Congruence research in behavioral medicine: methodological review and demonstration of alternative methodology

L. Alison Phillips

Received: May 20, 2011 / Accepted: January 25, 2012 / Published online: February 18, 2012
© Springer Science+Business Media, LLC 2012

Abstract Researchers in behavioral medicine are using methods to study the effect of congruence between two predictors (e.g., patient and provider preferences for patient-decision making) on outcomes (e.g., patient satisfaction and adherence) that may compromise the validity of their results and conclusions. The current paper reviews the methods used in behavioral medicine for the study of congruence effects and the problems associated with those methods—namely, that difference-score-based and artificial-group-based methods increase the risk of both Types I and II error and constrain the theoretical hypotheses that researchers are able to assess. The current paper explains and demonstrates a possible alternative method, polynomial regression, that may be used in some instances of congruence research and that avoids many of the problems of currently used methods; the current paper provides the first analysis of behavioral-medicine data using this method. Methodological advancement in health-related congruence research can help advance theory and optimize intervention-design.

Keywords Methodology · Congruence/concordance · Polynomial regression · Difference scores · Doctor-patient interactions

Introduction

Researchers in behavioral medicine focus on improving patient health through improving the care patients receive during the medical encounter with a provider; specific foci include the degree to which providers share information with the patient (e.g., Bosworth et al., 2009), share decision-making responsibilities with the patient (e.g., Joosten et al., 2008), and show empathy to the patient, both verbally and nonverbally (e.g., Roter et al., 2006) to predict patient satisfaction, adherence, and health outcomes following the medical encounter. Recently, although the topic has been studied for decades (e.g., Stewart et al., 1979), many scientists in the field have contributed to the research above with studies on the patients’ preferences for shared information, shared decision-making, empathy, and other qualities of the provider; specifically, researchers have shown how the congruence between patients’ preferences and their actual experiences may better predict outcomes than do the overall levels of these factors (Christensen et al., 2010; Cvengros et al., 2009, 2007). This recent “congruence research” has contributed theoretically to behavioral medicine, and its importance and contemporaneity is evident in the movement towards developing interventions to either match patients with providers who share or meet the patient’s preferences or train providers to first identify patient preferences and then tailor their behavior to the patient’s preferences (Christensen et al., 2010).

Before implementing interventions to match patients and providers, however, the field would benefit from first
verifying results that may be compromised by many congruence methods (i.e., difference scores and artificially categorized congruence-groups) and from exploring more theoretically complex congruence relationships that are not assessable using these methods (Edwards, 1994; Johns, 1981; MacCallum et al., 2002). Researchers within behavioral medicine have recognized problems with difference scores for the study of topics other than congruence (e.g., indirect comparative risk judgments; Ranby et al., 2010), but use of difference scores is still wide-spread within behavioral-medicine congruence research. Researchers in other fields, such as industrial psychology, have used alternative methods, such as those presented in the current paper (e.g., Kalliath et al., 1999; Shanock et al., 2010; Yang et al., 2008).

The purpose of the current paper is: (1) to review commonly used methods for assessing congruence hypotheses and their statistical and interpretational problems and (2) to introduce and demonstrate an alternative methodology that better assesses the congruence relationships inherently implied by currently used methods and allows assessment of congruence relationships that cannot be assessed by currently used methods—relationships that could affect how field interventions are designed and implemented in medical practice. The alternative methodology includes polynomial regression, which was primarily developed in seminal work by Edwards (1994, 2002), Edwards and Parry (1993), and response surface methodology, which may be used to help interpret some results of polynomial regression (Myers et al., 2009). Lastly, since the use of polynomial regression is not limited to just patient-provider interactions, a final aim of the current paper is to spark researchers’ interest in using polynomial regression to assess congruence hypotheses in widely varied areas of behavioral medicine research. Syntax files with detailed instructions for analysis in SPSS and Excel are provided in the online supplementary material to facilitate researchers’ use of the alternative methods. Drawbacks of polynomial regression and the benefits of other methods are lastly discussed.

The current paper is relevant for any researcher interested in the effect of congruence between two commensurate (measured on the same scale) predictors on a continuous outcome. The predictors may be either from the perspective of two different individuals rating the same content-domain (such as the patient and the provider, or the patient and his/her partner) or from the perspective of one person rating two related content-domains (such as a patient’s expectations and experiences). Relevant topics in behavioral medicine include the following, which have been used to predict patient satisfaction, adherence, and/or health outcomes: patients’ preferences for and providers’ actual sharing of medical information, display of empathy, and/or communication style during the medical encounter (Krupat et al., 2000, 2001; Cvengros et al., 2009); patients’ preferences for and perceived amount of participation in decision-making (Jahng et al., 2005); patients’ and their spouses’ perceptions of the patient’s level of control over his/her illness and treatment (Christensen et al., 2010; Sterba et al., 2008); and patients’ and providers’ beliefs about whether the patient’s illness is primarily psychological or physical (Greer & Halgin, 2006).

Overview of currently used methods

The most commonly used methods for studying congruence-hypotheses in behavioral medicine are described below, along with examples from the literature. Each has problems that researchers, reviewers, and readers should be aware of as they conduct and interpret congruence research. Problems specific to each method are described along with those methods, but most of the problems are shared between methods and are therefore described in the subsequent section.

Congruence indices based on the difference score

Algebraic difference score

The algebraic difference score is formed by subtracting one interval- or ratio-scaled predictor variable from another. The difference score is treated as a continuous/interval variable itself, whose zero-value represents strict congruence and off-zero values represent varying degrees of incongruence. Researchers use the algebraic difference score to obtain both absolute- and relative-effect size estimates for the effect of congruence on the outcome, and/or they use it to interpret the directional relationship between congruence and the outcome. For example, Greer and Halgin (2006) took the difference between patients’ and physicians’ beliefs regarding the degree to which the patient’s symptoms were psychological in nature (versus medical/physical) and concluded from correlations that there was no effect of belief-congruence on patients’ ratings of physician-cooperativeness but that “physicians were more likely to characterize the patient as difficult when they disagreed with the patient and believed that the presenting symptoms were more psychological in nature” (pp. 280–281). Krupat et al. (2001) took the difference between patients’ and physicians’ preferences for the amount of information- and power-sharing engaged in by physicians during the medical encounter (Patient-Practitioner Orientation Scale, PPOS; Krupat et al., 2000) and used it in a regression, concluding that patient-physician belief-congruence regarding the information and power sharing of the physician resulted in greater trust in and
endorsement of the physician but not greater patient satisfaction.

**Absolute-value and squared-difference difference scores**

In order to assess the effect of the “directionless degree” of incongruence on an outcome variable, some researchers have taken the absolute value or squared the values of the algebraic difference score and used the transformed congruence-index as a continuous predictor in correlation- or regression-based analyses. If the absolute-value or squared-difference difference score is found to be a significant predictor, with or without the component factors included as predictors, researcher have concluded that the degree of incongruence is related to the outcome. The interpretation of correlations or the regression coefficient for this type of congruence index is intuitive, because unit increases in the degree of incongruence are qualitatively equivalent, unlike unit increases in the absolute value difference score (for which negative, zero, and positive values represent qualitatively distinct changes in the relationship between the two predictors). However, unique to absolute-value difference scores is the problem that they do not truly represent “directionless” degree of incongruence if the number of positive and negative values of the algebraic difference score are unequal. The more disproportionate the number of positive/negative values becomes, the more the absolute value of those scores really represents the algebraic difference score, which poses theoretically distinct congruence hypotheses from what the researcher assumes he/she is testing (Edwards, 1994).

Sterba et al. (2008) used an absolute-value difference score to test the effect of degree of incongruence between women’s and their husbands’ perceptions of the women’s control over their rheumatoid arthritis on the women’s psychological adjustment to the condition. Using both correlations and regression, they concluded that women’s psychological adjustment was poorer the greater the degree of discrepancy between women’s and their husbands’ ratings. Sewitch et al. (2003) used absolute-value difference scores in correlation analyses with patient satisfaction to provide criterion validity to the items in their scale regarding patient and physician discordance on several content-dimensions.

**Profile similarity indices**

Profile similarity indices are congruence indices created from combining predictor variables from more than one latent construct into one variable. A type of profile similarity index used in recent behavioral medicine research is a scale (average or sum) of difference-score variables: Sewitch et al. (2003) created absolute-value difference scores between patients’ and physicians’ ratings on ten different domains (five regarding the patient’s health status, such as physical functioning and emotional well-being, and five regarding qualities of the medical visit, such as discussion of personal issues and patient satisfaction with the visit), which were combined into one score to represent overall patient-physician discordance.

Profile similarity indices are treated and interpreted the same way as either the algebraic or the absolute-value difference score in statistical analyses, i.e., as continuous variables representing direction and/or degree of incongruence, depending on the way the component constructs are combined. However, since multiple constructs or dimensions are combined into one index, the interpretation and conclusions made by researchers are limited to the overall effect of congruence rather than the domain-specific effects on the outcome. Profile similarity indices compound the disadvantages of single difference-scores (Edwards, 1993).

**Congruence-group methods**

Researchers have used group-based analyses, sometimes in order to avoid the problems with difference-score indices. Krupat et al. (2000) used a three-group method by artificially trichotomizing patients’ and practitioners’ scores separately on the PPOS and using them to place each dyad into one of the three congruence groups (incongruent X > Y, congruent, and incongruent Y > X). They then artificially dichotomized the outcome variable, patient satisfaction, and conducted a Chi-Square test to determine if a greater proportion of satisfied patients were in the congruent group compared to in either of the incongruent groups and if there was a difference in proportion of satisfied patients when the provider was more patient-centered than the patient or vice versa.

Researchers have also kept the predictors separate, labeling each component either high or low on the construct, and then using two-way ANOVAs to analyze the effect of congruence on the outcome. For example, Krupat et al. (2000), labeled patients as either “patient-centered” and “doctor-centered” based on a median split of their scores on the PPOS and gave them one of two medical encounter scenarios: one depicting a patient-centered doctor and the other a doctor-centered doctor, and had them rate how satisfied they would be if the doctor was their own. The authors concluded from the significant interaction term that congruence was important for patient satisfaction.

**Combination methods: creating congruence-groups from difference scores**

Many researchers have used a combination of difference-score based congruence indices and congruence-group methods to intuitively evaluate hypotheses regarding the
effect of incongruence direction and degree on the outcome of interest. To analyze the effect of incongruence degree on an outcome, researchers have artificially dichotomized the algebraic or absolute value difference score by placing zero-valued scores into the congruence group and all non-zero valued scores into the incongruence group, and this variable is typically used in regression analyses (e.g., Sterba et al., 2008). To assess the effect of direction of incongruence on the outcome(s), researchers have split the algebraic difference score into three groups (by tertiles) to represent congruence and directional incongruence between the predictors and then conducted ANOVAs and pairwise comparisons of group means on the outcome. Cvengros et al. (2007) assessed the congruence between (1) patient preferences for patient-centeredness of a physician and physician ratings of how patient-centered a physician should be and (2) patient-beliefs about and physician-ratings on the patient’s control over an illness to predict patient satisfaction and adherence. Cvengros et al. (2009) assessed the congruence between patient preferences for and their reports of the practitioner’s degree of information-sharing, behavioral involvement of the patient, and socioemotional support to predict patient satisfaction, patient-reported adherence, and HbA1c. Christensen et al. (2010) assessed the congruence between patients’ and physicians’ ratings of the patients’ control over his/her illness to predict medication refill adherence and diastolic blood pressure readings. When significant differences between congruence groups were found for the above outcomes, it was consistently the group representing the providers’ ratings being lower than the patients’ ratings (e.g., practitioner was less patient-centered than the patient preferred) that scored more poorly than either the group representing congruence or the group representing practitioner’s ratings being greater than the patient’s ratings (e.g. practitioner was more patient-centered than the patient preferred). In these cases, the authors concluded that both congruence and the direction of incongruence were important for the outcome, but as explained in the subsequent section, this difference in groups may be due to only one of the components (e.g., the patients’ beliefs) and not an effect of (in)congruence.

### Statistical and interpretational problems of commonly used methods

In general, the problems for congruence research caused by currently used methods are a result of those methods representing a naturally three-dimensional relationship.
between the two predictors and the outcome as an artificial two-dimensional relationship between a congruence-index (such as a difference score or a congruence-group variable) and the outcome; see Fig. 1a and b for an illustration of the algebraic difference score collapsing a naturally three-dimensional plane. Such congruence indices inherently confound statistical information and theoretical meaning for both predictors, which increases the risk of both Type I (e.g., concluding congruence is important for the outcome when it is not) and Type II (e.g., failing to detect a significant congruence-effect on the outcome) errors (Edwards, 2001; Johns, 1981; MacCallum et al., 2002). These fundamental problems are briefly explained here, and an empirical demonstration is given in the subsequent section, but the reader is referred to methodological reviews on the subject for more in-depth coverage (see Cronbach, 1958; Edwards, 1994; Johns, 1981; MacCallum et al., 2002; Wall & Payne, 1973).

Any method utilizing a difference score increases Type II error risk (and, equivalently, attenuates effect sizes) because of the constraints difference scores impose on the regression coefficients of the congruence predictors, which cause the loss of statistical information unique to each predictor (Johns, 1981). Table 1, adapted with permission from Edwards (1994), lists the constraints implied by three commonly used difference scores. The constrained equations in Table 1 are the linear regression models tested by researchers who use difference scores as predictors; the unconstrained equations are the models tested with polynomial regression and can be used to statistically test whether or not the constraints are accurate. Any method utilizing congruence-groups increases Type II error risk (and attenuates effect sizes) by losing statistical information of the continuous predictors as they are artificially categorized and/or split by sample- or scale-dependent values (Hunter & Schmidt, 1990).

The attenuation of effect-size estimates by these methods has led some researchers to claim that they provide conservative tests of hypotheses and are therefore stronger evidence for the significance of effects that are found in the data. However, in addition to increasing the risk of Type II error, these methods also increase risk of Type I error (concluding that congruence is important for the outcome when it is not), and so do not provide more conservative tests of hypotheses (Edwards, 2001; Maxwell & Delaney, 1993). Difference score methods, including congruence-group and combination methods, increase Type I error risk, because they confound the effects of the two component measures (e.g., the effect of physician gender separate from the effect of patient gender) on the outcome (Edwards & Parry, 1993). The relationship of the difference score or congruence-group variable to the outcome may be entirely explained by one of the predictor components, which is not typically checked by researchers who use difference scores methods (Edwards, 2001; Maxwell & Delaney, 1993).

These problems compromise theoretical advancement, because they may lead researchers either to keep a construct in their theoretical model that is not important for the outcome(s) of interest or to eliminate a construct from their model that is important for the outcome(s) (Edwards, 1994; Yang et al., 2008). Theoretical advancement also relies on researchers’ ability to do meta-analysis of congruence findings, which is compromised by the sample-dependent nature of difference-score methods and the fact that they

### Table 1 Regression equations for three difference-score models commonly used to test congruence hypotheses

<table>
<thead>
<tr>
<th>Model</th>
<th>Constrained equation</th>
<th>Unconstrained equation</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algebraic difference</td>
<td>[Z = b_0 + b_1(X - Y) + e]</td>
<td>[Z = b_0 + b_1X + b_2Y + e]</td>
<td>(b_1 = -b_2)</td>
</tr>
<tr>
<td>Absolute value</td>
<td>[Z = b_0 + b_1X - b_1Y + e]</td>
<td>[Z = b_0 + b_1X + b_2Y + b_3W + b_4WX + b_5WY + e]</td>
<td>(b_1 = -b_2)</td>
</tr>
<tr>
<td>Absolute difference</td>
<td>[Z = b_0 + b_1(X - Y)^2 + e]</td>
<td>[Z = b_0 + b_1X + b_2Y + b_3X^2 + b_4XY + b_5Y^2 + e]</td>
<td>(b_1 = 0)</td>
</tr>
</tbody>
</table>

The constrained equations are the regression models tested when each of the three types of difference scores is used as a predictor variable. The unconstrained equations are the polynomial regression models that may be used to test the accuracy of the constraints implied by the difference-score models. Adapted from Organizational Behavior and Human Decision Processes, volume 58, Jeffrey R. Edwards, The study of congruence in organizational behavior research: Critique and a proposed alternative, page 54, 1994, with permission from Elsevier.
yield inaccurate effect sizes. Lastly, problems associated with these methods do not allow the full range of relevant congruence hypotheses to be explored and may lead to suboptimal design of field interventions.

Demonstration of difference-score methods

Here, three commonly used difference-score methods are used to analyze the effect of congruence between two predictors on patient adherence. The data used in this demonstration are also used in the demonstrations of polynomial regression in subsequent sections; the data come from a larger study designed to assess patients’ illness and treatment representations (beliefs) and the role of the physician in addressing these beliefs for attaining better patient adherence. As part of this study, physicians (n = 34) were given a one-page questionnaire to fill out for each patient (n = 248) after the visit. Relevant to the current analyses, the physician was asked to give: (1) a rating of the patient’s general health (1 = Poor to 5 = Excellent); (2) an estimate of how the patient would rate his/her own health (1 = Poor to 5 = Excellent); and (3) ratings on several items regarding the degree to which the physician perceived patient agreement on the illness (e.g., its causes, associated symptoms) and prescribed treatment (e.g., appropriateness and likely effectiveness). The first two items are the predictor components, which were written to match a single item that is widely used to assess the patient’s self-rating of health and which has been shown to be highly predictive of patient morbidity and mortality (Idler & Benyamini, 1997). The third measure serves as one of the outcomes. The other outcome is patient adherence, which was measured 1 month after the medical visit with a patient self-report measure that has been validated in many medical contexts, the Medication Adherence Report Scale (Horne, 2004).

If the physicians’ estimates of the patient’s own rating of health is different from the physicians’ rating of the patient’s health, then this would indicate that the physician perceived some sort of disconnect between the patient’s objective health cues (including presenting symptoms and the patient’s medical history, etc., that is known by the physician) and the patient’s interpretation of these cues (e.g., overly optimistic or pessimistic, depending on the circumstances). Such incongruence might affect the physicians’ perceptions of agreement with the patient regarding the presenting illness and treatment, and such incongruence may also predict patients’ actual adherence to the prescribed treatment.

The algebraic difference score, the absolute-value difference score, and the tertile-split of the algebraic difference score (a combination method) were used to analyze exactly the same relationships, where X = the physician’s rating of the patient’s overall health, Y = the physician’s estimate of the patient’s self-assessed health, and Z = patient adherence. Figure 1a is the graph of the algebraic difference score results; the unstandardized regression coefficient on the difference score was .35, p < .001, and the $R^2 = .10$. The interpretation of this result typically given by researchers who use this method would be, “the effect of congruence on the outcome was significant”; the specific interpretation is that patient adherence increases as the physician’s estimate of the patient’s self-assessed health (X) approaches the physician’s rating of the patient’s health (Y) and continues to increase as X exceeds Y.

Figure 1c is the graph of the absolute value difference score results; the unstandardized regression coefficient on the difference score was .41, p < .001, $R^2 = .10$. The interpretation of this result is that patient adherence is maximized when the physicians’ ratings and estimates are equivalent (when X = Y) and decreases when either predictor is greater than the other.

Figure 1d is the graph of the combination method results; here the omnibus F-test was significant, $F(2,124) = 13.40$, p < .001, and post hoc Tukey tests indicated that only when X > Y was patient adherence significantly lower than for the other two congruence groups (the 95% confidence interval for the mean difference on patient adherence between the X > Y group and the X < Y group is $-1.10, -.02$ and between the X > Y group and X = Y group is $-.98, -.36$; these values are z-scores). The interpretation of this result is that patient adherence is similar when the physicians’ ratings and estimates are congruent and when the physician thinks the patient rates his/her health better than does the physician—adherence is only significantly lower when the physician thinks the patient rates his/her health as worse than does the physician.

All three, theoretically distinct, models were supported by the data—a problem which would not be detected by researchers who chose to use only one method and a problem that clearly demonstrates that these methods should not be used without checking their inherent assumptions, if at all. In the next section, an alternative analysis method, polynomial regression, is used to show that the algebraic difference score model (and not the absolute-value difference score model) is the accurate representation of the congruence relationship analyzed above. Polynomial regression is also used to assess the effect of congruence on the other outcome—physicians’ perceptions of agreement with the patient regarding the presenting illness and prescribed treatment, which cannot adequately be represented by any of the difference score models.
Alternative methodology: polynomial regression and response surface analysis

Polynomial regression allows for more accurate assessment of congruence hypotheses by avoiding the statistical and interpretational problems of difference-score-based methods; polynomial regression represents the predictor components as separate dimensions, thereby accurately reflecting the three-dimensional relationship that exists when two predictors are used to predict an outcome (traditionally X, Y, and Z, respectively; Edwards, 1994, 2002; Edwards & Parry, 1993; Shanock et al., 2010; Yang et al., 2008). That is, polynomial regression is conducted similarly to any regression analysis, but instead of the two predictor components being represented in single variable, they are left as separate variables and both entered into the regression model as predictors of the outcome; higher-orders of the variables (such as X², Y³, etc.) may also be added as predictors to better account for nonlinear relationships among the predictor variables and the outcome. Congruence hypotheses are then evaluated based on the relationships between the resultant regression coefficients, which may be plotted in a three-dimensional graph (a surface rather than a line). Unlike difference-score methods, polynomial regression preserves the absolute as well as the relative values of the two predictors so that perfect congruence between predictors is not just a single point but is a series of points for all values X and Y, where X = Y. Similarly, both direction and degree of incongruence are accurately captured by polynomial regression. A confirmatory approach is used when a researcher can specify expectations for these regression coefficients (such as the constrained relationships in Table 1); otherwise, an exploratory approach may be used.

Confirmatory polynomial regression

The confirmatory approach to polynomial regression is used when the researcher has a priori hypothesized that the congruence relationship between the commensurate predictors and the outcome fits a difference score model. Researchers who use difference scores or combination methods, are inherently assuming the same data structure as is tested with the confirmatory approach; therefore, a major benefit of the confirmatory approach is that it may be used to assess the validity of the difference-score assumptions—for example, by testing whether the constraints imposed by the difference score are accurate. If these assumptions are met, the data may be intuitively interpreted as they would be with a difference-score analysis. The confirmatory approach either supports or rejects an a priori hypothesized model.

Specific analysis steps

The specific steps required to conduct the analyses are described and demonstrated in this paper, but the mathematics behind the procedures have been developed and presented in great detail in other areas of psychology, to which the interested reader is referred (see: Edwards, 1994, 2002; Edwards & Parry, 1993; Kallith et al., 1999; Shanock et al., 2010; Yang et al., 2008).

The confirmatory approach is demonstrated in this section, but the following preparatory steps also apply to the exploratory approach described and demonstrated in subsequent sections: (1) measure both components on interval or ratio, commensurate scales; (2) verify the assumptions required for any multiple regression analyses (e.g., multivariate normality and reliability of measures; see Tabachnick & Fidell, 2007); (3) evaluate the presence of outliers and influential points, which may severely alter the regression coefficients of higher-order terms; and (4) scale-center both components and create all higher-order terms from these centered variables to facilitate analysis and interpretation of results.

With the confirmatory approach, a researcher may test two competing models (e.g., absolute versus algebraic difference models) or attempt to confirm one model. There are four criteria that must be met in order to find support for a model, otherwise the model is rejected: first, the unconstrained difference score model (see equations in Table 1) must explain significant variance in the outcome, and second, the individual terms in the model should be as expected in both significance and valence (see implied constraints in Table 1). The third criterion is that the constraints implied by the difference score model must hold true—or equivalently that the variance explained in the outcome must not significantly increase from the constrained to the unconstrained model. The last criterion is that no higher-order terms should explain significant incremental variance in the outcome to that explained by the unconstrained model. This criterion ensures that, even if the constraints are accurate, there are no more complex congruence relationships that better explain the data. Adequate statistical power is required to detect differences in $R^2$ between the constrained and unconstrained model and between the unconstrained model and the higher order terms (Edwards, 2002). If the confirmatory approach results in support of the model, then interpretation of the congruence relationship is intuitive and as implied by the model; for example, if the absolute-value difference model is supported by the polynomial regression, then one can conclude that the outcome is optimized with perfect congruence between predictors (and equally optimized for all absolute values of the congruent predictors) and is equally and negatively affected by incongruence in either direction.
If the confirmatory approach results in rejection of the model, then the exploratory approach can be used to assess the nature of any congruence relationships that may exist but that do not fit the difference score model(s).

Demonstration of confirmatory method

Presented here is a re-analysis of the fit between the empirical data and the algebraic and absolute value difference score models using the confirmatory approach to polynomial regression. Leverage values and Cook’s D statistic indicated that there were no data points that would disproportionately influence the higher-order terms. Multilevel modeling in SPSS indicated that the physician-level effect was non-significant for both outcomes (Wald Z—the default test in SPSS for tests of variance components in mixed-model analyses—for patient adherence = .44 and for physician perceptions of patient agreement = 1.90, \( p > .05 \)).

The results of the confirmatory approach for both outcomes for the algebraic and absolute-value difference score models, which are used to assess the four criteria required for support of a difference score model, are shown in Table 2.

**Algebraic difference score model**

(1) The unconstrained model explained significant variance in patient adherence (\( R^2 = .10, \ p < .01 \)) and in physician perceptions of patient-agreement (\( R^2 = .18, \ p < .001 \)). (2) The regression coefficients for \( X \) and \( Y \) were in the expected direction and had the expected significance for patient adherence (both coefficients were significant and in the opposite direction to each other) but not for physician perceptions of patient-agreement (only the coefficient on \( Y \) was significant). (3) The unconstrained model predicted significant incremental variance to the constrained model only for physician perceptions of patient-agreement (\( F(1,245) = 54.89, \ p < .001 \), not for patient adherence (\( F(1,245) = .84, \ p = .36 \)). This means that the constraint, \( b_1 = -b_2 \), was found to be inaccurate for physician perceptions of patient-agreement but accurate for patient adherence (both coefficients were significant and in the opposite direction to each other). (4) The higher-order terms predicted significant incremental variance to the unconstrained model for physician perceptions of patient-agreement (\( F(3,242) = 4.18, \ p < .01 \)) but not for patient adherence (\( F(3,242) = .83, \ p = .48 \)). All four criteria for support of the algebraic difference score model were met for patient adherence. Only one of the four criteria (the first) was met for physician perceptions of patient-agreement.

**Absolute-value difference score model**

(1) The unconstrained model explained significant variance in patient adherence (\( R^2 = .15, \ p < .01 \)) and in physician perceptions of patient-agreement (\( R^2 = .21, \ p < .001 \)). (2) The regression coefficients were as expected for patient adherence but not for physician perceptions of patient-agreement (the coefficient on \( X \), \( b_1 \), was not significant). (3) The constraints were accurate for patient adherence (\( F(4,242) = 1.75, \ p = .14 \)) but not for physician perceptions of patient-agreement (\( F(4,242) = 13.14, \ p < .001 \)). (4) The set of higher-order terms explained significant variance.

---

**Table 2** Results of the confirmatory analysis for the algebraic and absolute-value difference score models for two outcomes: patient adherence and physician perceptions of patient-agreement

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Constrained model</th>
<th>Unconstrained model</th>
<th>( F_c )</th>
<th>( F_h )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( X - Y )</td>
<td>( R^2 )</td>
<td>( X )</td>
<td>( Y )</td>
</tr>
<tr>
<td>Patient adherence</td>
<td>-.35***</td>
<td>.10***</td>
<td>-.33**</td>
<td>.39***</td>
</tr>
<tr>
<td>Physician perceptions of patient-agreement</td>
<td>-.05</td>
<td>.001</td>
<td>.10</td>
<td>.27**</td>
</tr>
<tr>
<td></td>
<td>(</td>
<td>X - Y</td>
<td>)</td>
<td>( R^2 )</td>
</tr>
<tr>
<td>Patient adherence</td>
<td>-.41***</td>
<td>.10***</td>
<td>-.47**</td>
<td>.48**</td>
</tr>
<tr>
<td>Physician perceptions of patient-agreement</td>
<td>-.32**</td>
<td>.04**</td>
<td>-.10</td>
<td>.42**</td>
</tr>
</tbody>
</table>

Values in the cells for \( X - Y \), \( |X - Y| \), \( X \), \( Y \), \( W \), \( WX \), and \( WY \) are unstandardized regression coefficients with \( X \) representing physicians’ ratings of the patient’s health, \( Y \) representing physicians’ estimates of how the patient would rate his/her own health, and \( W \) is a dummy code equal to 0 when \( X > Y \), 1 when \( X < Y \), and randomly assigned 0 or 1 when \( X = Y \). \( F_c \) is the F-ratio that tests the accuracy of the constraints imposed by the particular difference score model with degrees of freedom equal to \( (1, N - 3) \) for the algebraic difference score model and \( (4, N - 6) \) for the absolute value difference score model. \( F_c \) is the F-ratio that tests the difference in variance explained in the outcome by the unconstrained model and the higher-order terms for that model, with degrees of freedom equal to \( (3, N - 6) \) for the algebraic difference score model (for which the higher order terms are \( X^2 \), \( XY \), and \( Y^2 \)) and \( (6, N - 12) \) for the absolute value difference score model (for which the higher order terms are \( X^2 \), \( XY \), \( Y^2 \), \( WX \), \( WXY \), and \( WY^2 \)).

* \( p < .05 \); ** \( p < .01 \); *** \( p < .001 \)
incremental variance to the unconstrained model for patient adherence ($F(6,236) = 2.43, p < .05$) but not for physician perceptions of patient-agreement ($F(6,236) = 1.39, p = .22$). Three of the four criteria were met for patient adherence, and two were met for physician perceptions of patient-agreement. Therefore, the absolute-value difference score model is rejected for both outcomes. Since the algebraic difference score model was supported for patient adherence, no further analysis is required. Polynomial regression accurately determined the theoretical model that was best supported by the data.

Exploratory polynomial regression and response surface methodology

The exploratory approach is used when the researcher does not have an a priori hypothesis regarding the underlying model of the congruence relationships between the predictor components and the outcome or when the confirmatory approach results in rejection of the difference score model(s). The exploratory approach, like any other exploratory analysis (e.g. exploratory factor analysis), is a model-fitting approach that requires replication or cross-validation before hypotheses are built and tested from the results (see, e.g., Edwards & Harrison, 1993). In conjunction with response surface methodology, polynomial regression can be used to evaluate congruence hypotheses that behavioral medicine researchers have so far not explored and which may have implications for intervention design and implementation.

The exploratory method involves running polynomial regressions of increasing order (e.g., linear, quadratic, cubic, quartic) until a set of the same order terms no longer predicts significant incremental variance in the outcome. The final, best-fitting model is the model of highest-order that predicts incremental, significant variance in the outcome. If the final model is linear, then interpretation is intuitive. If however, the final model is quadratic or higher-order, interpretation of the results is more difficult and benefits from response surface methodology.

Response surface methodology is a technique that may be used to simplify the interpretation of three-dimensional surfaces—and therefore to evaluate congruence hypotheses (Edwards, 2002; Myers et al., 2009). Mere visual
representation of the response surface can greatly facilitate a researcher’s understanding of the effect of congruence/incongruence on the outcome, but statistical evaluation of surface characteristics is also possible in order to show support for congruence hypotheses. Since the majority of congruence relationships will best fit either a linear or quadratic surface (Edwards & Parry, 1993), and linear surfaces are intuitively interpreted, the focus of the current description will be on quadratic surfaces—specifically on concave (dome-shaped) surfaces for reasons of brevity and clarity, although the same characteristics can be assessed for convex (bowl-shaped) or saddle surfaces. Concave surfaces are most relevant for health outcomes that researchers want to maximize, such as patient adherence or satisfaction; the concave surface can be used to evaluate whether congruence between the components maximizes the outcome and whether incongruence is detrimental for the outcome.

Figure 2a–d help illustrate a range of possible congruence relationships and their corresponding response surface characteristics that would have implications for intervention-design. Figure 2a illustrates a hypothetical example in which the outcome, such as patient adherence, is maximized when patients receive (predictor Y) exactly as much information from their provider as they desired (predictor X); Fig. 2b illustrates when patient adherence is maximized when patients receive only slightly more information than they desired; and Fig. 2c illustrates when patient adherence is maximized when patients receive any range of information in excess of what they desired. Results resembling Fig. 2b would indicate a different type of intervention than would be indicated by results resembling Fig. 2c. Figure 2b results would indicate a patient-tailored intervention, in which providers are trained to identify the patient’s desired amount of information before interacting with him/her or in which patients are matched with providers based on information-giving style; Fig. 2c results would indicate instead a much simpler intervention in which patient-tailoring is less important—providers could merely be trained to err on the side of giving too much rather than too little information.

Polynomial regression and response surface methodology can detect all of these congruence relationships and others, using response surface characteristics, including the principal axes and the slope and curvature of the surface along the lines of congruence (when Y = X) and incongruence (when Y = −X). The first principal axis of a concave surface is the line in the XY-plane along which the outcome, or Z, has the lowest downward curvature—that is, it represents the values of X and Y for which the outcome is maximized and decreases the least. Therefore, if the goal of a researcher is to find the values of X and Y (e.g., the values of patients’ preferences for shared decision-making and the values of patients’ experiences of shared decision-making) along their entire scales that optimize Z (e.g., patient satisfaction), then the researcher is interested in the first principal axis and whether it parallels the line of congruence (i.e., when Y = X). In Fig. 2a these lines are exactly the same (parallel and with the same intercept); the solid line in the XY-plane is the line of congruence and also the first principal axis of the surface. A researcher can test to see if the first principal axis is shifted from the line of congruence such that they are parallel but with a constant difference between X and Y (see the note on significance tests below and the online supplementary material for syntax). Figure 2b shows such a congruence relationship (the dark/bold line in the XY-plane is the first principal axis; the thin solid line is X = Y).

The second principal axis of a concave surface is the line in the XY-plane along which the outcome decreases the fastest (i.e., along which the downward curvature of the surface is maximized; e.g., the combined values of patient and provider attitudes that are associated with the steepest decrease in patient satisfaction). If the second principal axis parallels the line of perfect incongruence (Y = −X), then this indicates that the outcome does indeed decrease the fastest when X is maximally discrepant from Y (as above, with a possible constant discrepancy between them). All Fig. 2a–d have a second principal axis equivalent to Y = −X.

Once the overall effect of (in)congruence has been established, additional information regarding the uniformity of the effect can be assessed with the slope and curvature of the surface along the line of congruence (Shanock et al., 2010). If the slope of the surface along the Y = X line at the origin is zero, then the surface is flat along the line of congruence, and the outcome is uniformly maximized for all values of X and Y; in terms of congruence effects, this means that the outcome is maximized when X and Y are congruent for all values of X and Y (e.g., Fig. 2a–c). If the slope is positive, then the outcome increases for greater values of X (and therefore Y); for example, patient satisfaction may be higher overall when patients’ and providers’ attitudes match than when they do not match and highest overall when both patients’ and providers’ attitudes are high versus low on the attitude scale (Fig. 2d illustrates such a relationship). If the slope is negative, then the opposite is true (the outcome is at its highest when X and Y are lower rather than higher). The slope is calculated at a single point, or coordinate of X and Y, because the slope of a surface is really the slope of a tangent line to the surface at that point. The origin is when X and Y both equal 0, and since the variables are centered at their mean-values, it is the point where both X and Y are at their mean-values. The slope of the surface at the origin
(when the centered predictors, X and Y = 0) along the line of congruence is calculated as \( b_1 + b_2 \), or the sum of the regression coefficients on X and Y.

To evaluate hypotheses regarding the direction of incongruence, a researcher can assess the slope of the surface along the line of incongruence. If the slope is negative, it means that the outcome/surface is higher when \( Y > X \) than when \( X > Y \) (Shanock et al., 2010), as in the hypothetic example in Fig. 2c. If the slope is positive, the opposite directional effect is supported. If the slope is flat, then the effect on the outcome is equal for both directions of discrepancy, as is depicted in the Fig. 2a, b, and d.

Finally, assessing the curvature of the surface along lines of interest indicates whether the effect of (in)congruence on the outcome is linear or non-linear—that is, whether the effects determined with calculations of the slope, described above, are linear or non-linear. For example, the increase in the outcome with increasing X and Y along the line of congruence, in Fig. 2d, is linear, not curvilinear; in terms of congruence effects, this means that unit increases and decreases in X and Y have a constant effect on the outcome; if the curvature was non-zero (for a concave surface, the curvature would be negative, if non-zero), this would mean that the unit increases in X and Y had a decreasing effect on the outcome (increases in X and Y from the origin—their average values—would have less and less benefit for the outcome, and decreases in X and Y from the origin would have less and less detrimental effects on the outcome). Similarly, if the curvature along the line of incongruence is significantly non-zero (negative in the case of a concave surface), then the effect of incongruence on the outcome decreases as the discrepancy between X and Y increases (the effect is not a constant decrease in the outcome as the discrepancy between X and Y increases, which would occur if the curvature of the surface along the line of incongruence was flat, or not significantly different from zero).

A note on the significance tests for the parameters: for parameters that are calculated as linear combinations of variables, such as the slope of the surface along the line of congruence mentioned above, linear contrasts may be used to test their significance (see the syntax in the online supplementary material). For parameters that are calculated from non-linear combinations of study variables, linear contrasts are not useable for testing their significance. Instead, bootstrapping may be used. Bootstrapping is a method by which parameter estimates are made from a large number of sub-samples of the original data. These estimates can be used to calculate confidence intervals for the parameters of interest—such as the slopes of the principal axes. Bootstrapping can also be used to calculate the confidence intervals for parameters that are linear combinations of variables (or to verify the results of linear contrasts) and has the advantage that it avoids violating distributional assumptions, such as the normality of the outcome variable (Efron, 1981). However, since the accuracy of the response surface shape depends on the data meeting distributional assumptions, polynomial regression should not be used if the assumptions are violated, as stated previously.

Demonstration of exploratory method

The exploratory method was used for physician perceptions of patient-agreement, because both the algebraic and the absolute-value difference score models were rejected for this outcome. The hierarchical regression testing increasing-order polynomial models yielded the following results: the linear and quadratic sets of terms significantly predicted the outcome \( (R^2 \text{ change} = .18 \) and \( .04, p < .001 \) and \( < .01, \) respectively), but the cubic and quartic sets did not \( (R^2 \text{ change} = .007 \) and \( .012, p = .67 \) and .58, respectively). Therefore, a quadratic model was the best-fitting model. Figure 2d is the graph of the surface, plotted in Excel from the regression coefficients \( b_0, b_1, b_2, b_3, b_4, \) and \( b_5, \) which correspond to the constant, X, Y, \( X^2, XY, \) and \( Y^2, \) respectively (see the online supplementary material).

The first principal axis, or the line along which the outcome was maximized overall (decreased the least), had a slope of \( .84; \) a bootstrap analysis of the regression coefficients with 10,000 samples yielded a 95% confidence interval for this value that contained 1 [using the percentile method recommended by Edwards (2002), due to typically skewed bootstrap samples]. Since the slope was not significantly different from 1, this indicates that the first principal axis is parallel to the line of congruence and therefore that congruence maximizes the outcome, as hypothesized. The second principal axis, or the line along which the outcome decreased the fastest, had a slope of \(-1.19; \) the bootstrap 95% confidence interval using the percentile method contained \(-1, \) indicating that the second principal axis is parallel to the line of incongruence. This means that the outcome is most negatively affected when \( Y = -X \) for all values of X and Y.

The slope and curvature of the surface at the origin along the line of congruence \( (Y = X) \) were \( .31 (p < .001) \) and \(.06 (p = .17), \) respectively. The significant positive slope means that physicians perceived greater agreement with the patient regarding the illness and treatment (i.e. the outcome was higher) when both the physicians’ ratings of the patient’s health and physicians’ estimates of the patient’s own rating of health were high (very good, excellent, e.g.) compared to low (poor, fair, e.g.). The non-significant curvature means that this increase in perceived agreement with higher ratings compared to lower ratings of health was linear rather than curvilinear. The slope and
curvature of the surface at the origin along the line of incongruence ($Y = -X$) were $-0.05 (p = .77)$ and $-0.80 (p < .01)$, respectively. The non-significant slope indicates that the effect of incongruence on the outcome is uniform, or equal, when $X > Y$ as when $Y > X$. The significant negative curvature indicates that the negative effect of incongruence on the outcome exponentially increases as the discrepancy between $X$ and $Y$ increases.

Taken together, these results indicate that physicians’ perceptions of patient agreement regarding the illness and its prescribed treatment were: (1) maximized overall when physicians’ estimate of the patient’s own rating of health was equal to the physicians’ rating of the patient’s health, (2) maximized specifically when the estimates and ratings are congruent and high (e.g., ‘very good’ health ratings) rather than low (e.g., ‘poor’ health ratings), (3) minimized overall, the more discrepant were the estimates and ratings, and (4) equally negatively affected when estimates > ratings as when ratings > estimates. In simplified terms, congruence was important for the outcome as were the overall levels of the predictors, the degree of incongruence was important for the outcome, but the direction of incongruence was not. While these results may not indicate an intervention design different from that indicated by Fig. 2a, the positive slope of the surface along the line of congruence does indicate that there may be an individual difference variable (perhaps a personality difference in preference for interacting with medical providers) that predicts variance in the outcome—a variable that one would want to assess in an intervention study in order to optimally assess the success of the intervention.

**Discussion**

The current paper has important implications for behavioral medicine research. First, existing congruence data can be re-analyzed in order to validate the findings already published in the literature; given the number of datasets that exist, it may be possible to cross-validate congruence relationships with polynomial regression by combining these datasets into one body of evidence. Second, existing and future congruence data can be analyzed to assess more complex relationships than have so far been investigated, which would help researchers to optimally design interventions to achieve the many goals of congruence research. The study of congruence is important for patient health care and may vastly benefit from a progression in methodology that allows for the valid evaluation of theoretical relationships between any two commensurate predictors and an outcome of interest.

Polynomial regression further allows for clustered data analysis, which is important for behavioral medicine research in which patients are often clustered within providers and/or providers are clustered within health facilities. As researchers have done in conjunction with the difference-score and congruence-group methods above, there are three general options with clustered data: first, the effect of the physician-level variable can first be ruled out (if it has no significant effect) and the polynomial regression can be run with the assumption of independence of patient cases (e.g., Greer & Halgin, 2006). Second, the physician-level effect (if significant) can be controlled for and adjusted coefficients may be plotted using standard hierarchical linear modeling or generalized estimating equations methodology (Cvengros et al., 2009; Krupat et al., 2001). Lastly, while no “how-to” guide exists for plotting response surfaces by second-order level, it is possible that each physician may have a surface distinguishable from the surfaces of other physicians (that is, if both component variables are patient-variables, otherwise the physician variable is constant); this would, however, require extremely large sample sizes to test (Hedeker et al., 1994).

Although polynomial regression avoids the major problems associated with currently used congruence methods in behavioral medicine, it has drawbacks of its own. Polynomial regression cannot be used to assess congruence as an outcome (see Edwards, 1995). Polynomial regression also assumes, as all standard regression methods do, that the independent variables are measured without error. Structural equation modeling (SEM) may be used to test congruence hypotheses without this assumption; however, the negatives of using SEM, including a requirement for much larger sample sizes, more sophisticated software, and understanding of more complex statistical models make polynomial regression more likely to be adopted by congruence researchers. Exploratory polynomial regression requires cross-validation of the fitted-model, just as does any other exploratory methodology (e.g. factor analysis).

Also, the methods currently being used by behavioral-medicine researchers are more intuitive and are more accessible to readers than polynomial regression, which are clearly reasons why many researchers would prefer the 2-dimensional methods to polynomial regression. However, the confirmatory approach to polynomial regression does not require cross-validation and makes exactly the same assumptions regarding the underlying data structure as do difference-score based and combination difference-score/congruence-group methods; confirmatory polynomial regression checks the assumptions that are inherent to these other methods, making it useful as a method to check these assumptions. Therefore, it is recommended that researchers who choose to use difference-score and/or congruence-group methods for their greater intuitiveness and simplicity
first utilize confirmatory polynomial regression to check the problematic assumptions of these methods. Since these assumptions are not often met (Edwards, 2002), exploratory polynomial regression affords one alternative method for analyzing the data that should not be analyzed with two-dimensional methods.

References


