Working with difference scores: An applied primer

Christopher J. L. Cunningham
The University of Tennessee at Chattanooga

THE UNIVERSITY of TENNESSEE at CHATTANOOGA

Objective & Rationale

- To educate regarding promises and pitfalls associated with difference scores
- Why? Because despite regularly published limitations, there are still times when difference scores are a natural and/or logical approach to conceptualizing and testing certain research questions.
- Many important occupational health issues involve issues of fit, discrepancy, congruence
  - Edwards & Van Harrison (1993)

What is a Difference Score (DiffScore)?

- "…discrepancies between distinct but conceptually related constructs (Tisak & Smith, 1994, p. 675)
- "…subtraction of one measure from another to create a measure of a distinct construct…" (Peter et al., 1993, p. 655)
- A.K.A., change scores, discrepancy indices, gain scores, deviation scores…

Varieties of DiffScores

- Algebraic
  - (X-Y)
- Absolute
  - |X-Y|
- Squared Difference
  - $(X-Y)^2$
- Profile Similarity Indices
  - Variations on the above

Reasons for DiffScores

- High face validity; intuitive nature and meaning
- Difference scores often assumed to represent something distinct from their component parts
- Frequently used in repeated measures analyses
- Frequently used to operationalize issues that matter to applied researchers, consultants, and professional practitioners

Common Uses of DiffScores

- To provide a dependent variable/outcome for intervention/ education studies
  - Discrepancy
  - Fit/Similarity
  - Congruence
  - Change
- To provide a measure of growth rate
- To indicate deviance in development

So What's the Big Deal?

- There has been a good deal of contention for many years about the use of difference scores (e.g., Cronbach & Furby, 1970; Edwards, 1994; Tisak & Smith, 1994).
- Limitations/challenges of working with difference scores are perhaps most strongly identified and discussed by Edwards (2002; 1995)

DiffScore Limitations: Reliability

- Reliability is complicated because X, Y often not from independent sources (Cronbach & Furby, 1970)
  - Influenced by correlation between the components, which are typically correlated (Johns, 1981; Peters et al., 1993)
  - If X and Y are both reliably measured, the actual reliability of (X – Y) will be lower if X and Y are moderately correlated → if average X, Y reliability = .70, and r_{XY} = .50, then (X – Y) reliability is closer to .40
  - Rogosa (1988) disagrees – may not always be the case

DiffScore Limitations: Uniqueness

- Are difference scores really, seriously different than their components?
  - More may be gained by considering both components separately as predictors in a regression framework (Cronbach & Furby, 1970; Edwards, 2002; 1994*)

DiffScore Limitations: Confusion & Inaccuracy

- Especially when used in correlation/regression:
  1) Low reliability poses interpretational challenges (Peters et al., 1993)
  2) Type of change/difference/misfit matters, right?
     - 3 general diffscore models: discrepancy, similarity, superiority
     - Important to consider underlying correlations and variances of the main effects separately
  3) Also, consider that if (X – Y) → Z, the regression weight of Y is constrained to be -1, which is probably not accurate (Cohen et al., 2003; Edwards, 2002)

Point to Ponder

"The practice of using difference score correlations alone to draw inferences about the benefits or costs of similarity or accuracy is like characterizing the movie Casablanca as a story about a man, a woman, and an airplane. The results of a difference score correlation are just as 'correct' as the foregoing summary of the movie, but in both cases the overall summaries leave the audience in the dark."

– Griffin et al. (1999), p. 517

What about Polynomial Regression?

- To avoid diffscore limitations and still answer interesting questions, we could use polynomial regression and response surface modeling (RSM)
- Can help to address the following questions
  - How does level of agreement/discrepancy between the two predictors relate to the outcome?
  - How does direction of discrepancy factor in?
    - What is the impact on my outcome (Z) if X > Y or if X < Y?

Response Surface Modeling

- A 3-d graphical approach to facilitating interpretation of polynomial regression results
  - Approach to analyzing features of surfaces that are based on polynomial regression equations
    - *Stationary point* – where surface is flat
    - *Principal axes* – describe the orientation of the plane
    - *Shape of surface* – relative to points on the X,Y plane, specifically those falling along the lines that fit Y = X and Y = -X

Great – How do I…?

- Yeah, about that…
- Although polynomial regression and RSM may address some diffscore limitations, these two techniques are not exactly introductory statistical issues
- Best instructional resources:
  - *For beginners*: Shanock et al., 2010 and this presentation
  - *For more advanced users*, Edwards (2002)
- On a high level, this is how it works…

Requirements

- Interest in examining how two predictor variables are related to an outcome
- Both predictors must represent the same conceptual domain (i.e., be *commensurate*)
- Predictors must be measured on same numeric scale (or be standardized first)
- All usual assumptions of multiple regression should be met

Polynomial Regression Elements

- Generate polynomial regression coefficients for the actual RSM

$$Z = b_0 + b_1X + b_2Y + b_3X^2 + b_4XY + b_5Y^2 + e$$

- Now use the *b* coefficients to illustrate the response surface pattern
- For a good full example in print, see Harris et al., 2008, *Journal of Applied Psychology*

Example (following Shanock et al., 2010)

1) Establish base rate of discrepancy
   - If X and Y are never discrepant, then it is probably not worth taking this approach
   - Can use Fleenor et al.'s (1996) technique:
     - Standardize X, Y variables
     - Any standardized score .5 SD +/- the standardized score of the other = "discrepancy"
     - Calculate % of "in agreement" and "discrepant" values

Example Step 1 SPSS Syntax

```
DESCRIPTIVES VARIABLES= VariableX VariableY .
    /SAVE
    /STATISTICS=MEAN STDDEV MIN MAX.

COMPUTE diff1= VariableX – VariableY.
EXECUTE.
RECODE diff1 (Lowest thru -.5=-1) (.5 thru Highest=1) (ELSE=0) INTO countdiff1.
EXECUTE.

FREQUENCIES VARIABLES=countdiff1
  /ORDER=ANALYSIS.
SORT CASES  BY countdiff1.
SPLIT FILE LAYERED BY countdiff1.

DESCRIPTIVES VARIABLES=VariableX  VariableY
  /STATISTICS=MEAN STDDEV MIN MAX.
```

Example: Predicting Work-Life Balance

Agreement Group	Percentage	Mean WNWE	Mean NWWI
NWWI > WNWE	21.4	1.86 (0.68)	2.41 (0.73)
WNWE = NWWI	30.2	2.68 (0.56)	1.86 (0.50)
WNWE > NWWI	48.4	3.66 (0.66)	1.38 (0.41)

N = 126
NWWI = Nonwork-to-work interference
WNWE = Work-to-nonwork interference

In these data it is evident that there is a variety of discrepancy and agreement between these two variables within this sample

Example cont'd

2) Polynomial regression

- Start by centering predictors on scale midpoint
  - If a 5-point scale, center at 3
- Construct new variables (X^2, Y^2, XY)
- Run regression
- If R^2 is significantly different from 0, test the 4 surface values (a_1, a_2, a_3, a_4)

Example Step 2 SPSS Syntax

```
Compute varxcntr = VariableX – [midpoint # value].
Compute varycntr = VariableY - [midpoint # value].
Execute.

COMPUTE xy = varxcntr*varycntr.
COMPUTE xsqrd = varxcntr*varxcntr.
COMPUTE ysqrd = varycntr*varycntr.
Execute.

REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS BCOV R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT  [outcome of interest]
  /METHOD=ENTER varxcntr varycntr xy xsqrd ysqrd.
```

Surface Test Values

- $a_1 = (b_1 + b_2)$
  - Slope of line of perfect agreement (X = Y)
- $a_2 = (b_3 + b_4 + b_5)$
  - Curvature along line of perfect agreement
- $a_3 = (b_1 - b_2)$
  - Slope of line of incongruence (indicates if X > Y or Y > X)
- $a_4 = (b_3 - b_4 + b_5)$
  - Curvature along line of incongruence (indicates degree of discrepancy between X, Y, and outcome)

- You can obtain simple significance tests of these *a* values using *t* formulas in Shanock et al. (2010)
- Advanced users can consider bootstrapping for potentially more precise estimates of statistical significance (Cafri et al., 2010; Edwards, 2002;)

Example Step 2 Results

	Predicting Work-Life Balance	
	Variable	*b* (se)
*b*0	Constant	4.770 (.193)
*b*1	X =Work-nonwork enhancement (WNWE)	0.166 (.176)
*b*2	Y =Nonwork-work interference (NWWI)	-1.213 (.326)
*b*3	X^2 = WNWE squared	0.049 (.069)
*b*4	XY = WNWE x NWWI	0.046 (.115)
*b*5	Y^2 = NWWI squared	-0.223 (.148)
	R^2	.33**
Surface Tests		
	a_1 = [b1 + b2]	-1.05*
	a_2 = [b3 + b4 + b5]	-0.13
	a_3 = [b1 – b2]	1.38
	a_4 = [b3 – b4 + b5]	-0.22

Example cont'd

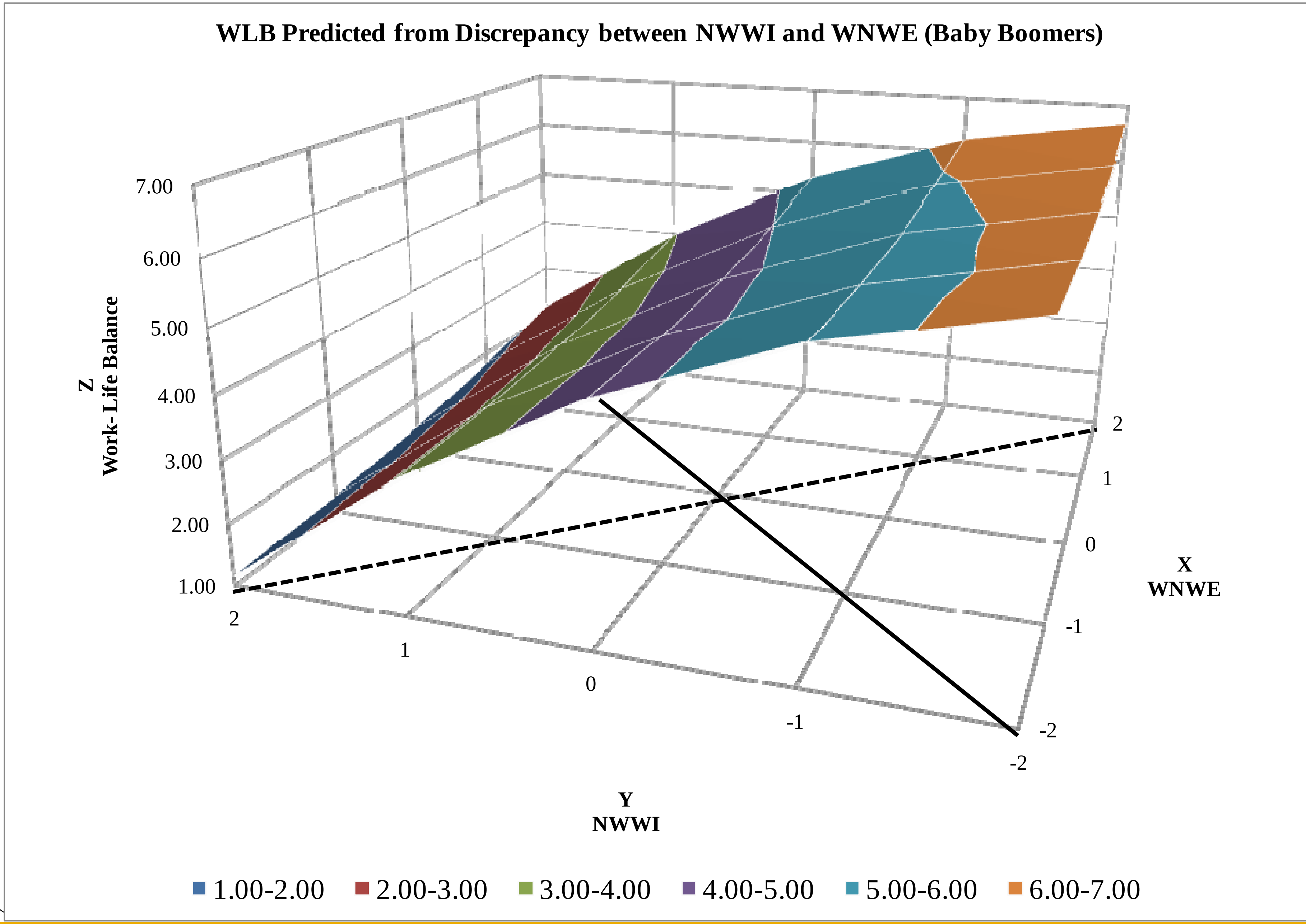
3) Plot the results in Excel
   - You can design your own workbook or use one of the several that already exist
     - Cafri et al. (2010)
     - Shanock et al. (2010)
     - *the one associated with this presentation*
   - You will need the *b* coefficients and some example points to plot first

Step 3: Graphing the Results

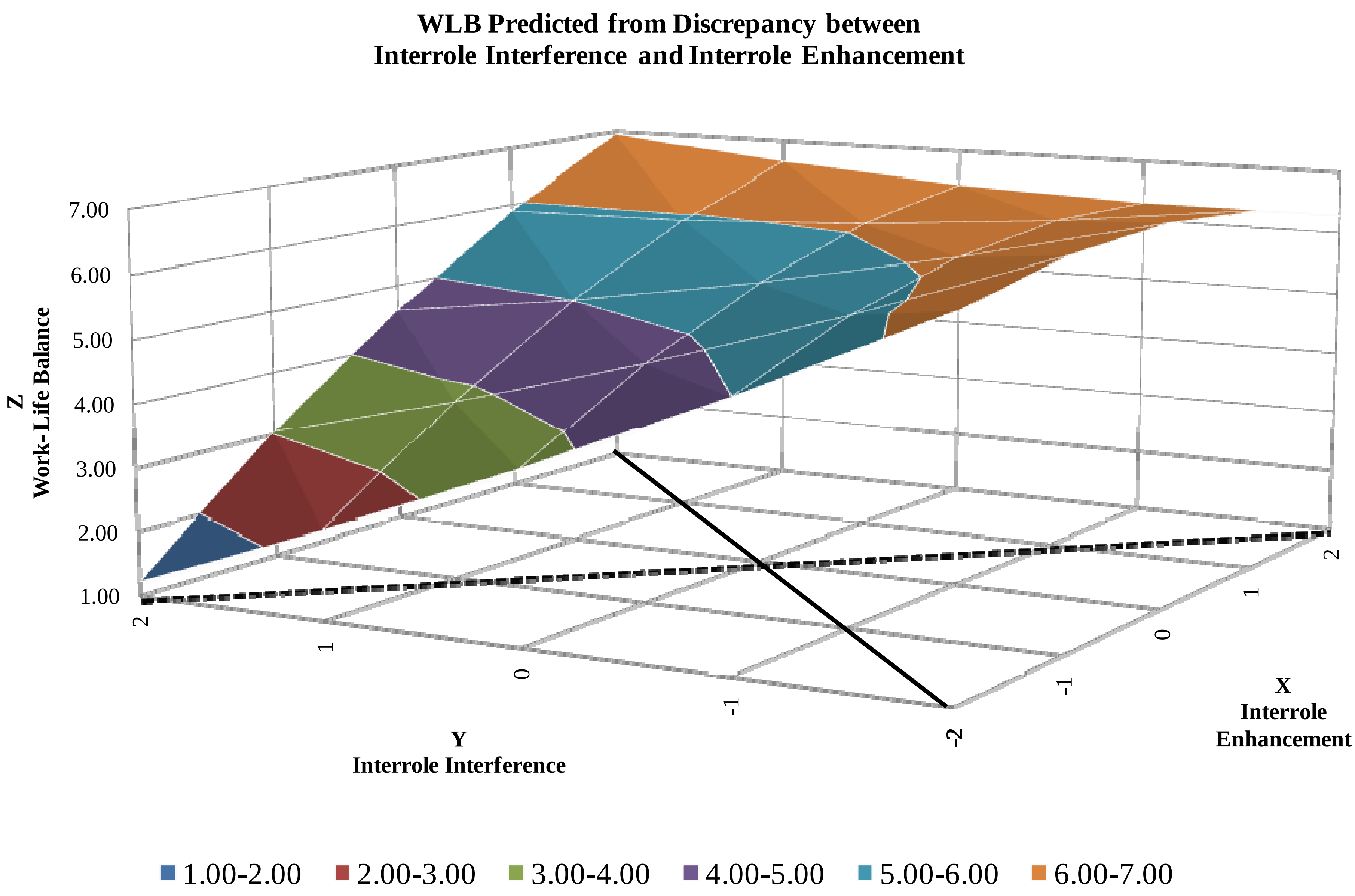
- Calculate a range of predicted values for the outcome, using the polynomial regression equation you obtained and plugging in a series of values depending on the range of your centered predictors:
  - For a 5-point Likert scale, use -2, 1, 0, 1, 2
  - For a 7-point Likert scale, use -4, -2, 0, 2, 4

Points to Plot

		X				
		-2	**-1**	**0**	**1**	**2**
	2	1.13	1.24	1.45	1.76	2.16
	1	3.11	3.17	3.33	3.60	3.95
Y	**0**	4.63	4.65	4.77	4.99	5.30
	-1	5.72	5.69	5.76	5.93	6.20
	-2	6.35	6.28	6.30	6.43	6.65



WLB Predicted from Discrepancy between NWWI and WNWE (Baby Boomers)
Z
Work-Life Balance
7.00
6.00
5.00
4.00
3.00
2.00
1.00
2
1
0
-1
-2
Y
NWWI
2
1
0
-1
-2
X
WNWE
1.00-2.00
2.00-3.00
3.00-4.00
4.00-5.00
5.00-6.00
6.00-7.00



WLB Predicted from Discrepancy between
Interrole Interference and Interrole Enhancement
7.00
6.00
5.00
4.00
3.00
2.00
1.00
Z
Work-Life Balance
2
1
0
-1
-2
Y
Interrole Interference
X
Interrole
Enhancement
1.00-2.00
2.00-3.00
3.00-4.00
4.00-5.00
5.00-6.00
6.00-7.00

Interpreting RSM Graphs

- The line reflecting X = Y (solid black line in the figures) reflects agreement
- The line reflecting X = -Y (dashed line in the figures) reflects discrepancy
- Can discuss the surface in terms of the relationships between your X and Y variables (along both lines of agreement and discrepancy)

Research & Practice Recommendations

- Difference scores may still be useful depending on your needs
  - Not because they "make intuitive sense" (Peters et al., 1993)
- Reframe research questions and consider hypothesizing regarding more direct effects, than indirect relationships
  - Maybe data can show us how variables are related better than we can with complex cognitive algebra (Peters et al., 1993)
- Stop using diffscores for measuring change – collect more than two points of data and a whole host of other analytical techniques is available to you that are much more appropriate and less contentious

Tangible Resources

- From Shanock et al. (2010)
  - See the article in *Journal of Business and Psychology*
  - Cunningham (this session), revised Excel workbook
- From Jeff Edwards
  - For RSM plotting: http://public.kenan-flagler.unc.edu/faculty/edwardsj/downloads.htm
  - For SPSS: http://public.kenan-flagler.unc.edu/faculty/edwardsj/resources.htm

Selected References

- Bedeian, A. G., & Day, D. V. (1994). Concluding statement. *Journal of Management, 20*(3), 695-698.
- Cafri, G., van den Berg, P., & Brannick, M. T. (2010). What have the difference scores not been telling us? A critique of the use of self-ideal discrepancy in the assessment of body image and evaluation of an alternative data-analytic framework. *Assessment, 17*(3), 361-376.
- Cronbach, L., & Furby, L. (1970). How we should measure 'change': Or should we? *Psychological Bulletin*, *74*(1), 68-80.
- Edwards, J. R., & Parry, M. E. (1993). On the use of polynomial regression equations as an alternative to difference scores in organizational research. *Academy of Management Journal, 36*, 1577-1613.
- Edwards, J., & Van Harrison, R. (1993). Job demands and worker health: Three-dimensional reexamination of the relationship between person-environment fit and strain. *Journal of Applied Psychology*, *78*(4), 628-648.
- Edwards, J. R. (2002). Alternatives to difference scores: Polynomial regression analysis and response surface methodology. In F. Drasgow & N. Schmitt (Eds.), *Measuring and analyzing behavior in organizations: Advances in measurement and data analysis* (pp. 350-400). San Francisco: Jossey-Bass.
- Edwards, J. R. (2001). Ten difference score myths. *Organizational Research Methods, 4*(3), 265-287.
- Edwards, J. R. (1995). Alternatives to difference scores as dependent variables in the study of congruence in organizational research. *Organizational Behavior and Human Decision Processes, 64*(3), 307-324.
- Edwards, J. R. (1994). Regression analysis as an alternative to difference scores. *Journal of Management, 20*(3), 683-689.
- Griffin, D., Murray, S., & Gonzalez, R. (1999). Difference score correlations in relationship research: A conceptual primer. *Personal Relationships, 6, 505-518.*
- Harris, M. M., Anseel, F., & Lievens, F. (2008). Keeping up with the Joneses: A field study of the relationships among upward, lateral, and downward comparisons and pay level satisfaction. *Journal of Applied Psychology, 93(3)*, 665-673.
- Johns, G. (1981). Difference score measures of organizational behavioral variables: A critique. *Organizational Behavior and Human Performance, 27,* 443-463.
- Peter, J. P., Churchill, G. A., & Brown, T. J. (1993). Caution in the use of difference scores in consumer research. *Journal of Consumer Research, 19*, 655-662.
- Tisak, J., & Smith, C. S. (1994). Defending and extending difference score methods. *Journal of Management, 20*(3), 675-682.
- Shanock, L. R., Baran, B. E., Gentry, W. A., Pattison, S. C., & Heggestad, E. D. (2010). Polynomial regression with response surface analysis: A powerful approach for examining moderation and overcoming limitations of difference scores. *Journal of Business and Psychology, 25*(4), 543-554.