

This article was downloaded by: [McMaster University]

On: 27 October 2014, At: 13:41

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Psychotherapy Research

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/tpsr20>

Expert reasoning in psychotherapy case formulation

Tracy D. Eells^a, Kenneth G. Lombart^b, Nicholas Salsman^c, Edward M. Kendjelic^d, Carolyn T. Schneiderman^e & Cynthia P. Lucas^f

^a University of Louisville, Department of Psychiatry and Behavioral Sciences, Louisville, Kentucky, USA

^b University of Massachusetts, Department of Psychology, Lowell, USA

^c Xavier University, Department of Psychology, Cincinnati, USA

^d VA Pittsburg Healthcare System, Neurobehavioral Program, Pittsburgh, Pennsylvania, USA

^e Private Practice, Louisville, USA

^f Private Practice, Nashville, USA

Published online: 15 Jan 2011.

To cite this article: Tracy D. Eells, Kenneth G. Lombart, Nicholas Salsman, Edward M. Kendjelic, Carolyn T. Schneiderman & Cynthia P. Lucas (2011) Expert reasoning in psychotherapy case formulation, *Psychotherapy Research*, 21:4, 385-399, DOI: [10.1080/10503307.2010.539284](https://doi.org/10.1080/10503307.2010.539284)

To link to this article: <http://dx.doi.org/10.1080/10503307.2010.539284>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Expert reasoning in psychotherapy case formulation

TRACY D. EELLS¹, KENNETH G. LOMBART², NICHOLAS SALSMAN³, EDWARD M. KENDJELIC⁴, CAROLYN T. SCHNEIDERMAN⁵, & CYNTHIA P. LUCAS⁶

¹University of Louisville, Department of Psychiatry and Behavioral Sciences, Louisville, Kentucky, USA; ²University of Massachusetts, Department of Psychology, Lowell, USA; ³Xavier University, Department of Psychology, Cincinnati, USA; ⁴VA Pittsburgh Healthcare System, Neurobehavioral Program, Pittsburgh, Pennsylvania, USA; ⁵Private Practice, Louisville, USA & ⁶Private Practice, Nashville, USA

(Received 1 May 2009; revised 31 October 2010; accepted 31 October 2010)

Abstract

Therapist reasoning in case formulation construction was investigated. Sixty-five psychodynamic or cognitive-behavioral therapists classified as experts, experienced, or novices generated “think aloud” formulations based on six standardized vignettes. Formulations were reliably transcribed, segmented into idea units, and content coded. ANOVA and sequential analysis compared formulation content and reasoning processes. Expert formulations contained more descriptive, diagnostic, inferential, and treatment planning information. They focused more on given and inferred symptoms, on adult relationship history, on inferred psychological mechanisms, on the need for further evaluation, and on plans to focus on treatment expectations and symptoms. They exhibited more forward (inferential) than backward (deductive) reasoning and, compared with non-experts, they exhibited more forward and backward reasoning. Results are discussed in terms of cognitive science models for expert problem solving and on implications for psychotherapy training, practice, and research.

Keywords: cognitive behavior therapy; psychoanalytic/psychodynamic therapy; process research; psychotherapist training/supervision/development; case formulation; forward reasoning

Psychotherapy case formulation is broadly recognized as an essential skill for psychotherapists (Binder, 2004; Caspar, Berger, & Hautle, 2004; Hersen & Porzelius, 2002). Bieling and Kuyken (2003) describe formulation as “the heart of evidence-based practice” (p. 53) and as occupying a “fundamental place in clinical psychology” (p. 53). Formulation, it is claimed, enhances psychotherapy effectiveness because symptoms and problems are understood and organized by a coherent theoretical structure (Benjamin, 2003).

The relatively small amount of research on case formulation provides limited support for its importance. Randomized, controlled studies comparing outcome in formulation-guided versus non-formulation-guided treatment are equivocal. Ghaderi (2006) assigned 50 patients with bulimia nervosa into either manual-based or formulation-guided cognitive-behavioral therapy. The latter group improved more on key outcome measures, although both groups achieved sustained improvements. Earlier, Schulte, Kunzel, Pepping, and Schulte-Bahrenberg

(1992) randomly assigned 120 phobic patients to an experimental group with individualized treatment planned by the therapist, a control group with standardized therapy (exposure in vivo), or a yoked control group. The standardized group had improved more at termination, but these gains diminished over time and no differences were found 2 years later.

Non-experimental studies have shown correlations between formulation-consistent interventions and both process and outcome measures. For example, an important component of psychodynamic formulation is an individual’s interpersonal wishes and expected responses from others to those wishes (Horowitz, 2005; Luborsky, 1977; Messer & Wolitzky, 2007). Crits-Christoph, Cooper, and Luborsky (1988) found that interventions consistent with patients’ main wishes and responses from others as identified with the Core Conflictual Relationship Theme, a formulation method developed by Luborsky (Luborsky & Barrett, 2007), correlated significantly and to a moderately strong degree with outcome in a sample of 43 patients

Correspondence concerning this article should be addressed to Tracy D. Eells, Department of Psychiatry and Behavioral Sciences, University of Louisville, 401 East Chestnut Street, Suite 610, Louisville, KY 40202, USA. E-mail: eells@louisville.edu

undergoing brief psychodynamic therapy. Similarly, adherence to a patient's Plan Formulation, a case formulation method based on Weiss's (1993) Control Mastery Theory, has been shown in multiple small sample studies to predict both outcome and depth of experiencing in therapy (Silberschatz & Curtis, 1993; Silberschatz, Curtis, & Nathans, 1989).

Relatively little research has been conducted on process and content aspects of case formulation. Evidence suggests that the process is difficult and that many clinicians may not formulate well. A study of psychiatry residents across four institutions found deficits in the ability to generate biopsychosocial formulations, although a concerted effort on the part of one institution led to significant improvement (McClain, O'Sullivan, & Clardy, 2004). Kuyken and colleagues (2005) measured the quality of formulations produced by 115 mental health practitioners attending a continuing education event and concluded that only 44% were "at least good enough." Eells and colleagues (2005) found that expert cognitive-behavioral and psychodynamic therapist clinicians produced higher-quality formulations than either experienced or novice counterparts. Quality was defined in multiple dimensions such as the degree of comprehensiveness, elaboration, and complexity of the formulation and evidence that the clinician followed a systematic case formulation process. Taken as a whole, these studies suggest that considerable variability exists in therapists' ability to develop psychotherapy case formulations.

With regard to formulation content, Eells, Kendjelic and Lucas (1998) analyzed randomly selected case formulations from an academic outpatient psychiatry clinic. They concluded that clinicians used the formulation primarily to summarize descriptive information rather than to integrate it into a hypothesis about the causes, precipitants, and maintaining influences of an individual's problems. In another study, Eells and Lombart (2003) found that psychodynamic therapists emphasize coping/defenses, childhood histories, strengths and treatment obstacles in their formulations, and are more likely than cognitive-behavioral therapists to explain problems in terms of early childhood events and stressors. Cognitive-behavioral therapists, in contrast, emphasize symptoms and problems more and explain them in terms of social learning, constitutional, and biological factors.

Expertise in Psychotherapy Case Formulation

The literature on expertise and expert performance provides tools to improve our understanding of the process of psychotherapy case formulation

(Ericsson, Charness, Feltovich, & Hoffman, 2006). Understanding exceptional ability of a specific type may help non-experts reach similar levels of ability. One approach is to study experts in comparison to novices (Chi, 2006). Experts may be identified by measures of academic qualifications (such as doctoral degree vs. graduate students), by years performing a task, or by some measure or index of ability in the area of study. Several studies have demonstrated that experience alone is insufficient for the development of expertise (Feltovich, Prietula, & Ericsson, 2006). They show that acquiring expertise in an area requires specific practice activities, focused reflection with the exploration of alternatives, repetition, and informative feedback, in addition to many years of experience.

Cognitive science has identified reliable characteristics of experts in several specific areas of performance including athletics, music, mathematics, chess, physics, and medicine. One of these characteristics is that experts excel mainly in their own domain rather than possessing superior general skills or abilities. In addition, they perceive large meaningful patterns in their area of expertise, are faster than novices at performing the skills of their domain, quickly solve problems with little error, see and represent a problem in their domain at a more principled level than novices, and spend considerable time analyzing problems qualitatively (Chi, 2006; Glaser & Chi, 1988).

Cognitive scientists have also studied the reasoning processes of experts when solving problems as compared to that of non-experts. Simon and Simon (1978) proposed that experts use forward reasoning, which involves moving from data to hypotheses until one reaches a solution. According to Simon and Simon, forward reasoning leads to efficient and accurate problem solutions when it is combined with a well-integrated and well-indexed representation of the problem in working memory and a vast knowledge base. An example of forward reasoning in the area of case formulation is: "He reports anger at his wife and says that as a child he was very close to his mother, describing her as extremely passive and doting; so, he likely expects all women to be like his mother and becomes anxious or angry when they are not." The clinician begins with descriptive information provided by the patient and draws an inference based upon it.

Backward reasoning has been associated with the problem solving of novices (Simon & Simon, 1978). It is characterized by the generation of problem solutions on the basis of a hypothesis for which supporting data are then sought (Buchanan, Davis, & Feigenbaum, 2006; Patel, Groen, & Arocha, 1990; Simon & Simon, 1978). A case formulation

example is: "She is borderline therefore I expect she was sexually abused as a child." Here, the clinician begins with an inference that the patient has borderline personality disorder, then speculates about a possible past event.

In a series of studies on medical expertise, Patel and colleagues (Patel & Groen, 1986; Patel, Groen, & Frederiksen, 1986) found that expert physicians generating diagnoses or interpreting test results use forward reasoning more than backward reasoning. Kaufman, Yoskowitz, and Patel (2008) concluded that medical experts use more forward reasoning when solving routine problems but that backward reasoning may be more characteristic of non-experts and experts solving non-routine problems. Elstein, Shulman and Sprafka (1978) found that experienced physicians used both forward and backward reasoning.

A common research methodology in the expertise literature is to present problems to experts and novices, then compare their thinking processes in reaching solutions (Ericsson & Simon, 1980, 1984, 1993). Specifically, individuals are presented with problems and instructed to "think aloud" as they solve them. Ericsson and Simon showed that transcripts can be reliably scored and that the process of thinking aloud does not ordinarily affect the cognitive processes being studied. The think aloud paradigm is often augmented with experimenter-provided verbal probes designed to elicit additional information.

The present study examines the process of generating psychotherapy case formulations using concepts and methods from the expertise literature. Using established principles, we sought to identify experts, experienced clinicians and novices in case formulation and used a think aloud methodology to examine how these clinicians developed formulations based on six clinical vignettes. Because the predominant research finding is that experts use more forward reasoning, we predicted that expert psychotherapy case formulators will exhibit more forward reasoning and less backward reasoning. Because experts are assumed to have vastly larger relevant knowledge bases, we also expected the formulations of expert therapists to make greater use of descriptive information and to generate more diagnostic, inferential, and treatment planning information than those of the other two groups.

Method

This study is part of a larger project on expertise in psychotherapy case formulation. The therapists, vignettes, transcription and content coding procedures are the same as described in Eells et al. (2005),

although the present study investigates different research questions and analyzes different data.

Therapists

Sixty-five psychotherapists participated in the project. Twenty-four were novices, defined as clinical psychology graduate students with less than 1500 hours of supervised psychotherapy experience. We chose the 1500 hour criterion to allow for a core of graduate-level education to become established, but to remain well below estimates that at least 10,000 hours of deliberate practice in an area is usually required to achieve expertise (Ericsson, Krampe, & Tesch-Romer, 1993). The novices were recruited from three American Psychological Association (APA) accredited graduate programs in clinical or counseling psychology. On average, they were in their third year of graduate study (range = 1st to 6th year). Eleven novices identified their current and predominant therapy orientation as psychodynamic (PD) and 13 as cognitive-behavioral (CB). Of the 11 PD novices (three males, eight females), nine further identified their orientation as within the interpersonal/relational school of psychodynamic psychology. The remaining two identified themselves further as oriented toward psychoanalysis. Of the 13 CB novices (seven males, six females), 10 further identified themselves as oriented toward Aaron Beck's cognitive-behavioral therapy model and three as eclectic/integrative in orientation. The mean age of the CB novices was 27.1 years and that of the PD novices was 37.3 years.

Experienced therapists were defined as clinical psychologists, counseling psychologists, or psychiatrists with 10 or more years of practice. They were recruited by referral and word of mouth. They resided primarily in the midwestern and southern United States. None of these therapists met additional criteria for expertise (described below). Eight of the 19 experienced therapists identified their current and predominant orientation as CB (seven male, one female) and 11 as PD (nine male, two female).

Of the 11 experienced PD therapists, six further identified with the relational/interpersonal or object relations approach, four with psychoanalysis, and one as eclectic/integrative. Seven of these 11 were psychiatrists and four were clinical psychologists. The mean age of the experienced PD therapists was 57.8 years, with 29.4 years of experience as therapists. They spent an average of 19.4 hours per week doing therapy. They reported a mean of 1.4 publications in the area of case formulation and reported teaching a mean of 2.5 workshops on case formulation.

Of the eight experienced CB therapists, six identified primarily with the Aaron Beck approach, one with rational-emotive therapy, and one with an eclectic/integrative orientation. All eight were doctoral-level psychologists. The mean age of the experienced CB therapists was 49.3 years, with 21.1 years of experience as therapists. They spent an average of 16.3 hours per week doing therapy. They reported a mean of 0 publications in the area of case formulation and reported teaching a mean of .3 workshops on case formulation.

For the expert category, we sought individuals who met our criteria for experienced therapists, but who also had achieved national or international recognition in the area of case formulation. We sought individuals who reasonably may have met Ericsson and colleagues' (1993) criteria for achieving a superior level of performance in a given field, primarily that they had engaged in 10,000 or more hours of deliberate practice in the field. Since we could not measure this directly, we sought indices of expertise. These indices were that (1) they had developed a method of psychotherapy case formulation, (2) they had published extensively in the area of case formulation, or (3) they had presented multiple workshops that included the topic of case formulation. Using these criteria we invited candidates to participate in the study, contacting them at national conferences or by telephone. Twenty-two therapists served as experts (PD = 11, CB = 11).

Of the 11 PD experts (nine male, two female), six were psychiatrists and five were clinical psychologists. Six identified further as relational or object relations in orientation, one as ego analytic, three as psychoanalytic, and one as eclectic/integrative. The mean age of the PD experts was 56.1 years, with 28.4 years as therapists. They reported a mean of 13.3 publications on case formulation and 23.4 case formulation workshops taught.

Of the 11 CB experts (seven male, four female), eight were clinical psychologists, two were psychiatrists and one was a counseling psychologist. Among the CB experts, eight further identified with the approach of Aaron Beck, one with dialectical-behavior therapy, and one with an eclectic/integrative approach. One did not provide additional information. They reported a mean of 8.1 publications on case formulation and 29.5 case formulation workshops taught. The ethnicity of all therapists was Caucasian.

Eells et al. (2005) studied the same therapists and provide further evidence of the difference in case formulation ability between the experts and experienced clinicians. That study found that the experts produced higher-quality formulations than both the experienced therapists and novices. The effect

size on a measure of overall quality comparing the experts and the experienced therapists across the six vignettes was .49, which is considered a "large" effect (Cohen, 1988).

Vignettes

Six vignettes (M words = 405; range = 368–424) were constructed to describe patients who met diagnostic criteria for one of three common psychological conditions: generalized anxiety disorder, major depressive disorder, or borderline personality disorder. One vignette for each disorder was highly characteristic of the condition (high prototypicality) and one was less characteristic but still met diagnostic criteria. Each vignette briefly contained identifying information, presenting condition, history of presenting condition, past history of mental health care, developmental history, social history, and mental status. Despite the brevity of the vignettes, the therapists rated them as moderately adequate on average. The mean adequacy rating was 5.4 ($SD = 1.91$) on a 9-point scale (1 = *completely inadequate*; 9 = *perfectly adequate*). As a manipulation check on whether we succeeded in preparing high and low prototypical vignettes for the three disorders, we asked the therapists how prototypical each vignette was of the disorder it was intended to depict (1 = *minimally prototypical*, 9 = *extremely prototypical*). The therapists rated the generalized anxiety vignette that was intended to be highly prototypical as more prototypical than the generalized anxiety vignette intended to be less prototypical ($M = 7.56$ vs. 4.63), $t(56) = 8.93$, $p < .001$. Similar findings were observed for the major depressive disorder vignettes ($M = 8.04$ for the intended highly prototypical major depressive disorder vignette vs. 4.42 for its low prototypical counterpart), $t(56) = 13.14$, $p < .001$, as well as for the borderline personality disorder vignettes ($M = 8.37$ vs. 4.09), $t(56) = 15.31$, $p < .001$.

Procedures

Case formulations of the therapists were obtained either in a face-to-face meeting or over the telephone with the first or fourth author. Audio-recordings of the vignettes in the voice of the first author were presented in a fixed random order with the constraint that vignettes describing the same disorder or presenting a disorder at the same level of prototypicality were not given consecutively. Therapists were given written copies of the vignettes and permitted to take notes while listening to the audiorecording. In the case of a telephone interview, written material including the vignettes was mailed to the therapist in advance and was sealed in an envelope with

instructions not to open the envelope until instructed to do so. After listening to each vignette, the therapist was given five minutes to “think aloud about your conceptualization of the patient ... [to] construct a case formulation ... as best you can, addressing whatever you feel is important.” After five minutes the therapist was interrupted by the interviewer and given two minutes to “think aloud about how you would treat the patient in psychotherapy.”

The 390 formulations (65 therapists \times 6 vignettes) were transcribed using standards recommended by Mergenthaler and Stinson (1992) and then segmented into “idea units” (IUs) following procedures described by Stinson, Milbrath, Reidbord, and Bucci (1994). The method involves three judges reading text and segmenting it into relatively small portions, usually a sentence or less, according to the judge’s decision of what represents a complete idea. Four graduate students and the first author served as text segmenters, working in teams of three. All were trained on practice vignettes until they achieved consistent percent-of-agreement levels of .85. For the study transcripts, the mean percent agreement among the judges was 87%, with a range of 86% to 88% among the vignettes. These results indicate good agreement on identifying idea units. Disagreements were resolved through consensus.

The formulations were then coded using the Case Formulation Content Coding Method (CFCCM) (Eells et al., 1998), a cross-theoretical and reliable tool for categorizing the content of psychotherapy case formulations. The reliability of the CFCCM has been assessed in two studies (Eells et al., 1998; Kendjelic & Eells, 2007). In both, a mean kappa of .86 (range .60 to 1.0 across both studies) was obtained when pairs of coders assessed the presence or absence of a case formulation element in a written case formulation. The CFCCM was revised and expanded for the current project so that it could be applied to each idea unit in a case formulation. As shown in Table 1, the revised system is organized into four major categories of information that might be contained in a case formulation: description, diagnosis, inferential information, and treatment planning. Each of these contains the subcategories shown in Table 1.

The goal of the content coding step was to produce a set of reliable, consensus codes for each idea unit in each formulation. The primary task was to determine which codes, if any, apply to an idea unit. Coding rules permitted each idea unit to receive a maximum of one code from each of the major coding categories (i.e., descriptive, diagnostic, inferential, and treatment planning information). However, only one subcategory within each major category could be applied to an idea unit. For

example, the statement “*Since this 35-year-old single patient describes his parents as critical, he may have a core belief that others will be critical*” would be coded for descriptive (shown in italics) and inferential information (shown in boldface), but not diagnosis or treatment. Coders would then need to decide which subcategory to apply within each major category. The statement “he may have a core belief that others will be critical” would likely be coded as an “inferred psychological mechanism” within the broader category of inferential information. Two subcategories of description apply: identifying information (“*35-year-old single patient*”) and developmental history (“*describes his parents as critical*”), although coders were constrained to choosing only one. They followed a set of rules prioritizing coding elements within each major category. In the above example, developmental history would take precedence over identifying information.

Six clinical or counseling psychology graduate students, working in teams of three, content coded the transcripts. They were unaware of the treatment orientation and expertise status of the therapist. The coders were trained by reading the coding manual, attending a series of training sessions led by the first author, and practicing on several formulations developed for training purposes. Each team held consensus meetings periodically during which each code was reviewed and disagreements were resolved through discussion. We calculated multi-rater kappa coefficients using the formula discussed in Siegel and Castellan (1988), obtaining values of .61, .81, .62, and .69, respectively, for codes within the descriptive, diagnosis, inferential, and treatment categories, reflecting good to excellent agreement beyond chance (Fleiss, 1981). Next, we recoded the sample of codes into dummy codes so they denoted only the presence or absence of a major category code, thus excluding consideration of whether agreement was found on a subcategory coding element within a major category. That is, we examined whether, for example, coders agreed as to whether descriptive information of any type was present, regardless of how one might label that descriptive information (e.g., identifying information or symptom identification). Resulting kappas were .64, .82, .70, and .81 for the descriptive, diagnosis, inferential, and treatment categories, respectively. A comparison of the two sets of kappa indicates more agreement about the presence or absence of a major coding category than about the specific code within the major category that should be applied. For example, they agreed more that a piece of inferential information was offered than about whether it indicated, say, a psychological mechanism, a

Table 1. Description of Case Formulation Content Coding Method

Major category	Subcategory
Descriptive information	(1) Identifying information; (2) symptom identification (information <i>given</i> in vignette); (3) history of present or previous episode of mental health problems or care (in self or family); (4) medical history; (5) developmental history; (6) adult life history; (7) mental status information; (8) other descriptive information; (9) more specific descriptive information needed.
Diagnosis	(1) <i>DSM-III-R</i> or <i>DSM-IV</i> Axis I diagnosis; (2) <i>DSM-III-R</i> or <i>DSM-IV</i> Axis II diagnosis; (3) Axis I and Axis II diagnosis in same idea unit; (4) alcohol/substance abuse or dependency.
Inferential information	(1) Inferred problems in global psychological, social, or occupational functioning; (2) inferred symptoms or problems; (3) predisposing experiences, events, traumas, stressors inferred as explanatory; (4) precipitating or current stressors and/or events; (5) inferred mechanisms: psychological (including problematic aspects/traits of the self; problematic aspects of relatedness to others; dysfunctional thoughts and/or core beliefs; affect regulation or dysregulation; defense mechanisms/problematic coping style; skills or social learning deficit); (6) inferred biological mechanisms; (7) inferred social or cultural mechanisms (including absence of or poor psychosocial support; demographic/cultural factors as source of a problem; role conflict: role strain; role transition; role dispute); (8) strengths in global psychological, social, or occupational functioning (including strengths/adaptive skills; aspects or traits of self; adaptive perceptions of or beliefs about others; positive motivation for treatment; adaptive wishes; hopes or goals; good psychosocial support); (9) identification of potential therapy-interfering events.
Treatment planning	(1) Type of treatment; (2) evaluation/assessment; (3) specific techniques; (4) possible red flag issues; (5) treatment contract/expectations; (6) therapist–patient relationship; (7) signs and symptoms; (8) predisposing experiences, events, or traumas; (9) psychological mechanisms; (10) social and/or cultural factors; (11) biological factors/psychopharmacology; (12) strengths in global psychological, social, or occupational functioning.

predisposing experience, or a problem in global functioning.

Because CFCCM coding procedures permitted both descriptive and inferential information to be coded within an IU, our plan to conduct sequential analysis required a determination as to whether the inferential idea or the piece of description occurred first in those IUs containing both. To make this determination, the first and second authors identified IUs that contained both descriptive and inferential information. Then, they made a judgment about which occurred first within the IU. Interrater agreement was 84.9%.

Results

A total of 14,499 codes were applied to the idea units, a mean of 44.9 ($SD = 15.3$) codes for each of the 390 formulations. Our analysis began by comparing the expert, experienced and novice case formulators with respect to the frequency of major CFCCM category codes, and then the subcategory codes. We conducted a series of analyses of variance examining possible effects for expertise, therapy mode and expertise \times therapy mode interactions, following procedures described by Appelbaum and Cramer (1974). We then conducted comparisons to assess whether the experts performed as predicted. Since our purpose for including multiple vignettes in the design was to ensure that we tested case formulation skill across a variety of diagnostic conditions, we collapsed all analyses across the six

vignettes. After investigating the CFCCM frequencies, we examined reasoning processes.

CFCCM Major Categories

Table 2 shows the mean frequency count for each major category of formulation information, by therapist group, averaged across the six vignettes. The major category means are based on the sum of the CFCCM subcategory codes for the respective major category. For example, the descriptive information mean frequency calculated for the CB novices ($M = 11.8$) is the total number of descriptive information subcategory codes generated by this group of therapists in all six vignettes, divided by six.

The best fitting linear models for descriptive information were an expertise main effect, $F(2,384) = 16.79$, $p < .001$, and a therapy mode main effect, $F(1,384) = 4.84$, $p < .05$. The comparisons showed that the experts generated more descriptive information than the experienced therapists, $t(244) = 5.71$, $p < .001$, and the novices, $t(274) = 2.97$, $p < .01$, as predicted. In addition, the formulations of the PD therapists contained more descriptive information than those of their CB counterparts.

For diagnostic information, the best fitting linear model was an expertise \times therapy mode interaction, $F(2,384) = 8.11$, $p < .001$, although we also observed main effects for expertise, $F(2,384) = 9.42$, $p < .001$, and therapy mode, $F(1,384) = 15.98$, $p < .001$.

Table 2. Case Formulation Content Coding Method (CFCCM) Items Coded On Average Greater Than Once Across the Six Vignettes and Statistically Significant

CFCCM major category:		Novices		Experienced		Expert		Statistical test results		
CFCCM subcategory:		CB	PD	CB	PD	CB	PD	Main effects and interactions	Contrasts	Effect size (<i>r</i>)
<i>Descriptive information^a:</i>	<i>M</i>	11.8	14.5	10.0	11.0	15.3	16.2	Expertise***	EXT > EXP***	.33
	<i>SD</i>	5.8	7.6	5.5	4.6	9.2	8.9		EXT > NOV**	.17
Symptom identification	<i>M</i>	1.37	1.03	.54	.68	1.62	.95	Therapy mode*	PD > CB	-.09
	<i>SD</i>	1.50	1.21	.80	1.01	2.29	1.63	Expertise	EXT > EXP	.20
Developmental history	<i>M</i>	1.81	2.41	1.60	2.03	1.91	2.62	Therapy mode	PD > CB	-.15
	<i>SD</i>	1.91	1.82	1.58	1.86	1.59	2.20			
Adult relationship history	<i>M</i>	3.36	3.73	2.00	2.70	3.79	3.48	Expertise	EXT > EXP	.25
	<i>SD</i>	2.29	2.57	1.82	2.15	3.05	2.23			
Need more information about...	<i>M</i>	3.24	4.67	3.38	2.89	5.21	6.42	Expertise	EXT > EXP	.27
	<i>SD</i>	4.08	4.90	3.68	2.71	5.15	6.48			
<i>Diagnostic information</i>	<i>M</i>	2.7	2.1	2.1	2.2	4.9	2.3	Expertise***	EXT > EXP***	.24
	<i>SD</i>	2.8	2.1	2.1	2.4	3.7	2.5		EXT > NOV**	.19
								Therapy mode***	CB > PD	.20
								Expertise × Therapy mode***	CB-EXT > mean of others***	.34
<i>Inferential information^b:</i>	<i>M</i>	16.4	16.3	12.9	15.0	19.4	18.1	Expertise***	EXT > EXP***	.35
	<i>SD</i>	5.9	5.9	4.2	5.5	8.1	5.5		EXT > NOV***	.19
Inferred symptoms/problems	<i>M</i>	1.83	1.85	1.46	1.47	3.02	1.79	Expertise	EXT > EXP	.25
	<i>SD</i>	1.46	1.62	1.41	1.44	2.32	1.54			
Predisposing experiences	<i>M</i>	4.56	4.86	3.96	4.98	3.95	5.98	Therapy mode	PD > CB	-.17
	<i>SD</i>	2.63	3.19	2.64	3.33	3.19	4.01			
Psychological mechanisms	<i>M</i>	6.63	7.08	5.23	4.77	8.21	6.59	Expertise	EXT > EXP	.29
	<i>SD</i>	4.26	4.56	2.88	3.26	5.46	3.54			
<i>Treatment planning^c:</i>	<i>M</i>	10.2	11.1	12.7	10.0	18.0	12.6	Expertise***	EXT > EXP***	.32
	<i>SD</i>	3.8	4.8	6.4	4.1	6.4	6.2		EXT > NOV***	.38
								Therapy mode***	CB > PD	.19
								Expertise × Therapy mode***	CB-EXT > mean of others***	.43
Evaluation/assessment	<i>M</i>	.32	.68	.31	.27	.44	.67	Expertise	EXT > EXP	.32
	<i>SD</i>	.80	2.30	.75	.57	1.70	.95		EXT > NOV CB-	.18
								Expertise × Therapy mode	EXT > mean of others	.27
								Expertise	EXT > NOV	.34
Focus on treatment contract/expectations	<i>M</i>	1.31	1.79	2.69	2.42	3.56	2.65	Expertise		
	<i>SD</i>	1.54	1.83	2.67	2.28	2.63	2.62			
Focus on signs/symptoms	<i>M</i>	.73	.68	.65	.41	1.76	.68	Expertise		
	<i>SD</i>	1.09	1.11	.96	1.04	2.94	1.15			
Focus on psychological mechanisms	<i>M</i>	3.14	2.35	2.42	1.53	3.24	1.80	Therapy mode	CB > PD	.25
	<i>SD</i>	2.39	2.12	2.40	1.60	2.50	1.79			
Focus on social/cultural factors	<i>M</i>	.42	.11	.19	.05	.45	.21	Therapy mode	CB > PD	.17
	<i>SD</i>	.95	.43	.45	.21	1.11	.60			

Note. EXT = expert, EXP = experienced, NOV = novice, CB = cognitive-behavioral, PD = psychodynamic, CB-EXT = cognitive-behavioral experts; * $p < .05$; ** $p < .01$; *** $p < .001$; ^aall descriptive information subcategory p values $< .008$; ^ball inferential information subcategory p values $< .006$; ^call treatment planning information subcategory p values $< .004$.

Analyses showed that the experts generated more information than both the experienced therapists, $t(244) = 5.46$, $p < .001$, and the novices, $t(274) = 2.81$, $p < .01$, as predicted. The interaction was driven by the greater attention given to diagnosis by the CB experts, as confirmed by comparing their results with the mean of all other therapists, $t(388) = 7.04$, $p < .001$.

For inferential information, the best fitting model was an expertise main effect, $F(2,384) = 18.10$,

$p < .001$. As predicted, the experts generated more inferences in their formulations than the experienced therapists, $t(244) = 6.17$, $p < .001$, and the novice therapists, $t(274) = 3.31$, $p < .001$.

For treatment planning, the best fitting linear model was an expertise × therapy mode interaction, $F(2,384) = 12.11$, $p < .001$. In addition, we observed main effects for expertise, $F(2,384) = 31.17$, $p < .001$, and therapy mode, $F(1,384) = 17.78$, $p < .001$. The experts produced more treatment

planning ideas than the experienced therapists, $t(244) = 5.29$, $p < .001$, and the novices, $t(274) = 6.87$, $p < .001$. The interaction is explained by the expert CB therapists generating significantly more treatment ideas than the other therapists, as shown by comparing their mean to that of all other therapists, $t(388) = 9.37$, $p < .001$.

CFCCM Subcategories

We continued our analysis by examining differences among the therapists in their use of specific types of descriptive, inferential, and treatment information in their formulations. To minimize the chance of experimentwise error, we first eliminated CFCCM items that were coded infrequently. We defined an infrequent item as one that appeared less than once, on average, across the six vignettes, i.e., those with a mean frequency less than .167 (1/6). Next, we conducted three multivariate analyses of variance (MANOVA). The first MANOVA treated the CFCCM descriptive information subcategory codes that met our frequency criterion as dependent variables. The second treated the CFCCM inferred information subcategory codes as dependent variables, and the third treated the CFCCM treatment subcategory codes as dependent variables. We did not pursue further analysis of diagnostic information. If the MANOVA analyses were statistically significant using the Wilks lambda criterion, we continued with separate ANOVAs of the dependent variables included in the MANOVA, applying a Bonferroni correction based on the number of dependent variables. If the ANOVAs were statistically significant, we conducted contrasts of the experts' results to those of the experienced and novice therapists, applying the same Bonferroni correction that was applied to the omnibus test. These procedures resulted in a conservative estimate of differences among the therapists on the CFCCM items analyzed.

The MANOVA of the descriptive information items showed main effects for expertise, $F(12, 758) = 3.94$, $p < .001$, and therapy mode, $F(6, 379) = 3.74$, $p < .001$. Table 2 shows the CFCCM descriptive information items that met our frequency criteria and for which we observed statistically significant differences among the groups. As indicated, ANOVA results showed expertise main effects for symptom identification, $F(2, 348) = 6.33$, $p < .002$, adult relationship history, $F(2, 348) = 10.05$, $p < .001$, and "need more information", $F(2, 348) = 11.55$, $p < .001$. The latter code was applied when the therapist expressed a wish for more descriptive information than was provided in the vignette. A series of contrasts showed that experts identified

more symptoms than the experienced therapists, $t(244) = 3.25$, $p < .001$; made greater use of adult life history than the experienced therapists, $t(244) = 4.03$, $p < .001$; and felt more in need of additional descriptive information than the experienced therapists, $t(244) = 4.43$, $p < .001$. PD therapists made more use of developmental history, $F(6, 379) = 3.74$, $p < .001$.

The MANOVA of inferred information items showed an expertise \times therapy mode interaction, $F(16, 754) = 2.51$, $p < .001$, and main effects for expertise, $F(16, 754) = 2.90$, $p < .001$ and therapy mode, $F(8, 377) = 4.02$, $p < .001$. The related ANOVAs showed no interactions that exceeded the Bonferroni threshold, but expertise main effects for inferred symptoms and problems, $F(2, 384) = 9.37$, $p < .001$, and for psychological mechanisms, $F(2, 384) = 8.44$, $p < .001$, that did. As indicated in Table 2, contrasts showed that the experts inferred more symptoms and problems than the experienced therapists, $t(244) = 4.09$, $p < .001$, as well as more psychological mechanisms than the experienced therapists, $t(244) = 4.10$, $p < .001$. The PD therapists inferred more past experiences as explanatory than the CB therapists, $F(1, 384) = 11.35$, $p < .001$.

The MANOVA results for the treatment planning items showed an expertise \times therapy mode interaction, $F(20, 750) = 3.64$, $p < .001$, and main effects for expertise, $F(20, 750) = 5.73$, $p < .001$, and therapy mode, $F(10, 375) = 7.15$, $p < .001$. As Table 2 shows, effects were observed for five CFCCM treatment planning items that met the frequency criterion defined above. For the item "evaluation/assessment," which was coded when a therapist recommended additional evaluation or assessment of any kind, the best fitting linear model was an expertise by therapy mode interaction, $F(2, 384) = 6.29$, $p < .002$, although we also observed an expertise main effect, $F(2, 384) = 10.96$, $p < .001$, for this item. The experts addressed the need for further evaluation or assessment more than the experienced therapists, $t(244) = 5.28$, $p < .001$. The interaction is due to the greater focus on further evaluation by the CB experts as compared to all other therapist groups, $t(388) = 5.45$, $p < .001$. For the treatment planning item "focus on treatment contract/expectations," the best fitting linear model was an expertise main effect, $F(2, 384) = 17.49$, $p < .001$. Contrasts showed that the experts focused on a treatment contract and expectations more frequently than did the novices, $t(274) = 5.95$, $p < .001$. For the item "focus on signs and symptoms," the best fitting model was an expertise main effect, $F(2, 384) = 6.63$, $p < .001$. None of the contrasts, however, was statistically significant below the Bonferroni threshold of $p < .004$. As shown in Table 2, the

CB therapists focused their treatment plans more on psychological mechanisms, $F(1, 384) = 22.49$, $p < .001$, and social/cultural factors, $F(1, 384) = 10.96$, $p < .001$, as compared to the PD therapists.

Effect Sizes

To understand the magnitude of the observed differences, we calculated effect sizes based on the Pearson product-moment correlation coefficient, correlating each dependent variable with index values (i.e., dummy codes) of the independent variables, as recommended by Cohen (1988). The results are shown in the last column of Table 2. Cohen suggested that an r of .10 be considered small, .30 medium, and .50 large. Based on these standards, the effect sizes range from small to medium. The strongest effect sizes are for the major CFCCM categories and relate to the experts' greater use of descriptive information, greater focus on diagnosis, greater number of inferences in their formulations, and more extensive treatment planning.

Reasoning Processes

Consistent with a conception of forward reasoning as an inferential process of working from data to hypotheses, we operationalized a unit of forward reasoning as a sequential description-to-inference link in the text of a therapist's formulation. Conversely, we operationalized a unit of backward reasoning as an inference-to-description link. Links could occur either between two idea units or within a single idea unit. To analyze forward and backward reasoning, we first combined all CFCCM subcategory codes into their respective major category. For example, all the subcategories of descriptive information (see Table 1) were recoded as "description"

and all the subcategories of inferential information were recoded as "inference."

We measured forward and backward reasoning in two ways. First, we assessed the strength of the association between contiguous pieces of descriptive and inferential information. In the case of forward reasoning, we asked the question: Given a piece of descriptive information, what is the probability that it will be followed by an inference? Conversely, in the case of backward reasoning, we asked: Given a piece of inferential information, what is the probability that it will be followed by a piece of descriptive information? To answer these questions, we conducted a lag one sequential analysis. As shown in Table 3, we chose the odds ratios as a measure of the strength of the relationship between descriptive and inferential information because it is readily interpretable. The odds ratio of 2.22 for the description-to-inference link indicates that for the therapists as a group a descriptive piece of information was 2.22 times as likely to be followed by inferential information as by additional description, by a diagnosis, or by treatment planning information, suggesting the presence of forward reasoning. Similarly for the inference-to-description link, an inference was twice as likely (odds ratio = 2.00) to be followed by descriptive information as by an additional inference, a diagnosis, or treatment planning information, suggesting the presence of backward reasoning. We observed odds ratios of similar magnitude for each of the therapist groups. These results show that both forward reasoning and backward reasoning were used by the therapists in their formulations without regard for treatment orientation or case formulation expertise.

To examine group differences in the strength of association between the components of forward and backward reasoning (i.e., description and inference), we calculated a Yule's Q for each therapist. Yule's Q is a transformation of the odds ratio that is more

Table 3. Lag One Sequential Analyses: Description-to-Inference and Inference-to-Description Links

Type of link		All therapists	Novice		Experienced		Expert	
			CB	PD	CB	PD	CB	PD
Forward reasoning								
Description →	<i>f</i>	6.59	6.69	6.71	5.00	5.98	7.11	7.58
inference	<i>SD</i>	2.80	3.24	2.81	2.67	2.78	4.22	3.63
	<i>OR</i>	2.22	2.73	1.88	2.60	2.51	2.11	1.92
Backward reasoning								
Inference →	<i>f</i>	6.36	6.45	6.70	4.63	5.88	6.65	7.39
description	<i>SD</i>	3.34	3.23	2.95	2.50	2.94	3.90	3.69
	<i>OR</i>	2.00	2.37	1.79	2.04	2.45	1.79	1.79

Note. CB = cognitive-behavioral, PD = psychodynamic, *f* = mean frequency of transitions per vignette, *SD* = standard deviation, *OR* = odds ratio pooled across vignettes.

useful for inference (Bakeman, McArthur, & Quera, 1996; Wickens, 1993). We pooled across each therapist's six formulations in order to maximize the chances of sufficient marginal sums in the resulting 2×2 tables (e.g., description or not description followed by inference or not inference). Excluding subjects whose marginal sums did not exceed five, we then conducted a series of analyses of variance using Yule's Q as the index of association and as the dependent variable (Bakeman et al., 1996; Wickens, 1993). None of these tests was statistically significant, indicating that the strength of association within the description-to-inference links and of the inference-to-description links did not differ between the groups.

The second way that we measured forward and backward reasoning was through frequency counts of description-to-inference links and inference-to-description links. We conducted two two-way analyses of variance to determine whether the experts generated more description-to-inference links than the other therapists and fewer inference-to-description links, as hypothesized. As shown in Table 3 and looking first at forward reasoning, we found a significant main effect for expertise, $F(2,384) = 9.34$, $p < .001$, with the experts producing more forward reasoning links than their experienced counterparts, $t(244) = 4.36$, $p < .001$, with an effect size of .25 as measured by r . No difference was found between the experts and the novices. For backward reasoning, we also observed a main effect for expertise, $F(2,384) = 9.08$, $p < .001$, although, contrary to our hypothesis, the experts used more backward reasoning than the experienced therapists, $t(244) = 4.22$, $p < .001$, with an effect size of .24, as measured by r . The experts did not differ from the novices in the frequency of backward reasoning links generated. We also observed a main effect for therapy mode, $F(1,384) = 4.52$, $p < .05$, with the PD therapists generating more backward reasoning links than the CB therapists ($r = -.09$).

Finally, to determine whether the experts generated more forward reasoning than backward reasoning, we next conducted a paired-group t test comparing the number of description-to-inference links with the inference-to-description links among the experts. (See Table 3.) As predicted, they produced more forward reasoning links than backward reasoning links $t(131) = 2.13$, $p < .05$. In contrast, no differences were observed among the novices, $t(143) = 1.07$, $p < .29$. Like the experts, the experienced therapists also produced more forward reasoning links than backward reasoning links, $t(244) = 2.05$, $p < .05$.

Discussion

Both hypotheses received support. Regarding the first, we found that the experts generated more forward reasoning than the other therapists when forward reasoning is measured by the frequency of description-to-inference links in the case formulations. We also found that the experts generated more forward reasoning links than backward reasoning links. Contrary to our hypothesis, the experts generated more backward reasoning, as measured by the number of inference-to-description links, than the other groups. The experts did not generate stronger description-to-inference associations than the other therapists, nor did they generate stronger inference-to-description associations.

Our second hypothesis was strongly supported: Expert case formulators generated more descriptive, diagnostic, inferential, and treatment planning information than the non-experts. In particular, they focused more on symptom identification and on the history of adult relationships. They were also more likely to express a need for additional descriptive information to develop their formulations. Experts inferred more symptoms, problems and psychological mechanisms. They were more likely to recommend further evaluation, to focus on the treatment contract and treatment expectations, and to focus treatment on symptoms.

Our findings are consistent with the literature on expertise, which shows that experts use more appropriate problem-solving strategies than novices (Chi, 2006), and, in general, more forward than backward reasoning (Hunt, 1989). To be effective, forward reasoning depends on a well-developed and rich knowledge base. Our finding that the experts generated a greater number of idea units in the case formulations is consistent with this idea.

Some studies have found that experts use more forward and more backward reasoning than non-experts (e.g., Elstein et al., 1978). Elstein, Shulman and Sprafka (1978) found that physicians who are presented with a list of symptoms and asked to think aloud about diagnostic possibilities tend to consider diagnosis very early in their thinking and to alternate between symptom and diagnostic considerations as they seek the correct diagnosis. Our finding that experts use more forward and backward reasoning than non-experts is consistent with Elstein, Shulman and Sprafka (1978), and suggests our case formulators may have followed a similar process. We explored this possibility more closely by comparing the formulations of the CB and the PD therapist generating formulations rated highest in quality with those of a CB and a PD therapist receiving quality ratings at the 25th percentile (Eells,

2010). We found that the formulations of the CB and PD expert followed a pattern similar to that described by Elstein, Shulman and Sprafka (1978).

Experts also have more accurate self-monitoring skills in terms of their ability to detect error and the status of their own comprehension (Chi, 2006). This is consistent with our finding that expert case formulators were more likely to seek additional descriptive information. Nevertheless, they made better use of the information that was available, as indicated by their greater number of diagnostic, inferential and treatment planning ideas.

The expert–novice paradigm is one of several approaches one might use to study expert performance. For example, Alexander (2004); and Alexander, Jetton, & Kulikowich, 1995) developed a model of domain learning that adds a developmental perspective and focuses more on personality, social, and motivational factors that are tied to the development of expertise. Kahneman (2003) posited two modes of cognitive functioning that may serve as a framework for understanding expert performance in psychotherapy. One is an intuitive mode that is rapid, relatively effortless, free flowing, emotional, and automatic; the second is a reasoning mode that is slow, serial, deliberate, effortful, rule-governed, and controlled. According to Kahneman, the best decision makers benefit from the easily accessible and free flow of ideas offered through the intuitive system, but balance this with a more deliberate, effortful reasoning system. Another alternative is that of Gigerenzer and colleagues (Gigerenzer et al., 2005; Gigerenzer, Todd, & ABC Research Group, 1999) who assert that a small set of simple but powerful “fast and frugal” heuristics facilitate accurate and efficient decision making, and may characterize the thinking of experienced physicians. Expert case formulation development may well follow processes similar to those proposed by Alexander, Kahneman, and Gigerenzer. These additional models provide further opportunities to understand how a high level of performance in psychotherapy case formulation can be achieved and maintained.

Although many more differences were observed based on level of expertise than on therapy orientation and we did not make specific hypotheses about therapy mode differences, a number of therapy mode findings emerged, many of which are consistent with the theoretical principles underlying each orientation. In particular, PD therapists focused more on the developmental history provided in the vignettes and made more inferences focused on predisposing experiences. This is consistent with the psychodynamic theoretical focus on early life events as shaping adult psychopathology. The CB

therapists focused more on diagnosis and said more about treatment planning, particularly the CB experts. These experts also tended to infer more symptoms and problems. In the area of treatment planning, CB therapists focused more on psychological mechanisms and on social or cultural factors.

Limitations

Several limitations should be kept in mind when considering these results. First, our method of identifying relatively small idea units and subjecting them to sequential analysis is readily applied to multiple long sequences of ideas and thus lends itself well to inference testing. However, a therapist’s meaning is not directly assessed through this method. Although the description and inference links were contiguous, it is possible that some were not conceptually related. For three reasons, we are convinced that this is not a large problem in the current dataset. First, our reasoning process conclusions are drawn from a large sample of links: 2570 forward reasoning links and 2480 backward reasoning links. Any “error” introduced by unrelated ideas should be randomly distributed across groups and thus is unlikely to significantly influence the statistical tests. Second, a sizable number of these links occurred within the same idea unit and were sequenced by coders, ensuring a conceptual relationship. Third, we spot checked several randomly identified forward and backward links and determined that they were conceptually meaningful. Nevertheless, reasoning patterns spanning multiple idea units or idea units that are hierarchically organized would not be detected by the sequential analysis method we used.

A second limitation concerns the sample. It is relatively small, which limits generalizability of the results. Further, the therapists differed in ways other than theoretical orientation and degree of experience/expertise, introducing possible confounds. We combined the therapists into two groups, PD and CB, when each of these orientations has diverse theoretical variations within it. Fortunately, we were able to explore the possibility of greater than desired within-group theoretical heterogeneity because we asked the therapists to identify a more narrow orientation within the broader one. Almost without exception, the CB therapists identified with the Beck tradition, whereas the PD therapists identified themselves within an interpersonal–relational tradition. Another consideration regarding the sample is our strategy for identifying experts. By design, the experts resembled the experienced therapists in several ways, such as years of experience providing

psychotherapy, number of hours per week doing therapy, educational background, age and gender. Their identification as “experts” hinged on presumed cognitive skills that are not directly observable. Our proxies for identifying the presence of these cognitive skills may have been flawed. It is possible that developing a method of case formulation or giving multiple workshops and publishing extensively on the topic are not good indicators of expertise in performing the case formulation task. Other methods for identifying experts could have been used, such as peer-nomination or therapy outcome scores, although these have limitations as well. For example, peers often do not see colleagues formulate cases and psychotherapy outcome involves many factors besides case formulation skill. We are reassured that our experts actually possess greater case formulation skills by the results of Eells et al. (2005), which found that the same sample of experts generated formulations that were reliably rated as higher in quality than those of the other therapists.

A third limitation is that the study is cross-sectional in nature, which leaves open the possibility that unknown differences in the cohorts of therapists account for the differences observed, not expertise. Another variable, such as motivational differences, may account for the observed differences. A better design would be a well-controlled longitudinal study in which the development of expert case formulation skill could be investigated over time.

Fourth, these results relate only to case formulation skills, not to the association between case formulations and psychotherapy process or outcome variables. It is possible that the expert case formulators are not experts at producing superior psychotherapy outcomes.

Fifth, some may object to the use of vignettes, claiming that they are too distant from the actual clinical situation. On the other hand, the use of vignettes provided several benefits. We were able to better control the information provided to the therapist, for example, by ensuring that similar categories of information were contained in each. Vignettes also permitted us to systematically vary disorder and prototypicality, which made possible the presentation of three different disorders at two distinct levels of prototypicality. Including multiple vignettes allowed us to assess case formulation skill generally rather than for a specific disorder or a small subset of disorders.

A sixth limitation may be posed by the experimental procedures. The therapists were under considerable time pressure and needed to think quickly. This time pressure may have enabled the experts to differentiate themselves more from the non-experts.

However, it might also limit generalizability to a non-experimental setting where these time constraints are less stringent.

Implications

Notwithstanding the above limitations, the results of this study have potential implications for psychotherapy training and research. Future training efforts might explicitly recognize forward and backward reasoning, along with a criterion-based approach in which model formulations are offered as standards (Caspar et al., 2004). Case formulation training might intentionally include facts-to-inference steps, while also allowing for the potential that a trainee may have a sudden insight that could then be supported by descriptive information. Trainees might engage in exercises specifying descriptive information used to generate alternate inferences and treatment ideas, thereby better learning to consider a wide range of potential explanations and treatment interventions to potentially apply to a case. Training should obviously continue to focus on acquiring a broad knowledge base, and skills related to the application of this knowledge to specific, real-time situations may be developed better in an applied rather than a didactic learning environment. Through deliberate practice, explicit attention might also be given to helping therapists develop the knowledge structures and representations that experts may possess. In order to provide feedback on case formulation, it is also possible that a version of the CFCCM could be developed to make it a less labor intensive and thus more practical and adaptable tool for measuring case formulation quality. Its cross-theoretical structure could facilitate training across multiple orientations to therapy.

These findings also have implications for psychotherapy research. One pertains to the role of case formulation in psychotherapy outcome. As reviewed earlier, the results of randomized clinical trials comparing manual-driven with formulation-driven therapy are inconclusive about whether case formulation plays the essential role in psychotherapy that has been claimed for it. These studies have multiple methodological problems that limit conclusions one can draw from them. One problem is a lack of heterogeneity between the levels of the independent variable since manual-driven therapies often allow for individualization and flexibility of the manual to the specific client (Wilson, 1996) and some manuals explicitly include a case formulation step (e.g., Clark, 1997; Ryle, 1990).

An alternative methodological approach would be to investigate case formulation expertise as a

mediating or moderating variable of psychotherapy outcome (Baron & Kenny, 1986; Frazier, Tix, & Barron, 2004). Viewing case formulation expertise as a mediator would allow one to explore its causal contribution to psychotherapy outcome. Specifically, one could hypothesize that case formulation expertise is a partial mediator of the causal relationship between general therapeutic competence and therapy outcome. The mediational hypothesis is that therapist competence will predict outcome, but that therapists with case formulation expertise will produce better outcomes. A related prediction is that a weaker relationship will exist between therapist competence and outcome once the relationship between case formulation expertise and outcome is controlled for. Easden and Fletcher (2010) have presented preliminary evidence supporting similar hypotheses. Viewing case formulation as a moderator, one could explore whether an expertly derived case formulation improves outcome with more difficult patients, as has been hypothesized (Persons, 2008), but is less important for more routine cases. The CFCCM in its current or modified form could serve as a measure of case formulation expertise.

Another implication for psychotherapy research pertains to the role of case-formulation guided therapy in an environment of evidence-based practice. Multiple professional organizations are calling for evidence-based practice (e.g., American Psychological Association, 2005). Case-formulation guided treatment could play an important role in such a practice environment since it is well-suited to the empirical demonstration of effectiveness (Eells & Lombart, in press). A sufficiently powered effectiveness study comparing case-formulation guided therapy, empirically supported treatment, and treatment as usual would provide a good test of the role of case-formulation guided therapy. It would be important to include a measure of case formulation expertise in the study. If case formulation skill plays the essential role in therapy that is claimed, one would predict superior outcomes for the case-formulation guided treatment condition. Equivalence between the case-formulation guided condition and the empirically supported treatment condition would also provide an empirical basis for a choice between these approaches to therapy. Persons and colleagues (2006) have conducted a small study demonstrating equivalence between these latter two conditions.

Conclusions

The present study, coupled with Eells et al. (2005), demonstrated that differences exist between experts and non-experts in case formulation skill and that

those who generate high-quality case formulations do so in predictable ways that are different from non-experts. Further, these skills were demonstrated across multiple diagnostic conditions and across both cognitive-behavior and psychodynamic approaches to psychotherapy.

Acknowledgements

This study was supported in part by grants from the National Institute of Mental Health (R03MH53996-01) and the Health Resources and Services Administration, US Department of Health and Human Services (6D14HP00135). The authors thank the therapists who served as subjects in this study, and the following individuals who provided assistance: Andrew Brothers, Camille Frey, Katherine Hoover, Greg Lynch, Maryrose Manshadi, Lisa Milliner, Richard Morris, Troy Raffield, Diane Taylor, Bernadette Walter, Julia Ward, and Steve Wisniewski. Copies of the CFCCM are available upon request. The vignettes are available for research purposes.

References

- Alexander, P.A. (2004). A model of domain learning: Reinterpreting expertise as a multidimensional, multistage process. In D.Y. Dai & R.J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 273–298). Mahwah, NJ: Lawrence Erlbaum Associates.
- Alexander, P.A., Jetton, T.L., & Kulikowich, J.M. (1995). Interrelationship of knowledge, interest, and recall: Assessing a model of domain learning. *Journal of Educational Psychology*, 87, 559–575.
- American Psychological Association. (2005). *Report of the 2005 presidential task force on evidence-based practice*. Washington DC.
- Appelbaum, M.I., & Cramer, E.M. (1974). Some problems in the nonorthogonal analysis of variance. *Psychological Bulletin*, 81, 335–343.
- Bakeman, R., McArthur, D., & Quera, V. (1996). Detecting group differences in sequential association using sampled permutations: Log odds, kappa, and phi compared. *Behavior Research Methods, Instruments Computers*, 28, 446–457.
- Baron, R.M., & Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Benjamin, L.S. (2003). *Interpersonal reconstructive therapy: Promoting change in nonresponders*. New York, NY: Guilford Press.
- Bieling, P.J., & Kuyken, W. (2003). Is cognitive case formulation science or science fiction? *Clinical Psychology: Science & Practice*, 10, 52–69.
- Binder, J.L. (2004). *Key competencies in brief dynamic psychotherapy: Clinical practice beyond the manual*. New York, NY: Guilford Press.
- Buchanan, B.G., Davis, R., & Feigenbaum, E.A. (2006). Expert systems: A perspective from computer science. In K.A. Ericsson, N. Charness, P.J. Feltovich & R.R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 87–103). New York, NY: Cambridge University Press.
- Caspar, F., Berger, T., & Hautle, I. (2004). The right view of your patient: A computer-assisted, individualized module for

- psychotherapy Training. *Psychotherapy: Theory, Research, Practice, Training*, 41, 125–135.
- Chi, M.T.H. (2006). Two approaches to the study of experts' characteristics. In K.A. Ericsson, N. Charness, P.J. Feltovich & R.R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 21–30). New York, NY: Cambridge University Press.
- Clark, D.M. (1997). *Cognitive therapy for social phobia: Some notes for therapists*. Oxford: University of Oxford, Department of Psychiatry.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Crits-Christoph, P., Cooper, A., & Luborsky, L. (1988). The accuracy of therapists' interpretations and the outcome of dynamic psychotherapy. *Journal of Consulting and Clinical Psychology*, 56, 490–495.
- Easden, M., & Fletcher, R. (2010). *The relationship between case conceptualization, treatment planning and depressive symptom change: Towards a model of therapist competence in CBT for depression*. Paper presented at the World Congress of Behavioral and Cognitive Therapies.
- Eells, T.D. (2010). The unfolding case formulation: The interplay of description and inference. *Pragmatic Case Studies in Psychotherapy* [Online].
- Eells, T.D., Kendjelic, E.M., & Lucas, C.P. (1998). What's in a case formulation? Development and use of a content coding manual. *Journal of Psychotherapy Practice and Research*, 7, 144–153.
- Eells, T.D., & Lombart, K.G. (2003). Case formulation and treatment concepts among novice, experienced, and expert cognitive-behavioral and psychodynamic therapists. *Psychotherapy Research*, 13, 187–2004.
- Eells, T.D., & Lombart, K.G. (in press). Theoretical and evidence-based approaches to case formulation. In P. Sturmey & M. McMurren (Eds.), *Forensic case formulation*. New York, NY: Wiley.
- Eells, T.D., Lombart, K.G., Kendjelic, E.M., Turner, L.C., & Lucas, C. (2005). The quality of psychotherapy case formulations: A comparison of expert, experienced, and novice cognitive-behavioral and psychodynamic therapists. *Journal of Consulting & Clinical Psychology*, 73, 579–589.
- Elstein, A.S., Shulman, L.S., & Sprafka, S.A. (1978). *Medical problem solving: An analysis of clinical reasoning*. Cambridge, MA: Harvard University Press.
- Ericsson, K.A., Charness, N., Feltovich, P.J., & Hoffman, R.R. (Eds.). (2006). *The Cambridge handbook of expertise and expert performance*. New York, NY: Cambridge University Press.
- Ericsson, K.A., Krampe, R.T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100, 363–406.
- Ericsson, K.A., & Simon, H.A. (1980). Verbal reports as data. *Psychological Review*, 87, 215–251.
- Ericsson, K.A., & Simon, H.A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Ericsson, K.A., & Simon, H.A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge, MA: MIT Press.
- Feltovich, P.J., Prietula, M.J., & Ericsson, K.A. (2006). Studies of expertise from psychological perspectives. In K.A. Ericsson, N. Charness, P.J. Feltovich & R.R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 41–67). New York, NY: Cambridge University Press.
- Fleiss, J.L. (1981). The measurement of interrater agreement. In *Statistical methods for rates and proportions* (2nd ed.) (pp. 212–236). New York, NY: John Wiley & Sons.
- Frazier, P.A., Tix, A.P., & Barron, K.E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology*, 51, 115–134.
- Ghaderi, A. (2006). Does individualisation matter? A randomised trial of standardised (focused) versus individualised (broad) cognitive behaviour therapy for bulimia nervosa. *Behaviour. Research and Therapy*, 44, 273–288.
- Gigerenzer, G., Kurzenghäuser, S., Bibace, R., Laird, J.D., Noller, K.L., & Valsiner, J. (2005). *Fast and Frugal Heuristics in Medical Decision Making* Science and medicine in dialogue: Thinking through particulars and universals (pp. 3–15). Westport, CT: Praeger Publishers/Greenwood Publishing Group.
- Gigerenzer, G., Todd, P.M., & Research Group., ABC (1999). *Simple heuristics that make us smart*. New York, NY: Oxford University Press.
- Glaser, R., & Chi, M.T.H. (1988). Overview. In M.T.H. Chi, R. Glaser & M.J. Farr (Eds.), *The nature of expertise*. xv–xxviii). Hillsdale, NJ: Lawrence Erlbaum.
- Hersen, M., & Porzelius, L.K. (Eds.). (2002). *Diagnosis, conceptualization, and treatment planning for adults: A step-by-step guide*. Mahwah, NJ: Lawrence Erlbaum.
- Horowitz, M.J. (2005). *Understanding psychotherapy change: A practical guide to configurational analysis*. Washington DC: American Psychological Association.
- Hunt, E. (1989). Cognitive science: Definition, status, and questions. *Annual Review of Psychology*, 40, 603–629.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58, 697–720.
- Kaufman, D.R., Yoskowitz, N.A., & Patel, V.L. (2008). Clinical reasoning and biomedical knowledge: Implications for teaching. In J. Higgs, M. Jones, S. Loftus & N. Christensen (Eds.), *Clinical reasoning in the health professions* (3rd ed.) (pp. 137–150). Philadelphia, PA: Elsevier.
- Kendjelic, E.M., & Eells, T.D. (2007). Psychotherapy case formulation training improves formulation quality. *Psychotherapy: Theory, Research, Practice, Training*, 44, 66–77.
- Kuyken, W., Fothergill, C.D., Musa, M., & Chadwick, P. (2005). The reliability and quality of cognitive case formulation. *Behaviour Research and Therapy*, 43, 1187–1201.
- Luborsky, L. (1977). Measuring a pervasive psychic structure in psychotherapy: The core conflictual relationship theme. In N. Freedman & S. Grand (Eds.), *Communicative structures and psychic structures* (pp. 367–395). New York, NY: Plenum Press.
- Luborsky, L., & Barrett, M.S. (2007). *The core conflictual relationship theme: A basic case formulation method* Handbook of psychotherapy case formulation (2nd ed.) (pp. 105–135). New York, NY: Guilford Press.
- McClain, T., O'Sullivan, P.S., & Clardy, J.A. (2004). Biopsychosocial formulation: Recognizing educational shortcomings. *Academic Psychiatry*, 28, 88–94.
- Mergenthaler, E., & Stinson, C.H. (1992). Psychotherapy transcription standards. *Psychotherapy Research*, 2, 125–142.
- Messer, S.B., & Wolitzky, D.L. (2007). The traditional psychoanalytic approach to case formulation. In T.D. Eells (Ed.), *Handbook of psychotherapy case formulation* (2nd ed.) (pp. 67–104). New York, NY: Guilford Press.
- Patel, V.L., & Groen, G.J. (1986). Knowledge based solution strategies in medical reasoning. *Cognitive Science*, 10, 91–116.
- Patel, V.L., Groen, G.J., & Arocha, J.F. (1990). Medical expertise as a function of task difficulty. *Memory and Cognition*, 18, 394–406.
- Patel, V.L., Groen, G.J., & Frederiksen, C.H. (1986). Differences between medical students and doctors in memory for clinical cases. *Medical Education*, 20, 3–9.
- Persons, J.B. (2008). *The case formulation approach to cognitive-behavior therapy*. New York, NY: Guilford Press.
- Persons, J.B., Roberts, N.A., Zalecki, C.A., & Brechwald, W.A.G. (2006). Naturalistic outcome of case formulation-driven

- cognitive-behavior therapy for anxious depressed outpatients. *Behaviour Research and Therapy*, 44, 1041–1051.
- Ryle, A. (1990). *Cognitive analytic therapy: Active participation in change*. Chichester: John Wiley & Sons.
- Schulte, D., Kunzel, R., Pepping, G., & Schulte-Bahrenberg, T. (1992). Tailor-made versus standardized therapy of phobic patients. *Advances in Behaviour Research and Therapy*, 14, 67–92.
- Siegel, S., & Castellan, N.J. Jr. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed). New York, NY: McGraw-Hill.
- Silberschatz, G., & Curtis, J.T. (1993). Measuring the therapist's impact on the patient's therapeutic progress. *Journal of Consulting and Clinical Psychology*, 61, 403–411.
- Silberschatz, G., Curtis, J.T., & Nathans, S. (1989). Using the patient's plan to assess progress in psychotherapy. *Psychotherapy*, 26, 40–46.
- Simon, D.P., & Simon, H.A. (1978). Individual differences in solving physics problems. In R. Siegler (Ed.), *Children's thinking: What develops?* (pp. 215–231). Hillsdale, NJ: Erlbaum.
- Stinson, C.H., Milbrath, C., Reidbord, S.P., & Bucci, W. (1994). Thematic segmentation of psychotherapy transcripts for convergent analysis. *Psychotherapy*, 31, 36–48.
- Weiss, J. (1993). *How psychotherapy works: Process and technique*. New York, NY: Guilford Press.
- Wickens, T.D. (1993). Analysis of contingency tables with between-subjects variability. *Psychological Bulletin*, 113, 191–204.
- Wilson, G.T. (1996). Manual-based treatments: The clinical application of research findings. *Behaviour. Research and Therapy*, 34, 295–314.