figure A6

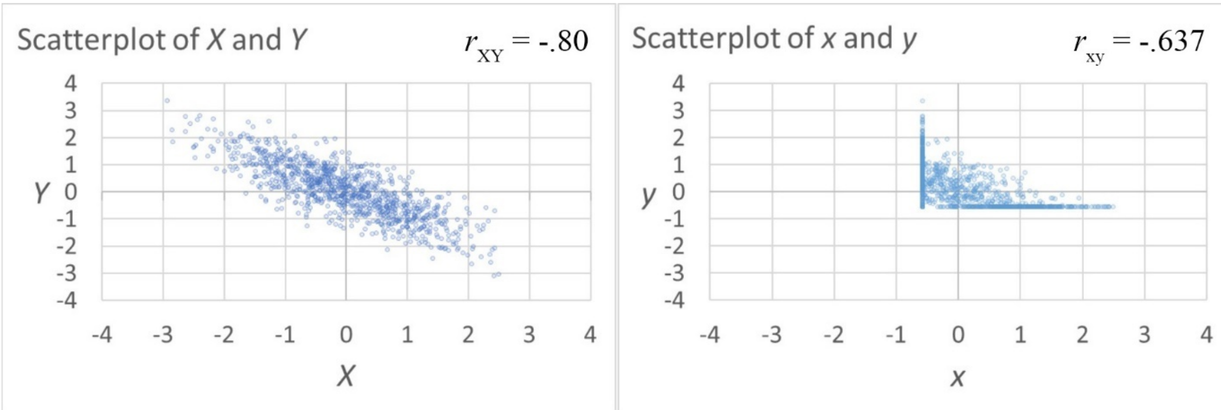
Correlation between x and y as a Function of ρ_{XY} and the Proportion of Non-zero Values in the Smaller Part, When x Covers the Top p% of X and y Covers the Top (1-p)% of Y



Note. The solid blue line assumes ρ_{XY} = -1 (θ = 180°). The dashed red line assumes that ρ_{XY} = -.707 (θ = 135°). Both lines assume that x and y cover the top parts of X and Y, respectively, that the proportion of X that is covered by x is p, and that the proportion of Y that is covered by y is 1-p, as shown in Figure 4 for ρ_{XY} = -1 and Figure 11 for $\rho_{XY} \neq$ -1. The horizontal axis shows the smaller part, which is the minimum of 1 and 1-p. The vertical axis shows the correlations between x and y, which were estimated using 500,000 cases that were randomly generated using CensorCorr, which assumes X and Y have a multivariate normal distribution.

Figure A7

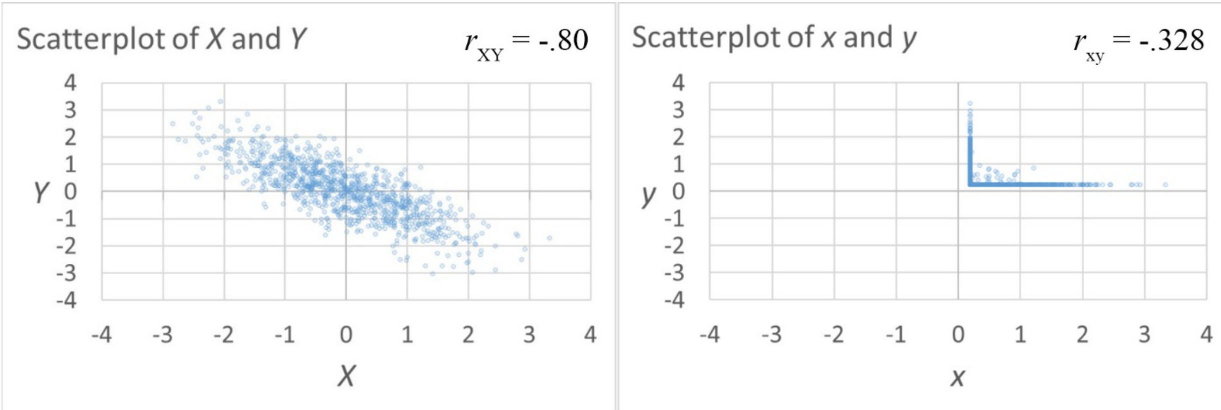
A Case in Which x Covers the Top p% of X and y Covers the Top p% of Y: Large Portions



Note. Here, x and y cover the top parts of X and Y, respectively, and the proportions covered are equal: Thus, the proportion of X that is covered by x is p, and the proportion of Y that is covered by y is also p. θ is the angle between X and Y. X and Y are zero-centered; therefore, the correlation between X and Y is $\cos(\theta)$. However, x and y are not zero-centered, and so ρ_{xy} does not generally equal $\cos(\theta)$. Instead, if X and Y have a bivariate normal distribution, CensorCorr can be used to estimate the correlation between x and y. The scatterplots (created using CensorCorr with 10,000 cases that were randomly generated assuming a bivariate normal distribution) show the case where x covers 70% of X, y covers 70% of Y, and $\rho_{XY} = -.80$, for which ρ_{xy} was estimated as $-.637$.

Figure A8

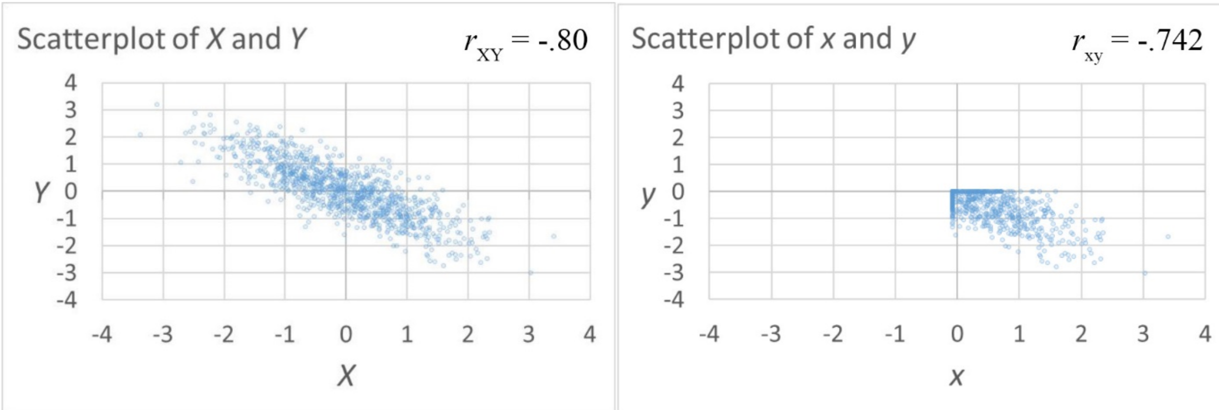
A Case in Which x Covers the Top p% of X and y Covers the Top p% of Y: Small Portions



Note. Here, x and y cover the top parts of X and Y, respectively, and the proportions covered are equal: Thus, the proportion of X that is covered by x is p, and the proportion of Y that is covered by y is also p. θ is the angle between X and Y. X and Y are zero-centered; therefore, the correlation between X and Y is $\cos(\theta)$. However, x and y are not zero-centered, and so ρ_{xy} does not generally equal $\cos(\theta)$. Instead, if X and Y have a bivariate normal distribution, CensorCorr can be used to estimate the correlation between x and y. The scatterplots (created using CensorCorr with 10,000 cases that were randomly generated assuming a bivariate normal distribution) show the case where x covers 40% of X, y covers 40% of Y, and $\rho_{XY} = -.80$, for which ρ_{xy} was estimated as $-.328$.

Figure A9

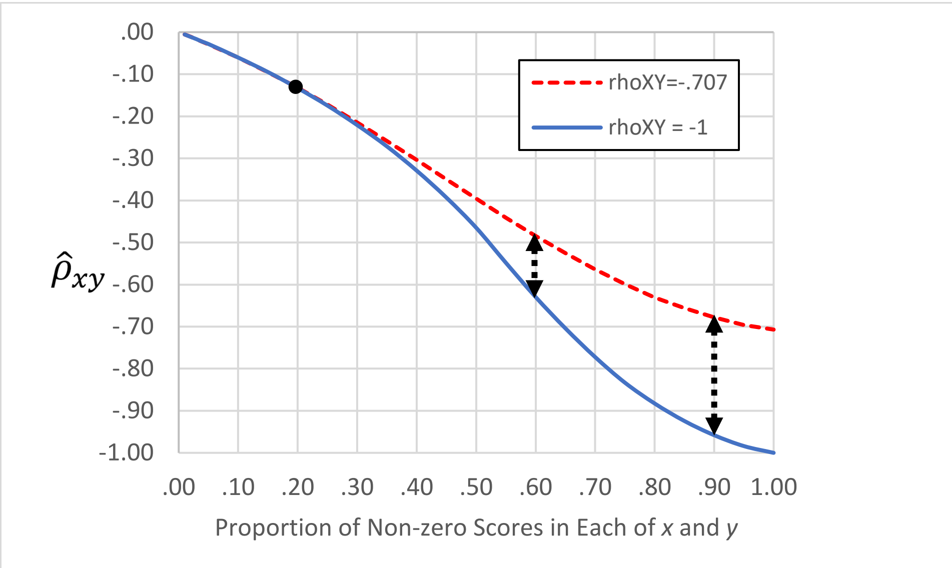
A Case in Which x Covers the Top Half of X and y Covers the Bottom Half of Y



Note. Pictured is a special case in which x covers the top half of X and y covers the bottom half of Y. Thus, $x = X$ for $X > 0$ and 0 otherwise; y = max-y − Y for $Y < 0$ and max-y otherwise. θ is the angle between X and Y. X and Y are zero-centered; therefore, the correlation between X and Y is $\cos(\theta)$. However, x and y are not zero-centered, and so ρ_{xy} does not generally equal $\cos(\theta)$. Instead, if X and Y have a bivariate normal distribution, CensorCorr can estimate the correlation between x and y. The scatterplots (created using CensorCorr with 10,000 cases that were randomly generated assuming a bivariate normal distribution) show the case where $\rho_{XY} = -.80$, for which ρ_{xy} was estimated as $-.742$.

Figure A10

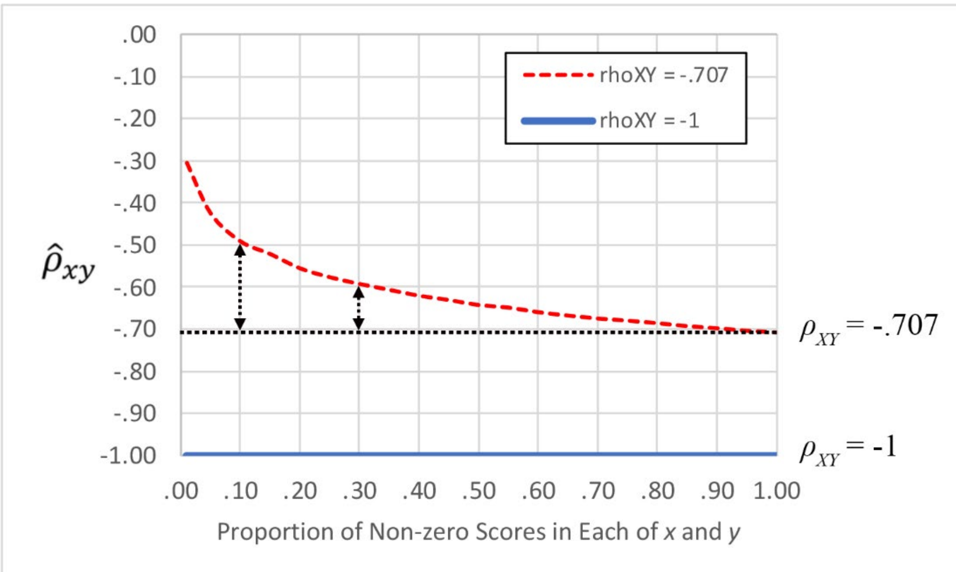
Correlation between x and y as a Function of ρ_{XY} and the Proportion of Non-zero Values in Each of x and y, When x Covers the Top p% of X and y Covers the Top p% of Y



Note. The solid blue line assumes $\rho_{XY} = -1$ (θ = 180°). The dashed red line assumes $\rho_{XY} = -.707$ (θ = 135°). Both lines assume that x and y cover the top parts of X and Y, respectively, and that the proportion of X that is covered by x is the same as the proportion of Y that is covered by y as shown in Figures 1D, 1E, 5, and 6 for $\rho_{XY} = -1$ and Figures A7 and A8 for $\rho_{XY} \neq -1$. The vertical axis shows the correlations between x and y, which were estimated using 500,000 cases that were randomly generated using CensorCorr, which assumes X and Y have a multivariate normal distribution.

Figure A11

Correlation Between x and y as a Function of ρ_{XY} and the Proportion of Non-zero Values in Each of x and y, When x Covers the Top p% of X and y Covers the Bottom p% of Y



Note. Here, x covers the top p% of X, and y covers the bottom p% of Y. The solid blue line assumes ρ_{XY} = -1 (θ = 180°); hence, $x = B$, $Y = -B$, and x and y cover identical sections of B. The dashed red line assumes ρ_{XY} = -.707 (θ = 135°); hence, B does not exist. See Figure A9 for an example where x and y cover 50% of X and Y, respectively. The vertical axis shows the correlations between x and y, which were estimated using 500,000 cases that were randomly generated using CensorCorr, which assumes X and Y have a multivariate normal distribution.