

Inadequate and Harmful Clinical Supervision: Testing a Revised Framework and Assessing Occurrence

The Counseling Psychologist

2014, Vol. 42(4) 434–472

© The Author(s) 2013

Reprints and permissions:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/0011000013508656

tcp.sagepub.com



**Michael V. Ellis¹, Lauren Berger¹, Alexa E. Hanus¹,
Erin E. Ayala¹, Brett A. Swords¹,
and Michael Siembor¹**

Abstract

Two studies were conducted to revise and empirically test Ellis's framework for inadequate and harmful supervision, and to determine the occurrence of inadequate and harmful clinical supervision from the supervisees' perspective. For Study 1, we delineated 10 criteria for minimally adequate clinical supervision and defined inadequate and harmful supervision by differentiating self-identified and de facto supervision for each. Ratings from 34 supervision experts were used to generate a taxonomy of 16 de facto inadequate and 21 de facto harmful supervision descriptors. Because harmful supervision was distinct from, yet subsumed by, inadequate supervision, we revised the taxonomy and definitions accordingly. In Study 2, the occurrence of inadequate and harmful supervision was assessed for 363 supervisees; 93.0% were currently receiving inadequate supervision and 35.3% were currently receiving harmful supervision. Over half of the supervisees had received harmful clinical supervision at some point. Implications for research, training, and practice are discussed.

¹University at Albany, Albany, NY, USA

Corresponding Author:

Michael V. Ellis, ED 220, Division of Counseling Psychology, Department of Educational and Counseling Psychology, University at Albany, 1400 Washington Avenue, Albany, NY 12222, USA.

Email: mvellis@albany.edu

Keywords

clinical supervision, psychotherapy supervision, counselor training, therapist training, professional ethics

In recent years, the phenomenon of “clinical supervision that goes badly” has received increasing attention (e.g., Gray, Ladany, & Ancis, 2001; Nelson & Friedlander, 2001; Ramos-Sánchez et al., 2002). In fact, a 5-year review of the clinical supervision literature found that the largest number of articles pertained to harmful supervision (Goodyear, Bunch, & Claiborn, 2005). There are many reasons that bad or harmful clinical supervision is an important topic for discourse, not the least of which is potential harm to clients and supervisees. The supervisee is in an inextricably vulnerable relationship—an evaluative, hierarchical relationship where the supervisor holds the supervisee’s professional career in his or her hands (Bernard & Goodyear, 2014; Koenig & Spano, 2003). Supervisees also may be vulnerable due to unavoidable multiple relationships (Cobia & Boes, 2000; Gottlieb, Robinson, & Younggren, 2007; Hall, 1988). Thus, the supervisee is at risk of harm should a supervisor act in unethical or harmful ways. Although authors are exploring negative experiences in clinical supervision (e.g., Gray et al., 2001; Greer, 2002; Ladany, Friedlander, & Nelson, 2005; Nelson & Friedlander, 2001; Ramos-Sánchez et al., 2002), the findings are scattered, equivocal, and restricted to small samples (e.g., Ellis, 2001; Goodyear et al., 2005; Hutt, Scott, & King, 1983).

One of the problems in the literature has been the lack of clearly defined constructs to conceptualize and study supervision that goes badly or harms supervisees. For example, Ellis (2001) identified over a dozen different terms that continue to be used to describe supervision that goes badly, including negative supervision experiences (Ramos-Sánchez et al., 2002), bad supervision (Jacobsen & Tanggaard, 2009), ineffective supervision (Ladany, Mori, & Mehr, 2013), hindering events (Kaduvettoor, O’Shaughnessy, Mori, Beverly, & Ladany, 2009), and unsuccessful supervisory behaviors (Dressel, Consoli, Kim, & Atkinson, 2007). Unfortunately, the inconsistencies among these constructs across the various studies prohibited a synthesis or comparison of the findings (Ellis, 2001).

Ellis (2001) attempted to bring clarity to the topic by offering a unified framework—a continuum—of two constructs: harmful clinical supervision and bad clinical supervision. Ellis defined harmful supervision as supervisory practices that result in psychological, emotional, and/or physical harm or trauma to the supervisee (e.g., the supervisor’s sexual intimacy, sexual

harassment, or sexual improprieties with a supervisee; aggressive and abusive behavior; violation of the supervisee's boundaries; microaggressions). Ellis defined bad supervision as ineffective supervision that does not traumatize or harm the supervisee, and that is characterized by one or more of the following: the supervisor's disinterest and lack of investment in supervision, the supervisor's failure to provide timely feedback or evaluation of the supervisee's skills, the supervisor's inattention to the supervisee's concerns or struggles, the supervisor does not consistently work toward the supervisee's professional growth or training needs, or the supervisor does not listen and is not open to the supervisee's opinions or feedback. In addition to proffering definitions for bad and harmful supervision, Ellis theorized that bad and harmful supervision could either comprise the poles of a continuum or be related though separate constructs (i.e., a two-dimensional framework).

Ellis's (2001) framework and definitions, however, are problematic. We reasoned that Ellis's constructs needed to be revised to accommodate varying criteria of harmful and bad supervision. Although Ellis's framework provides structure to the topic of "supervision that goes badly," it does not resolve the issue of workable and testable definitions of this phenomenon. The definition of bad supervision lacks a theoretical basis and is vague and not well delineated. Hence, both constructs are difficult to operationalize and test empirically. Perhaps this is why the data regarding bad supervision are scattered and sparse (e.g., Giddings, Cleveland, & Smith, 2007; Greer, 2002; Ladany et al., 2013; Saccuzzo, 2002; Watkins, 1997), and few data exist regarding the occurrence of harmful clinical supervision (e.g., Allen, Szollos, & Williams, 1986; Anonymous, 1991; Burkard, Knox, Hess, & Schultz, 2009; Goodyear et al., 2005).

The deleterious effects of harmful supervision on supervisees may parallel the detrimental effects of harmful therapy to clients (Barlow, 2010; Castonguay, Boswell, Constantino, Goldfried, & Hill, 2010; Dimidjian & Hollon, 2010; Koenig & Spano, 2003; Mays & Frank, 1985). Paralleling the psychotherapy literature (Barlow, 2010), we need to agree upon definitions of harm and bad that are specific to clinical supervision. Nowhere in the literature has this been established here-to-fore. Until a more suitable framework is developed, theoretical, empirical, and clinical progress in understanding bad and harmful supervision will be hindered (cf. Barlow, 2010). Once a viable framework replete with a taxonomy of inadequate and harmful supervision descriptors exists (i.e., operational definitions), initial data regarding the occurrence of inadequate and harmful clinical supervision can be obtained. Thus, given the shortcomings in the literature, the purpose of the current two-study project was (a) to test empirically a framework for inadequate and harmful supervision, and (b) to obtain preliminary data on the

occurrence of inadequate and harmful clinical supervision from the supervisees' perspective.

Study 1: Reconceptualizing and Testing a Framework for Inadequate and Harmful Clinical Supervision

We modified Ellis's (2001) framework in two fundamental ways: (a) by reconceptualizing bad supervision into inadequate clinical supervision (hereafter, inadequate supervision), and (b) by incorporating more objective criteria and self-identification into the definitions of inadequate and harmful supervision. Whether the framework was uni- or multidimensional remained open to empirical investigation. Hence, the purpose of Study 1 was to test the revised framework and attendant constructs, in particular to develop operational definitions of inadequate and harmful clinical supervision that are grounded in theory and consider multiple perspectives (Dimidjian & Hollon, 2010).

Inadequate Clinical Supervision

Theoretical basis for inadequate supervision. To revise the construct of bad supervision, we provided a theoretical basis by anchoring it to the definition of clinical supervision. Bernard and Goodyear's (2014) definition—perhaps the most widely accepted one (Falender & Shafranske, 2004)—states that clinical supervision is

an intervention that is provided by a more senior member of a profession to a more junior colleague or colleagues who typically (but not always) are members of that same profession. This relationship is evaluative and hierarchical, extends over time, and has the simultaneous purposes of enhancing the professional functioning of the junior person(s), monitoring the quality of professional services offered to the clients she, he, or they see, and serving as a gatekeeper for the particular profession the supervisee seeks to enter. (p. 9)

Second, we modified inadequate supervision to fit better with current ethical standards, standards for therapist training, and standards for clinical supervision. To identify what constitutes inadequate supervision, we reasoned that we first needed to delineate *minimally adequate clinical supervision* because it is not explicated in the literature.

Minimally adequate clinical supervision. Regrettably, the American Psychological Association (APA) has not delineated separate ethical or practice standards for clinical supervision (cf. APA, 2002, 2007) as has been done in allied mental health professions. Nonbinding clinical supervision “benchmarks” have been articulated for psychology (Fouad et al., 2009); however, these are insufficient to define minimally adequate clinical supervision. Hence, we drew upon ethical guidelines, accreditation, and licensure standards, and standards for clinical supervision for U.S. psychologists (APA 2002, 2007; Association of Psychology Postdoctoral and Internship Centers [APPIC], 2009a, 2009b; Association of State and Provincial Psychology Boards [ASPPB], 2003, 2009; Crespi & Lopez, 1998), from other U.S. mental health professions (American Association for Marriage and Family Therapy [AAMFT], 2007; American Counseling Association [ACA], 2005; Association for Counselor Education and Supervision [ACES], 1995; Center for Credentialing and Education [CCE], 2009; Council for the Accreditation of Counseling and Related Educational Programs [CACREP], 2009; National Association of Alcoholism and Drug Abuse Counselors [NAADAC], 2008, 2011; National Association of Social Workers [NASW], 2008; National Council on the Practice of Clinical Social Work [NCPCSW], 2003; Substance Abuse and Mental Health Services Administration [SAMHSA], 2007) and our psychology counterparts in other Anglophone countries (Australian Capital Territory Psychologists Board, 2009; Australian Psychological Society [APS], 2003; British Association for Counseling [BAC], 1988; British Psychological Society [BPS], 2003, 2006; Canadian Psychological Association [CPA], 2009; New Zealand Psychologists Board [NZPB], 2009a, 2009b, 2010).

In reviewing the ethics codes for these professional organizations (e.g., APA, 2002), we found limited information regarding clinical supervision. Therefore, we expanded the search to include the requirements and standards for accreditation and licensure, certification, and guidelines and standards for clinical supervision (e.g., AAMFT, 2007; ACES, 1995; BAC, 1988; BPS, 2003, 2006; CCE, 2009; NCPCSW, 2003). Using a consensus validation approach, the team coalesced a list of supervision requirements articulated by and across these documents. Documents that expressed common ideas or concepts using different terminology were combined. Using these standards and requirements, we defined *minimally adequate clinical supervision* (see Figure 1). The 10 criteria in Figure 1 are likely not sufficient for most disciplines; however, we believe these components constitute the bare minimum necessary for clinical supervision, as currently articulated by professional organizations. The criteria for minimally adequate supervision provide the foundation to define inadequate clinical supervision.

The supervisor

- Has the proper credentials as defined by the supervisor's discipline or profession;
- Has the appropriate knowledge of and skills for clinical supervision and an awareness of his or her limitations;
- Obtains a consent for supervision or uses a supervision contract;
- Provides a minimum of 1 hr of face-to-face individual supervision per week;
- Observes, reviews, or monitors supervisee's therapy/counseling sessions (or parts thereof);
- Provides evaluative feedback to the supervisee that is fair, respectful, honest, ongoing, and formal;
- Promotes and is invested in the supervisee's welfare, professional growth and development;
- Is attentive to multicultural and diversity issues in supervision and in therapy/counseling;
- Maintains supervisee confidentiality (as appropriate); and
- Is aware of and attentive to the power differential (and boundaries) between the supervisee and supervisor and its effects on the supervisory relationship.

Figure 1. Criteria for minimally adequate clinical supervision across disciplines.

Drawing on the definition of clinical supervision (Bernard & Goodyear, 2014) and the criteria for minimally adequate supervision, *inadequate clinical supervision* occurs when the supervisor is unable, or unwilling, to meet the criteria for minimally adequate supervision, to enhance the professional functioning of the supervisee, to monitor the quality of the professional services offered to the supervisee's clients, or to serve as a gatekeeper to the profession. In addition, inadequate supervision may include, but is not limited to, the behaviors and descriptors delineated in Ellis's (2001) definition of bad supervision.

Self-identified and de facto definitions. To incorporate subjective and more objective perspectives of inadequate supervision, we differentiated *self-identified* and *de facto* inadequate supervision (SIIS and DFIS). As the name suggests, SIIS occurs if, after reading the definition of inadequate supervision, a supervisee declares that he or she has received inadequate supervision. De facto inadequate supervision (DFIS) is defined as the supervisor's failure to provide the minimal level of supervisory care as established by his or her discipline or profession, by law (Giddings et al., 2007; Greer, 2002; Sacuzzo, 2002) or by failure to meet the minimally adequate supervision criteria (Figure 1). Thus, for DFIS, a supervisee does not have to identify or label his or her supervision as inadequate. Rather, the supervisee's endorsements

of supervision descriptors, including supervisor's behaviors or inactions, become the criteria for establishing if the supervision he or she received was inadequate. Usually, inadequate supervision refers to an ongoing supervisory situation or relationship—it may encompass one truly inadequate session or incident. Inadequate supervision can occur in individual, group, or supervisor supervision. It may entail a poor-quality supervisory relationship, and may be harmful to the supervisee's client

Harmful Clinical Supervision

Paralleling the revisions to inadequate supervision, we expanded Ellis's (2001) definition of harmful supervision to go beyond self-identification by incorporating the supervisor's actions or inactions that are "known" to cause harm (a general consensus that the action or inaction typically results in harm). We defined *harmful supervision* as supervisory practices that result in psychological, emotional, and/or physical harm or trauma to the supervisee. Harmful supervision can be through self-identification (*self-identified* harmful supervision [SIHS]) or occur when the supervisor's behavior (or inaction) meets specific criteria (*de facto* harmful supervision [DFHS]). The two essential components of harmful supervision are (a) that the supervisee was genuinely harmed in some way by the supervisor's inappropriate actions or inactions, or (b) the supervisor's behavior is known to cause harm even though the supervisee may not identify the action as harmful. Thus, harmful supervision may result from the supervisor acting inappropriately or with malice, supervisor negligence, or the supervisor clearly violating accepted ethical standards and standards of practice and care (e.g., Dye & Borders, 1990; Ladany, Lehrman-Waterman, Molinaro, & Wolgast, 1999).¹

Harmful supervision should be distinguished from those instances where a supervisee struggled with painful issues in supervision, or when a supervisor gave painful to hear, emotionally upsetting feedback about the supervisee's professional incompetence that was necessary for the supervisee's professional growth (Ellis, 2001), or to protect client or public welfare (i.e., serving in the gatekeeping role; for example, Nelson, Barnes, Evans, & Triggiano, 2008). We are attempting to differentiate between the supervisor's actions that were respectful of the supervisee's boundaries, and focused on the supervisee's professional development within the context of a positive supervisory relationship, from those instances where the supervisee's best interests were not primary. Harmful supervision can consist of one or more incidents, or can be an ongoing supervisory situation. Harmful supervision can occur in individual or group supervision, clinical supervision, or supervisor supervision, and with one or more supervisors.

Keeping with Ellis (2001), harmful supervision may include sexual improprieties or sexual intimacy with the supervisee (e.g., Bartell & Rubin, 1990; Celenza, 2007; Lamb, Catanzaro, & Moorman, 2003); the supervisor acting physically, emotionally, or psychologically aggressive and abusive; violating the supervisee's boundaries (e.g., emotional intimacy forced upon the supervisee, revealed personal information about the supervisee to his or her clients; Koenig & Spano, 2003); using power for personal gain at the supervisee's expense; making macro- and microaggressions toward the supervisee (e.g., blatant racism, homophobia; Burkard et al., 2009; Sue et al., 2007); publicly humiliating and deriding the supervisee; demeaning, critical, and vindictive attitude toward the supervisee; engaging in an exploitative multiple relationship that caused the supervisee harm (Gottlieb et al., 2007; Hall, 1988); and failing to take action resulting in harm to the supervisee or client.

The effects of harmful supervision incidents or experiences include symptoms of psychological trauma (e.g., prevailing sense of mistrust, debilitating fears, or excessive shame, guilt, and self-derogation), conspicuous loss of self-confidence, functional impairment in the supervisee's professional or personal life, and a significant decline in the supervisee's general mental or physical health. The effects of the harmful experience may last a short time (a couple of days) or may persist for months to years even after seeking therapy to deal with the supervisee's reactions to the situation. Harmful supervision practices may harm clients as well.

Testing the Definitions and Framework

With the revised framework and definitions in place, the next step was to test the framework and constructs. Of particular importance was to establish the criteria that constitute DFIS and DFHS. Making judgments determining whether supervision was inadequate or harmful is an onerous task because of the potentially dire consequences for supervisors (e.g., supervisor's professional reputation; civil or criminal legal action) and for supervisees (e.g., harmed clients). It seemed prudent, therefore, to use criteria that supervision experts judge as clearly inadequate or harmful. Thus, we sought to develop a taxonomy (classification system) of de facto inadequate and harmful supervision descriptors; we were not developing a new scale or measure.

Given the construct definitions, we hypothesized that supervision experts would (a) rate supervision descriptors as either highly inadequate or harmful, and (b) judge the specific supervision descriptors for inadequate supervision as clearly inadequate and substantively different from harmful supervision, and vice versa for harmful supervision. Furthermore, we formulated competing hypotheses regarding the dimensionality or structure of the inadequate

and harmful supervision framework. As stipulated by Ellis (2001), either the two constructs anchor the ends of a continuum of inadequate to harmful clinical supervision (i.e., one dimension), or they are related yet distinct dimensions (i.e., two correlated dimensions). That is, if the framework were unidimensional, ratings of supervision descriptors on the inadequate construct would be inversely related to ratings of the same descriptor on the harmful construct and would fall along a single dimension. If the framework were two-dimensional, ratings on each construct would be correlated with the ratings substantially higher on the target construct than the other construct.

Method

Power analysis. To determine the desirable sample size given the research design, we performed an a priori statistical power analysis (Cohen, 1988). Due to the lack of empirical literature on inadequate and harmful supervision to determine an effect size, we used a large effect size for the counseling psychology literature— $\rho^2 = .189$ (ρ^2 is the shrunken effect size; Haase, Ellis, & Ladany, 1989). Thus, a sample of at least 35 participants with a Type I error rate of $\alpha = .05$, and an estimated population effect size of $\rho^2 = .189$, yielded an expected a priori statistical power of .82 for the series of *t* tests.

Participants. Our target population was clinical supervision experts. To be considered a clinical supervision expert, a participant had to identify himself or herself a supervision expert, be formally trained in clinical supervision (e.g., coursework in clinical supervision, supervisor practicum, training in clinical supervision), or have accrued more than 20 years of experience as a clinical supervisor in the absence of formal training, have accrued more than 3 years of experience as a clinical supervisor, and have supervised more than 10 therapist–counselor supervisees in their career. A few well-established experts in clinical supervision (e.g., Drs. Janine Bernard, Carol Falender, Micki Friedlander, M. Lee Nelson) endorsed these inclusion criteria as appropriate (via informal discussions with the authors). Participants who failed to meet these criteria were excluded from the study; 46 supervisors completed the research materials, of whom 34 met the criteria as a supervision expert. The recomputed a priori statistical power for 34 participants was .81.

The 34 supervision experts had a mean age of 51.29 years old ($SD = 9.66$); 72.7% were female. The majority were non-Hispanic White (87.9%); 6.1% were Hispanic and 6.1% were Biracial. About two thirds (68.7%) earned a PhD, PsyD/DPsy, or EdD (28.4% MSW, MA, MS, or MEd, and 2.9% BA). Their professional fields included Counseling Psychology (32.4%), Clinical Psychology (20.6%), Rehabilitation or Mental Health Counseling (11.7%),

Social Work (17.6%), or Counselor Education (5.9%). About 80% of the supervisors were currently licensed (5.9% not licensed), and 43.8% were currently certified (18.8% not certified). On average, they had worked 22 years as a mental health provider ($SD = 10.25$, $Mdn = 24.0$) in an academic setting (32.4%), private practice (23.5%), college counseling center (8.8%), community agency (17.4%), or a substance abuse facility (17.4%). At the work sites, 70.6% of the supervisors identified their primary roles as clinical supervisors, 44.1% as instructors/professors, 32.4% as administrators, 38.2% as counselors, and 17.6% in management. The participants reported supervising for an average of 12.09 years ($SD = 8.51$, $Mdn = 9.08$). Supervisors endorsed an eclectic theoretical orientation (23.5%), followed by cognitive-behavioral (20.6%), humanistic (17.6%), psychodynamic (11.8%), systems (11.8%), and interpersonal (8.8%) orientations.

When reporting information about their specific history as a clinical supervisor, they had on average 15.56 years of experience as a clinical supervisor ($SD = 9.49$, $Mdn = 14.25$). Overall, they supervised an average of 47.26 supervisees ($SD = 32.58$, $Mdn = 40.0$) while currently supervising 7.64 supervisees ($SD = 18.82$, $Mdn = 5.0$). The supervisors were providing an average of 4.09 individual supervision hours per week ($SD = 4.36$, $Mdn = 2.5$) and 1.54 group supervision hours per week ($SD = 1.84$, $Mdn = 1.0$). All but one were formally trained in clinical supervision (97.0%), receiving an average of 3.29 years of supervision training ($SD = 4.21$, $Mdn = 2.0$ years). In terms of the clinical supervision training received, 91.2% of the supervisors attended workshops, 85.3% acquired continuing education, 64.7% took a course, 64.7% completed supervised supervision practical, 50% were self-taught, and 11.6% were trained through consultation or research. On average, they attended 5.12 courses ($SD = 12.24$, $Mdn = 2.0$) focused on multicultural clinical supervision. They also authored supervision articles in peer-reviewed journals (35.3%), non-peer-reviewed journals (23.5%), books (8.8%), and book chapters (17.6%). They led supervision courses (50%), presented on the topic of supervision at workshops (47.1%), and presented supervision posters/papers at conferences (47.1%).

Measures. The research team comprised eight graduate students and one counseling psychologist professor in an APA-accredited counseling psychology program. Following a consensus validation approach, the team derived a pool of supervision descriptors using the revised construct definitions and criteria for minimally adequate supervision. The team evaluated the extent to which each proposed descriptor matched the inadequate or harmful definition and revised the pool of descriptors until consensus was achieved. This

process resulted in a comprehensive taxonomy of 50 descriptors—27 for DFIS and 23 for DFHS (see Table 1).

Participants referred to definitions of the two constructs on each page and rated the 50 supervision descriptors “To what extent is this Harmful supervision?” and “To what extent is this Inadequate supervision?” Specifically, participants were asked to “rate each item on each of the two constructs using the 1 to 7 rating scale below. A 1 means *not at all*, 4 means *moderately*, and 7 means *totally*. Decide to what extent each item is Harmful Supervision and is Inadequate Supervision.”

Procedures. We solicited mental health counselors and therapists who self-identified as meeting the criteria as a clinical supervision expert, including professionals and advanced supervisors-in-training, who were at least 18 years of age and currently providing clinical supervision or supervisor supervision to participate in the study. Potential participants were solicited via listservs (e.g., the Society of Counseling Psychology Supervision and Training Section listserv, Council of Counseling Psychology Training Programs [CCPTP] listserv, ACES listserv) and individual emails to colleagues and professionals in clinical supervision. Participants were directed to a password-protected website to complete the informed consent and research materials. The response rate was unknown because we could not estimate the number of supervisors reached by the listservs and emails.

Results

The means and standard deviations for the ratings on the 50 inadequate and harmful supervision descriptors are reported in Table 1. To test the hypothesis that supervision experts would rate the DFIS descriptors significantly and substantively higher on the inadequate supervision construct than the harmful construct, we conducted a series of dependent samples one-tailed *t* tests. To control the study-wise Type I error rate, we used a modified Bonferroni procedure that preserves statistical power (Holland & Copenhaver, 1988). Per Table 1, all but 2 of the 27 tests were statistically significant and observed an effect size greater than $\bar{\rho}^2 = .190$. One of the two supervision descriptors not achieving criteria was rated higher on the harmful supervision construct than on the inadequate supervision construct (i.e., supervisor pathologizes me in evaluations), suggesting that it pertained to harmful supervision rather than inadequate supervision.

Similarly, we performed a series of one-tailed dependent samples *t* tests to test the hypothesis that supervision experts would rate the DFHS descriptors significantly and substantively higher on the harmful construct relative to the

Table 1. Inadequate and Harmful Supervision Descriptor Ratings for Studies 1 and 2.

Supervision Descriptor	Study 1							Study 2	
	Harmful		Inadequate		<i>t</i> (33)	<i>p</i>	$\bar{\rho}^2$	<i>M</i>	<i>SD</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>					
Inadequate supervision									
Does not know what to do ^l	3.41	1.78	6.82	0.39	-11.42	.0001	.792	1.70	1.31
Supervising my supervisor ^l	3.88	1.94	6.76	0.43	-8.76	.0001	.690	1.47	1.03
Never spend time improving skills ^l	2.91	1.66	6.71	0.63	13.28	.0001	.838	2.00	1.49
Clients suffered emotional trauma because of supervision	5.44	1.89	6.65	0.69	-4.05	.003	.312	1.17	0.78
Oblivious to cultural background ^l	4.71	1.59	6.62	0.78	-6.58	.0001	.554	1.93	1.52
Refuses to address issues ^l	4.21	1.51	6.50	0.62	-9.68	.0001	.731	1.58	1.08
Does not discuss difficulties with clients ^l	3.38	1.63	6.50	0.71	11.04	.0001	.780	5.33	1.57
Not provided adequate supervision for clients ^l	3.82	1.49	6.47	0.83	11.87	.0001	.804	1.94	1.50
No evaluative feedback ^l	3.03	1.62	6.36	0.74	-11.87	.0001	.809	2.03	1.44
Supervision is waste of time ^l	3.18	1.57	6.15	1.23	-9.28	.0001	.714	1.73	1.41
No interest in cultural background ^l	3.97	1.79	6.15	0.96	-7.12	.0001	.593	1.97	1.64
Oblivious to interpersonal process ^l	3.38	1.46	6.12	0.88	-10.20	.0001	.752	1.76	1.42
Behaves unethically ^l	5.62	1.61	6.12	1.34	-1.35	.187	.023	1.30	1.02
Does not meet for 1 hr per week ^l	2.47	1.71	6.09	1.11	-11.55	.0001	.795	1.75	2.87
Not committed ^l	3.32	1.43	6.00	0.92	-10.06	.0001	.746	1.74	1.42
Does not listen	3.82	1.40	5.85	0.92	-7.05	.0001	.588	1.69	1.35
Frequently distracted	2.88	1.45	5.76	0.90	-13.57	.0001	.847	2.22	1.45
Locked in conflict	4.12	1.86	5.71	1.72	-3.20	.003	.213	1.41	1.04
Discusses strengths ^R	2.94	1.48	5.66	1.31	-9.43	.0001	.733	4.88	1.71
Unclear what to do	2.65	1.23	5.50	0.99	-12.31	.0001	.815	2.08	1.46
Never discusses professional development	2.44	1.19	5.47	1.19	-15.13	.0001	.870	1.88	1.39
Never observed sessions ^l	2.65	1.63	5.35	1.56	-10.46	.0001	.761		
Highly skilled ^R	2.79	1.37	5.32	1.34	-8.63	.0001	.683	5.48	1.70
Focus only on diagnoses	2.15	1.02	5.18	1.40	-11.88	.0001	.804	2.19	1.29
Not use consent or contract ^l	2.94	1.67	5.12	1.63	-9.19	.0001	.710		
Relationship is cold and distant	4.18	1.64	5.09	1.55	-3.21	.003	.219	1.65	1.41
Treats me with respect ^R	3.97	1.43	4.58	1.48	-1.95	.060	.078	5.81	1.51
Harmful supervision									
Threatened me physically ^H	7.00	0.00	5.91	1.75	3.63	.001	.263	1.01	0.21
Have a sexual relationship ^H	6.91	0.38	5.88	1.67	3.74	.001	.276	1.02	0.32
Have been sexually intimate ^H	6.85	0.44	6.06	1.63	2.69	.011	.155	1.00	0.00
Is aggressive and abusive ^H	6.85	0.44	5.91	1.64	3.44	.002	.241	1.20	0.83
Harmed by supervisor's actions ^H	6.84	0.44	5.53	2.11	1.96	.001	.264	1.39	1.00
Traumatized by supervision ^H	6.76	0.50	5.50	1.94	3.83	.001	.286	1.39	1.01
Dual relationship was harmful ^H	6.74	0.57	5.53	1.80	3.97	.0001	.302	1.28	1.07
Supervisor sexually inappropriate ^H	6.74	0.62	5.94	1.61	2.81	.008	.168	1.03	0.34
Supervision is harmful ^H	6.74	0.67	5.62	1.78	3.64	.001	.265	1.43	1.12

(continued)

Table 1. (continued)

Supervision Descriptor	Study 1						Study 2		
	Harmful		Inadequate		t(33)	p	p ²	M	SD
	M	SD	M	SD					
Safe from exploitation ^{R,H}	6.68	0.73	5.50	1.97	3.37	.001	.248	5.74	2.11
Harmed by inactions ^H	6.59	0.44	5.91	2.11	3.53	.058	.077	1.48	0.98
Feel exploited ^H	6.55	0.75	5.52	1.66	3.37	.002	.238	1.42	1.23
Is cruel ^H	6.47	0.83	4.85	1.94	4.75	.0001	.388	1.24	0.80
Evaluations are victimizing ^H	6.47	0.86	5.53	1.83	3.01	.005	.191	1.23	0.90
Violated sense of safety ^H	6.44	0.86	5.50	1.81	3.12	.004	.204	1.40	1.13
Feel guilt, embarrassment, shame, or blame ^H	6.44	0.75	4.76	2.05	4.59	.0001	.371	1.50	1.19
Avoids exploitative dual roles ^{R,H}	6.26	0.90	5.65	1.81	1.89	.068	.070	5.67	2.01
Used drugs together ^H	6.24	1.21	5.91	1.69	0.95	.348	.000	1.00	0.00
Publicly humiliated ^H	6.18	1.19	5.21	1.93	2.87	.007	.179	1.27	0.95
Discriminating toward me ^H	6.15	1.60	6.32	1.34	-0.46	.650	.000	1.14	0.65
Pathologizes me ^H	6.06	1.20	5.18	1.78	2.64	.013	.149	1.61	1.17
Drunk together	5.94	1.56	5.56	1.78	1.02	.316	.001	1.04	0.38
Feel safe with supervisor ^R	5.82	1.24	5.12	1.80	1.99	.055	.080	5.77	1.65

Note. Means were derived across the entire sample, not subdivided by those who self-identified as receiving harmful or inadequate supervision or by those who met criteria for de facto harmful or inadequate supervision.

^HDe facto harmful.

^IDe facto inadequate.

^{RH}Originally inadequate, but switched to harmful.

^{RI}Originally harmful but switched to inadequate.

^RReverse scored.

inadequate construct. Seven of the 23 *t* tests were nonsignificant and 3 more failed to achieve a shrunken effect size of $\bar{p}^2 = .190$ or larger (see Table 1). Contrary to predicted differences, two of these had higher ratings on inadequate supervision than harmful supervision (i.e., supervisor behaves unethically; supervisor is blatantly discriminating) suggesting that these supervision descriptors tapped inadequate supervision rather than harmful supervision. Testing the dimensionality of the inadequate and harmful supervision framework, the correlation between the harmful construct mean ratings and the corresponding inadequate mean ratings was not significant, $r(50) = -.184$, $p = .20$, $\bar{p}^2 = .014$. This, plus an inspection of Table 1, revealed that contrary to our hypotheses, the pattern of ratings for the supervision descriptors did not conform to either a unidimensional or a two-dimensional conceptualization as theorized by Ellis (2001). Rather, in general, harmful descriptors were rated as inadequate and harmful, whereas inadequate descriptors were rated as solely inadequate. Thus, the supervision experts distinguished between

harmful and inadequate supervision, albeit, in a different way than previously conceptualized.

Reconceptualizing the Framework and Selecting de facto Supervision Descriptors

The pattern of supervisor mean ratings required us to reconceptualize the framework for inadequate and harmful clinical supervision. A perusal of Table 1 revealed that all but three supervision descriptors were rated as high-moderately inadequate (i.e., means greater than 5.0; a score of 4 was moderately inadequate) regardless of its original designation (harmful or inadequate). Across all supervisor descriptors, the mean ratings ranged from 2.15 to 7.00 on the harmful construct and from 4.58 to 6.82 on the inadequate construct. For the originally designated harmful supervision descriptors, the mean harmful ratings ranged from 5.62 to 7.00, whereas the mean inadequate ratings ranged from 2.15 to 6.06. Hence, the supervision experts differentiated harmful supervision from inadequate supervision. Taken in combination with the correlation of the inadequate and harmful ratings ($p^2 = .014$), the data suggested that inadequate clinical supervision subsumes harmful clinical supervision. That is, all harmful supervision is by definition inadequate supervision. In response to an open-ended question about the study, it is noteworthy that 4 out of 18 supervisors (22.2%) who responded to the question proffered this conceptualization of inadequate and harmful supervision. The team also switched two supervision descriptors that were rated higher on one construct to the corresponding construct. Thus, supervision descriptors were categorized as either inadequate or harmful and not both (i.e., mutually exclusive categories).

The research team for this phase of the project included one counseling psychologist with expertise in clinical supervision, six counseling psychology doctoral students, and three masters' counseling students. We took a conservative approach to determine operational definitions for DFHS and DFIS. That is, the experts in Study 1 had to judge the supervision descriptor as unmistakably inadequate and/or harmful. The team identified and used four criteria to evaluate the supervision descriptors. First, the supervision descriptor had to be vital to its respective construct definition, and in the case of DFIS, it also had to be an obvious violation of minimally adequate supervision (Figure 1). Second, the minimum criteria to be selected for DFIS and for DFHS were a mean rating of 6.0 or higher on its respective construct. The team reasoned that a mean rating of 6.0 on a 7-point scale, where a 7.0 is totally inadequate or harmful was clear evidence that the supervision

descriptor was inadequate or harmful, respectively. To be consistent with the revised definitions and framework, the selection criteria for DFHS stipulated that the supervision descriptor also be rated 4.75 or higher on the inadequate construct (at least high-moderately inadequate).

Third, the team initially sought supervision descriptors with mean ratings that were statistically significant and substantively distinct. The statistical significance and effect size criteria differed, however, for the two sets of de facto supervision descriptors due to reconceptualizing the inadequate-harmful supervision framework. Because harmful supervision descriptors were also rated highly on the inadequate construct, the mean ratings on both constructs were close to the maximum rating (e.g., between 5.0 and 7.0 on a 7-point scale). Because attenuated (smaller) effect sizes are not likely to be statistically significant except for very large samples, the team did not automatically exclude harmful descriptors that evidenced nonsignificant effect sizes less than $\bar{p}^2 = .189$. For DFIS, effect sizes needed to be unequivocally inadequate and thus exceed $\bar{p}^2 = .50$ (accounting for more than 50% of the variance in the differences of the mean ratings).

Fourth, the team classified supervision descriptors that had a mean inadequate rating of 6.0 or greater and a harmful rating between 3.8 and 6.0 (at least moderate harm) as DFIS and in the boundary between harmful and inadequate supervision. That is, these supervision descriptors were clearly indicative of DFIS but were not de facto harmful (i.e., inadequate descriptors that lie in the boundary between inadequate and harmful supervision).

Ultimately, the team selected 16 DFIS descriptors (superscript I in Table 1) and 21 DFHS descriptors (superscript H in Table 1). Recall that the supervision experts judged 35 of the 37 supervision descriptors as undeniably inadequate supervision. The two exceptions were “not use consent or contract” and “never observed sessions.” We included these two supervision descriptors for mainly two reasons. First, nearly every discipline mandates or calls for the use of a supervision consent or supervision contract (e.g., AAMFT, 2003; Bernard & Goodyear, 2014; Sutter, McPherson, & Geeseman, 2002; Thomas, 2007), largely due to legal liability and ethical concerns (e.g., Saccuzzo, 2002, 2003). Second, an essential component of effective supervision is observing the supervisee’s in-session work (e.g., Bernard & Goodyear, 2014; Huhra, Yamokoski-Maynhart, & Prieto, 2008; Noelle, 2002)—failure to do so is also a legal liability concern.

Supervision descriptors classified in the inadequate-harmful boundary included oblivious to cultural background, no interest in cultural background, refuses to address issues, supervising my supervisor, not provided adequate supervision for clients, and behaves unethically. One descriptor met the criteria but was not included in the operational definition of DFIS—clients

suffered emotional trauma because of supervision. The team judged it an outcome of inadequate supervision—its focus is on the client rather than supervision itself.

We retested the inadequate and harmful supervision framework using the de facto supervision descriptors. The correlation between the harmful construct mean ratings and the corresponding inadequate mean ratings for the 37 final de facto supervision descriptors was significant, $r(36) = -.59, p = <.001$, $\rho^2 = .300$. On the surface, this result appeared to support the hypothesis of a bipolar dimensional framework. The pattern of the 37 mean ratings in Table 1, however, suggested otherwise—all supervision descriptors were DFIS, hence, the correlation was consistent with and affirmed the reconceptualized framework.

Discussion

In Study I, we sought to operationalize and test Ellis's (2001) framework of inadequate and harmful clinical supervision as well as create operational definitions for DFIS and DFHS. To facilitate interpreting the results within the context of the strengths and limitations of the study, these are discussed first.

Limitations. First, we acknowledge that the inclusion criteria used in defining supervision experts were derived subjectively. In part, this was due to a general lack of explicitly defined guidelines for such a qualification. The guidelines used for the current study included similar, if not more stringent, guidelines than other published clinical supervision articles with expert supervisor samples (e.g., Grant, Crawford, & Schofield, 2012). A different pattern of harmful and inadequate supervision descriptors might have emerged if a different set of inclusion criteria for determining expertise was used. When we implemented more and less stringent criteria for a supervision expert, however, the results evidenced no salient changes. The results could also vary with a different or larger sample of supervisors. We note that common practice is for a task force or committee to develop taxonomies; however, these are not tested empirically as was done here (e.g., Falender et al., 2004; Fouad et al., 2009).

A second limitation was that we took a conservative approach in selecting supervision descriptors to define DFHS and DFIS. Some might argue that descriptors with mean rating values less than 6.0 ought to be included as either inadequate or harmful supervision (see Table 1). We, however, sought unambiguous criteria and definitions for DFIS and DFHS that few could remonstrate as not being harmful or inadequate. Nevertheless, others may

consider some supervision descriptors excluded from the de facto list as clearly inadequate or harmful (e.g., not feel safe with my supervisor).

Finally, the proposed definitions may not account for the relevance of context in distinguishing between inadequate and harmful supervision. In other words, DFIS and DFHS may be dependent on multiple situational factors not accounted for by simple application of the proposed de facto definitions. For instance, a potentially harmful impasse that is successfully resolved between the supervisee and supervisor would not be considered harmful supervision. If it were not resolved appropriately, however, the consequences could meet criteria for harmful supervision. Thus, the potential to harm supervisees is dependent on multiple criteria, and it is possible that these supervision descriptors are inadequate and minimally harmful in some situations and clearly harmful in others.

Strengths. We undertook a multifaceted approach to conceptualizing and testing empirically inadequate and harmful supervision and their operational definitions. Specifically, we (a) explicated a priori theorizing and falsifiable hypotheses (Wampold, Davis, & Good, 1990); (b) synthesized the professional ethics codes and standards for practice, accreditation, and certification-licensure from psychology in the United States and internationally, as well as allied U.S. mental health professions to formulate 10 criteria for minimally adequate supervision; (c) grounded the definition of inadequate supervision on the definitions of clinical supervision and minimally adequate clinical supervision; (d) used ratings from supervision experts to derive a taxonomy of DFIS and DFHS descriptors; (e) implemented team consensus validation procedures throughout the study (e.g., to revise the construct definitions, to select the de facto inadequate and harmful supervision descriptors, and create the taxonomy); and (f) systematically controlled study-wise Type I and II error rates.

Although taking a conservative approach to define DFHS and DFIS imposed some limitations, we believed this was also a strength of the study because it decreased ambiguity within the constructs and their operational definitions. We included the supervision descriptors captured in minimally adequate supervision (i.e., the professional ethical codes and standards) and that had supervisors' mean ratings of 6.0 or greater on the 7-point scale. We also assessed the degree to which descriptors were clearly harmful or inadequate by considering effect sizes. Hence, the DFIS and DFHS descriptors were theoretically and empirically based.

Major findings. Three findings from Study 1 were salient. First, the definitions and framework for inadequate and harmful clinical supervision were in part

inaccurate and required revision. Specifically, the previous conceptualizations of harmful and inadequate supervision assumed a level of mutual exclusivity, such that negative occurrences in supervision may be *either* inadequate *or* harmful (Ellis, 2001). However, the current data provided by 34 supervision experts suggested that inadequate supervision subsumes harmful supervision (i.e., harmful supervision is by definition inadequate supervision). As such, inadequate supervision has the potential to induce up to a moderate level of harm before crossing the threshold of clearly harmful supervision (see Table 1).

The boundary between inadequate and harmful supervision merits attention because the data suggested that the boundary might not be sharply demarcated. Whereas harmful supervision is inadequate and harmful, the boundary area encompasses those supervision descriptors that were judged inadequate and moderately harmful. These descriptors were nonetheless classified as DFIS versus DFHS. In short, the first major finding was an empirically and theoretically grounded conceptualization and framework for inadequate and harmful supervision.

The second major finding was the compilation of an empirically based taxonomy of 37 supervision descriptors as viable operational definitions of DFIS and DFHS. We modified Ellis's (2001) framework by differentiating and incorporating self-identified and de facto perspectives into the definitions of inadequate and harmful supervision. Overall, the inadequate and harmful supervision descriptors seemed to capture adequately the definitions for DFHS and DFIS. The evidence further affirmed that the construct definitions for DFHS and DFIS were largely consistent with the supervisors' working notion of these concepts (i.e., rating data). However, more inadequate descriptors failed to meet criteria for inclusion in the taxonomy than harmful descriptors. Supervision experts judged the taxonomy of 16 inadequate descriptors and 21 harmful descriptors as clearly inadequate or harmful, respectively. Thus, the results suggested that the taxonomy of supervision descriptors could be used to evaluate and classify supervision as DFIS and/or DFHS independent of the supervisee's self-identification—named the Taxonomy of Inadequate and Harmful Clinical Supervision (TIHCS).

The third set of notable findings concerned ratings by the supervisors that seemed inconsistent with established ethical guidelines. That is, although the majority of professional ethics and standards mandated that supervisors use a supervision consent or contract (e.g., AAMFT, 2003; ACA, 2005; APA, 2002; ASPPB, 2003; Bernard & Goodyear, 2014; Sutter et al., 2002; Thomas, 2007), the U.S. supervisors rated the failure to do so as moderately inadequate ($M = 5.12$). In addition, they rated the supervisor's failure to observe the supervisee's therapy/counseling sessions as moderately inadequate ($M =$

5.35). Somewhat puzzling, most of the supervisors used a consent or contract for supervision (64.7%) and directly monitored their supervisee's sessions (73.5%), yet did not rate the lack of these as clearly indicative of inadequate supervision. Further, a few supervisors (11.8%) considered sexual contact with a supervisee as less than totally harmful even though such behavior is unethical (e.g., Bartell & Rubin, 1990). In addition, approximately one third of the supervisors indicated that using drugs or being drunk with a supervisee was not at all to moderately harmful or inadequate supervision. Together, these ratings revealed a disconnect between what is delineated in ethical and accreditation guidelines as well as the clinical supervision literature regarding adequate supervision, and what is endorsed by clinical supervision experts.

Why does this apparent disconnect exist? We do not know. Two explanations seem plausible. First, the supervisor ratings may be due to a lack of guidelines for psychologists in the United States regarding the practice of clinical supervision. Fortunately, such guidelines are currently being developed, which ultimately may lead to implementing competency-based clinical supervision more pervasively (APA Board of Educational Affairs [BEA] Task Force on Supervision Guidelines, 2013). The gap may also be a reflection of the supervision literature. Heretofore, criteria for minimally adequate and harmful supervision have not been well delineated, especially for psychology in the United States, and the literature lacked coherent, theoretically and empirically derived constructs and definitions thereof for inadequate, harmful, and minimally adequate supervision. Thus, supervisees and supervisors have been largely uninformed about these aspects of supervision practice. Without question, further research is needed to understand the disconnect.

Study 2: Occurrence of Inadequate and Harmful Clinical Supervision

Hence, with a viable taxonomy of DFIS and DFHS descriptors and the revised framework and definitions of inadequate and harmful supervision, we turned to Study 2. Few data exist regarding the perceived occurrence of inadequate or harmful clinical supervision (e.g., Allen et al., 1986; Anonymous, 1991; Ellis, 2001; Hutt et al., 1983). Some evidence suggests that 33% to 50% of supervisees are likely to encounter harmful supervision and that 7% to 10% of supervisees will leave the field due to harmful supervision (e.g., Barnett-Queen & Larrabee, 2000; Gray et al., 2001; Ladany et al., 1999; Larrabee & Miller, 1993; Moskowitz & Rupert, 1983; Nelson & Friedlander, 2001). These investigators studied a variety of constructions of "problematic

supervision” and lacked a coherent conceptual framework to guide their investigations, nor did they differentiate inadequate supervision from harmful supervision. Thus, the findings pertained to self-identified problems and were somewhat ambiguous and potentially misleading, leaving a deficiency in the literature. The purpose of Study 2, therefore, was to obtain initial data regarding the occurrence of inadequate and harmful clinical supervision from a diverse sample of supervisees in mental health fields.

Method

Participants. The sample, which was a subsample of a larger study of the supervisory relationship, consisted of 363 supervisees. A majority of the sample was female (81.8%), with a mean age of 34.76 years ($SD = 10.98$, $Mdn = 31.0$). The majority (79.9%) of the participants were non-Hispanic White, with 4.7% African American, 4.5% Asian/Pacific Islander, 4.5% Hispanic/Latina, 1.1% Native American, 0.6% Middle Eastern, and 4.7% Other. Most of the participants (56.7%) held a master’s degree (e.g., MA, MS, MEd, or MSW) and 7.0% had earned their doctorate (PhD, PsyD, or EdD). Of those currently in an academic program (74.2%), participants were on average in their second year of study ($M = 2.55$, $SD = 1.48$) in a doctoral program (42.9%; 23.8% PhD, 18.83% PsyD/DClinPsy, 0.3% EdD), with 20.9% pursuing an MA/MS/MEd, 3.5% MSW, or 4.0% undergraduate degree (AA or BA/BS). Participants were from various fields of study, including Clinical Psychology (26.3%), Counseling Psychology (19.3%), Mental Health Counseling (14.5%), Social Work (8.9%), Substance Abuse (9.2%), School Psychology (6.1%), and School Counseling (3.6%).

In terms of clinical experience, participants were in a pre-practicum (3.5%), first practicum (11.7%), advanced practicum (12.4%), master’s internship (10.5%), post-master’s internship (12.1%), pre-doctoral internship (13.3%), post-doctoral internship (6.7%), other (2.2%), or “Not Applicable” (27.6%). The placement settings included community mental health centers (36.2%), substance abuse treatment facilities (13.8%), community-based agencies (9.6%), college counseling centers (9.0%), university-based training centers (7.1%), primary or secondary schools (7.1%), hospitals (3.5%; public, Veterans Administration [VA], or private), private practice (4.0%), forensic/prisons (3.1%), or other settings (6.8%). The supervisees had an average of 5.88 years of clinical training ($SD = 5.95$, $Mdn = 4.3$ years), 4.27 years of supervised training ($SD = 4.83$, $Mdn = 3.2$), and had worked with an average of 5.14 clinical supervisors ($SD = 3.35$, $Mdn = 4.0$). At the time of the study, the participants had an average of 1.57 clinical supervisors ($SD = 0.91$) and had been working with the supervisor

identified for the study for a median of 7 months ($M = 1.45$, $SD = 2.17$ years) with the expectation of working with this supervisor for another (Mdn) 6 months ($M = 1.09$, $SD = 1.62$ years). In regard to hours spent in clinical supervision, 91% of participants received at least one individual hour of supervision each week ($M = 1.72$, $SD = 2.87$, $Mdn = 1.0$). Supervisees identified their theoretical orientations as cognitive-behavioral (35.6%), eclectic (18.1%), humanistic/existential (12.4%), psychodynamic/psychoanalytic (11.0%), interpersonal (6.8%), systems (5.6%), behavior (4.5%), cognitive (2.5%), and other (3.5%). The mean number of clients discussed in supervision was 3.34 ($SD = 2.27$, $Mdn = 3.0$).

Supervisees also reported the demographics of their clinical supervisors. Of the supervisors, 60.7% were female. With regard to race, 85.5% were non-Hispanic White, 5.4% were African American, 4.8% were Hispanic/Latina, 1.7% were Asian/Pacific Islander, and 2.6% were other (Native American, Middle Eastern, etc.). The supervisor's reported theoretical orientations were cognitive-behavioral (29.1%), psychodynamic/psychoanalytic (14.3%), eclectic (13.7%), systems (6.9%), humanistic/existential (6.3%), interpersonal (6.0%), behavioral (2.9%), unknown (14.9%; supervisees reported not knowing their supervisor's theoretical orientation), or other (5.9%). The supervisor's degrees included master's (MA, MS, MEd, MSW; 38.9%), PhD (32.0%), PsyD (12.2%), other (14.5%), while 2.3% of the supervisees did not know their supervisor's degree. The supervisees believed that over half of the supervisors (64.9%) were trained in supervision, 6.6% were not trained, and 28.6% of the supervisees did not know whether their supervisors were trained. The supervisees reported that 6.0% of the supervisors were never licensed and 75.9% were licensed at the time of study; 1.1% selected either licensure pending or previously licensed, while 14.9% of supervisees did not know their supervisor's licensure status. Supervisor's reported field of study included Clinical Psychology (31.7%), Social Work (14.6%), Counseling Psychology (13.1%), Substance Abuse (5.4%), School Psychology (4.3%), Marriage and Family (2.9%), and 18.9% reported "other," which included Rehabilitation Counseling, Child/Adolescent Psychology, Mental Health Counseling, School Counseling, and Neuropsychology. Another 9.1% of supervisees reported not knowing their supervisor's field of study.

Variables. Following the definitions formulated in Study 1, inadequate clinical supervision and harmful clinical supervision were operationalized in two ways: self-identified and de facto. The taxonomy of supervision descriptors from Study 1 was included in a larger study whose purpose was to develop and test a new measure of the supervisory relationship. The items and supervision descriptors for the larger study were randomly ordered. For SIIS and

SIHS, we asked about the participant's experiences receiving clinical supervision as a supervisee. After reading the respective definitions, supervisees identified whether they had received inadequate or harmful supervision with their current supervisor as well as with other supervisors. Follow-up questions pertained to the context, severity, and impact of the experience.

Inadequate clinical supervision. Perceived occurrence of inadequate clinical supervision consisted of supervisees who reported receiving either self-identified, DFIS, or both with their current primary clinical supervisor.

SIIS: SIIS included those who responded *yes*—he or she received inadequate clinical supervision from their current or other supervisors (complete definitions of minimally adequate and inadequate supervision were provided). Participants selected from *yes*, *no*, or *maybe*, but only a *yes* response was classified as SIIS.

DFIS: Employing the taxonomy derived in Study 1, 16 inadequate supervision descriptors comprised DFIS. Participants rated “the extent to which each statement describes how you currently think or feel about your relationship with your clinical supervisor.” Supervision descriptors were rated on a 7-point fully anchored scale, where 1 was *not at all describes*, 4 was *sometimes describes*, and 7 was *totally describes*. Consistent with Study 1, we took a “supervisor conservative approach.” That is, to be counted as DFIS or DFHS, participants had to rate the supervision descriptors a five (*often describes*) or greater; or three (*occasionally describes*) or less for reverse scored descriptors. Three de facto criteria were assessed in the demographics questionnaire: used a supervision consent or contract, supervisor observed or reviewed recordings of supervisee therapy/counseling sessions, and number of hours per week received of one-on-one clinical supervision. Taking a supervisor conservative approach, using a rating of five or higher on any one supervision descriptor, receiving less than 1 hr of individual supervision per week, failing to use a supervision consent or contract, or not directly overseeing supervisee's sessions constituted DFIS.

Harmful clinical supervision. Perceived occurrence of harmful clinical supervision consisted of supervisees who reported receiving either self-identified, DFHS, or both with their current primary clinical supervisor.

SIHS: SIHS occurred if the person responded *yes* to receiving clinical supervision from their current or previous supervisors that was harmful after reading the definition. Participants selected from *yes*, *no*, or *maybe*; *yes* was classified as SIHS.

DFHS: The taxonomy of 21 harmful supervision descriptors from Study 1 defined DFHS. Using the same 7-point anchored rating scale described

above, the criteria was a rating of five or higher (three or lower for reversed scored behaviors). Given the severity of the supervisor behavior and its aversive affects, the criteria for four supervisor behaviors was a rating of two (*rarely describes*) or greater: supervisor physically threatened the supervisee; have a sexual relationship; use or have used drugs together; or are (or have been) sexually intimate. If one of these four behaviors occurred at all, it was DFHS.

Procedures. Participants were solicited primarily via email listservs with the stated purpose of developing and testing a new measure of the supervisory relationship. Harmful or inadequate supervisory experiences were not mentioned in the cover letter. The listservs included APA, ACA, NASW, and various graduate training programs. Participants were eligible for a random drawing of US\$150. The online research materials consisted of a cover letter and informed consent, a larger item pool inclusive of the taxonomy of 37 inadequate and harmful supervision descriptors, and a demographics questionnaire. The cover provided a link to a password-protected website (Psych-Data.com), and asked people to distribute the research participation request to others. We were unable to identify a response rate because the actual number of people reached by email was indeterminate.

Results

Inadequate clinical supervision

SIIS. Nearly one in four participants (24.6%; $n = 86$) identified currently receiving SIIS. In addition, 49.0% ($n = 179$) of the participants identified that they had received SIIS from another clinical supervisor. Taken together, 61.4% ($n = 265$) of the participants identified that they had received SIIS. These participants judged their current inadequate supervision as somewhat harmful to their clients ($M = 3.29$, $SD = 2.14$, where 1 = *not at all harmful*, 9 = *totally harmful*); 49% of the participants identified that they had received inadequate supervision from another clinical supervisor and that it was moderately harmful to their clients ($M = 4.35$, $SD = 2.39$).

DFIS. Overall, 90.1% ($n = 326$) of the participants met criteria on one or more of the 16 DFIS descriptors, constituting DFIS. Among the supervisees who were receiving DFIS, 45.5% ($n = 163$) endorsed multiple inadequate descriptors ($M = 2.28$, $Mdn = 1.0$, $SD = 2.24$ descriptors). Notably, 54.2% ($n = 197$) of the supervisees reported their current supervisor did not use either a consent or contract for clinical supervision; 39.7% ($n = 144$) reported their sessions were not observed, monitored, or reviewed; 12.5%

($n = 45$) of the supervisees indicated that the supervisor at most occasionally discussed the major difficulties they were facing with their clients; 8.8% ($n = 32$) reported receiving less than 1 hr of individual supervision per week; and 5.9% ($n = 22$) of the supervisees did not receive individual clinical supervision on a weekly basis. Two other DFIS descriptors pertained to the supervisee's cultural background (oblivious to cultural background; no interest in cultural background)—endorsed by 5.5% ($n = 20$) and 7.7% ($n = 28$) of the supervisees, respectively.

Aggregating SIIS and DFIS, 93.0% ($n = 337$) of the participants were receiving inadequate supervision in their current supervisory relationship. Of these, 1.2% ($n = 4$) reported receiving SIIS but did not meet criteria on any DFIS descriptors, 73.5% ($n = 248$) met criteria on at least one DFIS descriptor but did not identify as receiving SIIS, and 25.2% ($n = 85$) both reported SIIS and met criteria on at least one DFIS descriptor. Combining SIIS from the current or another supervisor with DFIS, 96.3% of the supervisees had received inadequate supervision. That is, 350 of the 363 participants were categorized as receiving inadequate supervision at some point during their career.

Harmful clinical supervision

SIHS. One in every eight participants (12.4%; $n = 43$) identified currently receiving SIHS. These supervisees rated their experiences as moderately harmful ($M = 4.56$, $Mdn = 4.0$, $SD = 2.63$, where 1 = *not at all harmed*, 4 = *moderately harmful*, and 9 = *totally harmed*). Over a fourth of the participants (27.4%, $n = 100$) reported receiving SIHS from another clinical supervisor. They judged these supervisory experiences as clearly harmful ($M = 5.87$, $Mdn = 6.0$, $SD = 2.45$). Collectively, 36.2% ($n = 132$) of the participants reported receiving SIHS from at least one of their supervisors.

Of the participants who reported receiving SIHS, 67.4% ($n = 245$) indicated that it was an ongoing situation versus a single incident or one supervision session, and a majority (62.8%; $n = 154$) did not report their harmful experience to agency staff. Of the supervisees who reported receiving SIHS from another supervisor, 86.0% ($n = 211$) indicated it was an ongoing situation. More than half (55.9%; $n = 203$) of these supervisees disclosed the other harmful supervision experiences to agency staff.

DFHS. Using the taxonomy of 21 harmful supervision descriptors, 28.1% ($n = 110$) of the supervisees were currently receiving DFHS. Among these supervisees, 39.2% ($n = 43$) endorsed more than one harmful descriptor ($M = 2.53$, $Mdn = 1.0$, $SD = 3.25$). Two of the most frequently endorsed harmful descriptors involved exploitation (safe from exploitation; avoids

exploitative dual roles), rated by 13.5% ($n = 49$) and 11.6% ($n = 42$) of the supervisees as *not at all* (reversed scored), respectively. One supervisee indicated that the current supervisor had often threatened him or her physically. Another supervisee indicated a current sexual relationship with his or her supervisor. None of the supervisees reported using drugs with their current supervisor.

Combining SIHS and DFHS, 35.3% ($n = 129$) of the supervisees were categorized as experiencing harmful supervision in their current supervisory relationship. Of these, 63.9% ($n = 83$) met criteria on at least one DFHS descriptor but did not report receiving SIHS. Aggregating SIHS by the current or another supervisor and DFHS, half of this sample of supervisees (50.9%; $n = 185$) were categorized as receiving harmful clinical supervision at some point during their career.

Discussion

The chief purpose of Study 2 was to obtain initial data on the perceived occurrence of inadequate and harmful clinical supervision, using the taxonomy of supervision descriptors derived from Study 1, as well as supervisees' self-identification. To interpret the results and findings from Study 2 within the context of the strengths and limitations of the study, these are discussed first.

Limitations. Perhaps the most significant limitation to the present study is that the data were from the perspective of the supervisee only. That is, responses to the DFIS and DFHS descriptors and self-identified items were ultimately based on supervisee self-report. Neither independent observational data nor supervisor data were obtained. While this limits the results to the subjective perspective of supervisees, it is also the case that about half of the supervision descriptors in the taxonomy (TIHCS) involve specific, observable supervisor behaviors (e.g., threatened me physically; no evaluative feedback). It seems unlikely that the supervisees' perspectives of whether these specific behaviors occurred would be inaccurate or invalid (see Figure 1).

Second, supervisees provided the data about themselves and their current supervisor. Variables such as the supervisee's and supervisor's gender, degree, and race, we presume, were credible. Other variables could be inaccurate or of questionable validity (e.g., supervisor theoretical orientation, and supervisor training in clinical supervision). Interestingly, for some supervisor variables, 9% to 29% of the supervisees responded "don't know" (i.e., 9.1% field of study/discipline, 14.9% licensure status, 15.1% theoretical orientation, and 28.6% supervisor trained in clinical supervision). Why did the supervisees

not know this information? Although many possible reasons exist, it seemed reasonable to infer tentatively that supervisors were not discussing this information with their supervisees. Recall that most supervisees (81%) indicated that their supervisors did not use a consent for supervision, which typically includes the supervisor's professional disclosure information (Bernard & Goodyear, 2014). Readers should recall that supervisor characteristics are based on supervisee report.

Finally, participants were solicited via listservs and email, so it was not possible to determine a response rate. We assumed that the sample was biased and not representative of the target population of mental health clinical supervisees; however, the nature of the bias was not evident. Recall that subjects for Study 2 were recruited as part of a larger study to develop a measure of the supervisory relationship. It should be noted that a small percentage of supervisees may represent countries other than the United States. As the practice of professional psychology may differ internationally, cross-cultural inferences may not be appropriate. It is possible that some participants chose to take part in the study based on either a particularly good or a bad supervisory experience, which may have biased the data. Without existing occurrence data for inadequate or harmful clinical supervision, we lacked comparative data to assess potential biases in the sample. The characteristics of the supervisees, however, were similar to the samples of other recently published articles (e.g., Amerikaner & Rose, 2012), for example, including master's and doctoral trainees with 35% of the sample being between the ages of 30 and 40 years old from clinical and counseling psychology programs. Nonetheless, the results should be generalized with caution beyond this sample, at least until the results are replicated.

Strengths. In an attempt to conduct a conceptually and methodologically rigorous study, we attended to threats to hypothesis, statistical conclusion, internal, and construct validities of the study (per Ellis & Ladany, 1997). Specifically, we used comparative data (e.g., self-identified as receiving harmful supervision or not), defined the constructs a priori, operationalized the primary constructs using self-identified and de facto inadequate and harmful supervision (i.e., included more objective data), and obtained a diverse sample of mental health provider supervisees drawn from multiple professional disciplines. The supervisees exhibited a broad range of supervised clinical experiences from pre-practicum to post-license (i.e., a supervisee developmental level proxy; Bernard & Goodyear, 2014), racial backgrounds (20% non-Caucasian), settings, and professional disciplines (albeit 57% psychology).

General Discussion

Major Findings

Rather than a separate discussion of the Study 2 findings, we integrated them with Study 1 into a general discussion of the major findings of this two-study project. Two findings were prominent: a revised framework for inadequate and harmful supervision, and initial occurrence data for inadequate and harmful clinical supervision.

Revised definitions and framework for inadequate and harmful clinical supervision. A major finding was an empirically grounded, revised conceptualization and framework for inadequate and harmful supervision. Specifically, the revised framework (a) offers construct definitions that differentiate self-identified from de facto inadequate and harmful supervision, and (b) differentiates harmful supervision from and subsumes it within inadequate supervision. While perhaps obvious to some readers, nevertheless, this conceptualization did not exist before in the published literature. Incorporated into the revised framework was the TIHCS—a taxonomy of 37 supervision descriptors as viable operational definitions of DFIS and DFHS. A coherent, theoretical, and empirically supported model of inadequate and harmful supervision has not existed to guide research and supervision practice. Indeed, very few instances of empirically grounded definitions and framework exist in the clinical supervision (see Bernard & Goodyear, 2014; Falender & Shafranske, 2004; cf. Milne, Aylott, Fitzpatrick, & Ellis, 2008) or broader literature.

The necessity of de facto and self-identification criteria to assess inadequate and harmful clinical supervision. The data suggested that assessing inadequate and harmful clinical supervision appeared to require the application of the de facto taxonomy in combination with self-identification. A clear discrepancy was observed between self-identified and de facto occurrence rates for inadequate and harmful clinical supervision. Although the occurrence of SIIS and SIHS was high in our opinion (24.6% for inadequate supervision and 12.4% for harmful supervision), an additional 68.3% and 21.8% of the supervisees, respectively, met criteria for DFIS or DFHS, but did not self-identify as currently receiving inadequate or harmful supervision. In other words, more than half of the supervisees may have unknowingly received inadequate and/or harmful supervision. Supervisees may have been unaware of what constitutes inadequate or harmful clinical supervision (Ellis, 2001). For example, given that many supervisors did not use a consent or contract for supervision, or did not monitor their sessions, supervisees may

be uninformed about their rights, what constitutes minimally adequate supervision, and the supervisor's responsibilities (Thomas, 2007). The lack of knowledge may compromise their ability to identify the extent to which the supervision they are receiving was inadequate or harmful. However, of the supervisees who reported receiving SIIS and SIHS, 4.7% and 53.5%, respectively, did not meet criteria for DFIS and DFHS. This could be the result of the conservative criteria used to classify a supervision descriptor as *de facto* inadequate or harmful—to meet criteria the descriptor had to be rated a five or greater (7 is *totally*).

Occurrence of inadequate and harmful clinical supervision. Foremost, the data suggested that the occurrence of inadequate and harmful clinical supervision were high (cf. Ellis, 2001). Fully 36% of supervisees in our sample were categorized as currently receiving harmful supervision, and over half were identified as receiving harmful clinical supervision at some point in their career. These percentages were at the upper end of the percentages previously found for similar constructs in smaller samples (33%-50%; for example, Barnett-Queen & Larrabee, 2000; Ladany et al., 1999; Larrabee & Miller, 1993; Nelson & Friedlander, 2001). The data also suggested that experiences with harmful supervision were not isolated events. Notably, the low mean ratings for the majority of descriptors suggest that each descriptor by itself did not occur frequently, which is positive. But the combination of descriptors shows a different view.

Inadequate supervision was even more common. A large majority (93%) of the supervisees in our sample were identified as currently receiving inadequate supervision, while 96.3% received inadequate supervision at some point in their careers. Notably, these findings were largely driven by supervisees' endorsements on two specific supervision descriptors: the supervisor's failure to observe or monitor supervisee sessions (39.7%), and failure to use a supervision consent or contract (54.2%).

Practical Implications and Future Directions

Reflecting on the findings and results from the two studies presented herein, several implications seemed warranted. The most striking findings of the present study were the observed occurrence rates of inadequate and harmful clinical supervision. The studies on the various aspects of "supervision that goes badly" in combination with the more rigorous data reported here are compelling. Perhaps, we should no longer question whether "bad" supervision occurs (Jacobsen & Tanggaard, 2009); instead, we need to focus on how to detect, solve, and prevent what appears to be a major problem in the field

(cf. Barlow, 2010; Castonguay et al., 2010; Dimidjian & Hollon, 2010; Lilienfeld, 2007; Mays & Frank, 1985).

Part of the solution may come from continuing to identify those variables that predict the occurrence of harmful and inadequate supervision. Once these factors are better understood, methods for preventing inadequate and harmful supervision can be augmented and refined. For instance, inadequate supervision may be less likely to occur if supervisors are mandated to receive training in supervision that includes supervision of supervision (Borders, 1989; Borders et al., 1991). This may also serve to protect supervisors from possible litigation (e.g., Guest & Dooley, 1999; Saccuzzo, 2002). The research presented herein only scratches the surface of such issues; the major factors leading to inadequate and harmful supervision remain unknown. Therefore, as with harmful therapy (e.g., Barlow, 2010), we challenge researchers to discover the variables that predict and explain inadequate and harmful clinical supervision.

A second part of the solution may come from better educating supervisees about the supervisory process. As hypothesized, many more participants met criteria for DFIS and DFHS than self-identified as such even though using our supervisor-protective approach. In other words, many supervisees may have unknowingly received inadequate and/or harmful supervision. If supervisees were more aware of their basic rights in supervision (e.g., the supervisor's responsibilities), they may be more cognizant of and able to identify when harmful or inadequate supervision is taking place. Supervisors routinely using a written consent and contract for supervision that includes a supervisee bill of rights (Bernard & Goodyear, 2014) could accomplish this. It may be equally important to train supervisees not only to identify inadequate and harmful supervision but also how to respond appropriately should they encounter either one. Training programs implementing a role induction for supervision prior to clinical practicum might achieve this (e.g., Bahrnick, Russell, & Salmi, 1991).

A few implications of the results reported here for clinical supervisors warrant a brief discussion. Indeed, evidence continues to suggest that most supervisors are either not formally trained or have received minimal training in clinical supervision (Bernard & Goodyear, 2014). Few supervisors have supervised experience conducting supervision (Bernard & Goodyear, 2014). As a result, many supervisors may lack knowledge about some critical features of supervision, and are likely not well informed about adequate, inadequate, and harmful supervision. Thus, it is not surprising that most of the supervisees in this study received inadequate supervision and many received harmful supervision. The 10 criteria for minimally adequate supervision described here in combination with implementing the APA guidelines for

clinical supervision (APA BEA Task Force on Supervision Guidelines, 2013) could be beneficial for educating supervisors and training programs. Incorporating these guidelines and criteria in formal supervision training and ensuring that supervisors are aware of the behaviors, actions, and inactions that constitute inadequate and harmful supervision could potentially reduce the occurrence of inadequate and harmful supervision.

Finally, it is noteworthy that problems arising from clinical supervision including inadequate supervision are the seventh most frequently reported reason for disciplinary actions by licensing boards (ASPPB, 2013). Perhaps this is one reason that ASPPB is currently drafting supervision guidelines to be incorporated into regulations for licensing psychologists (ASPPB, 2013). Hence, at a more systemic level, perhaps APA could implement more stringent criteria into the accreditation guidelines (cf. APA, 2007) to ensure that all accredited programs provide more than nominal exposure to the supervision literature (e.g., require a one semester supervision practicum in conjunction with a course on the theory, research, and practice of supervision). These changes would be consistent with the current ASPPB draft supervision guidelines. It seems ironic that few psychologists would permit someone to practice therapy with no training or with a 3- to 6-hr continuing education seminar. Yet as a profession, someone can engage in unsupervised supervisory practice without demonstrating competency in supervision knowledge, skills, and attitudes (Fouad et al., 2009). From this perspective, a seminar plus a semester of supervised supervision seems paltry.

Three additional issues merit a brief discussion. As noted previously, the taxonomy and de facto definitions for harmful supervision used in Study 2 took a supervisor conservative approach. That is, a rating of 5 (*often describes*) or greater was required to meet criteria for DFIS and DFHS. Using a supervisee-protective approach (i.e., *any* occurrence of a known harmful supervisor behavior; a rating of 2—*rarely describes*—or greater), the percentage of DFHS increased to 65.6% (vs. 28.1%) and the percentage of supervisees currently receiving harmful supervision increased to 66.4% (vs. 35.3%).

The question elicited here is how the field will balance a supervisee-protective stance and a supervisor-protective stance. The issue is complex. For example, issues of supervisee incompetence are likely a contributing or complicating factor (e.g., Falender, Collins, & Shafranske, 2009). Yet, we need to be cognizant to not victimize the victim; that is, assume that the fault lies with the supervisee versus the supervisor. At the same time, it seems prudent to be careful not to condemn or label a supervisor as inadequate or harmful prior to further investigation. Thus, the question of balancing supervisees' and supervisors' well-being is a political and empirical issue for leaders and researchers in clinical supervision to deliberate.

Another issue to address is that we derived the definition of minimally adequate supervision, and thus the construct definition of inadequate supervision and the taxonomy of DFIS descriptors, are based on existing literature and current supervision guidelines (e.g., ACA, 2005; APA, 2002, 2007; Bernard & Goodyear, 2014). Because the current literature and guidelines lack clear explication of adequate and inadequate supervision, the construct definition and taxonomy delineated here are open to debate. For example, many current guidelines for clinical supervision in psychology in the United States (e.g., APA, 2002, 2007; APPIC, 2009a, 2009b; ASPPB, 2003, 2009) do not explicitly call for supervisors to observe their supervisees' in-session behaviors (e.g., observe or monitor therapy sessions). Nevertheless, we included this criterion in the definition of minimally adequate supervision, the taxonomy of DFIS descriptors, and construct definition for inadequate supervision because many believe that it is vital for the supervisee's competence and professional development (e.g., AAMFT, 2003, 2007; Bernard & Goodyear, 2014; Falender & Shafranske, 2004; Huhra et al., 2008; Noelle, 2002). If therapists are to demonstrate competency in therapeutic skills, supervisors need to observe and provide feedback to the supervisee on what they are doing in therapy sessions (e.g., Bernard & Goodyear, 2014; Huhra et al., 2008). Monitoring supervisees' sessions is the preferred method to ensure that supervisees do not harm clients and deliver an adequate level of care. It is notable that nearly 40% of the supervisees in the Study 2 reported that their supervisor did not observe and provide feedback on the supervisee's in-session actions. Perhaps more troubling, the evidence suggested that clients may be harmed because of inadequate clinical supervision. Although the findings presented here are preliminary and require replication and substantiation, they, in combination with previous research, point to a problem in the profession.

As a final note, an important distinction is warranted—inadequate versus ineffective clinical supervision. As in the psychotherapy literature (Barlow, 2010; Dimidjian & Hollon, 2010), the adequacy of supervision is arguably related to, yet independent from the efficacy of supervision. They are distinct constructs. That is, it is conceivable to receive supervision that is adequate yet ineffective, whereas the inverse seems less plausible (effective and inadequate supervision). Nevertheless, these constructs deserve further investigation.

We encourage researchers to test further the framework for inadequate and harmful supervision to advance our understanding of the current status of clinical supervision. Continual assessment of the adequacy of the constructs and definitions offered herein, and the extent to which they appropriately capture inadequate and harmful clinical supervision experiences is also

important. Finally, we encourage readers to reflect on their own supervisory experiences, as a supervisee and if applicable, as a supervisor. In particular, supervisors are encouraged to examine critically their supervision practices with a keen eye to minimally adequate supervision and harmful supervision behaviors.

Conclusion

Worthington (1987) stated that “A good theory of lousy supervisor behaviors is missing” (p. 203). Over 25 years later, progress is slowly being made (e.g., Goodyear et al., 2005). Our hope is that the preliminary, cross-discipline definition of minimally adequate clinical supervision in combination with a revised framework and constructs that differentiate self-identified inadequate and harmful supervision from *de facto* inadequate and harmful clinical supervision presented here may augment our understanding. The TIHCS, a theoretically and empirically founded taxonomy for *de facto* inadequate and *de facto* harmful clinical supervision, may stimulate investigations of exceptional supervision. Finally, initial occurrence data for inadequate and harmful supervision illuminate problems in the current practice of clinical supervision that beckon to be acknowledged, investigated, and remedied.

Authors’ Note

Earlier versions of this manuscript were presented at the 4th and 6th International Interdisciplinary Conference on Clinical Supervision and the 117th and 118th American Psychological Association Conference.

Acknowledgment

We are grateful to Michael Zlatev, Sasha Lerner, and Amy Santell for their insightful comments on earlier drafts of this manuscript.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Note

1. We identify “*de facto* harmful supervisor descriptors” even though not all *de facto* harmful supervision (DFHS) behaviors assessed in the two studies presented here

may actually be perceived to be traumatic or result in significant distress or harm to the supervisee (e.g., Frazier et al., 2009). DFHS behaviors may or may not lead to harm due to a variety of factors, including the resilience of the person experiencing the harmful event.

References

- Allen, G. J., Szollos, S. J., & Williams, B. E. (1986). Doctoral students' comparative evaluations of best and worst psychotherapy supervision. *Professional Psychology: Research and Practice, 17*, 91-99. doi:10.1037/0735-7028.17.91
- American Association for Marriage and Family Therapy. (2003). *Approved supervisor designation standards and responsibilities handbook*. Washington, DC: Author.
- American Association for Marriage and Family Therapy. (2007). *Approved supervisor designation: Standards and responsibilities*. Washington, DC: Author.
- American Counseling Association. (2005). *ACA code of ethics*. Alexandria, VA: Author.
- American Psychological Association. (2002). *Ethical principles of psychologists and code of conduct*. Retrieved from <http://www.apa.org/ethics/code/index.aspx>
- American Psychological Association. (2007). *Guidelines and principles for accreditation of programs in professional psychology*. Washington, DC: Author.
- American Psychological Association Board of Educational Affairs Task Force on Supervision Guidelines. (2013). *Guidelines for competency-based clinical supervision in health service psychology education and training programs*. Unpublished draft document.
- Amerikaner, M., & Rose, T. (2012). Direct observation of psychology supervisees' clinical work: A snapshot of current practice. *Clinical Supervisor, 31*, 61-80. doi:10.1080/07325223.2012.671721.
- Anonymous. (1991). Sexual harassment: A female counseling student's experience. *Journal of Counseling & Development, 69*, 502-506. doi:10.1002/j.1556-6676.1991.tb02632.x
- Association for Counselor Education and Supervision. (1995). Ethical guidelines for counseling supervisors. *Counselor Education and Supervision, 34*, 270-276.
- Association of Psychology Postdoctoral and Internship Centers. (2009a). *APPIC membership criteria: Doctoral psychology internship programs*. Retrieved from <http://www.appic.org/About-APPIC/Joining-APPIC/Members/Internship-Membership-Criteria>
- Association of Psychology Postdoctoral and Internship Centers. (2009b). *APPIC membership criteria: Postdoctoral training programs*. Retrieved from <http://www.appic.org/About-APPIC/APPIC-Policies/Postdoc>
- Association of State and Provincial Psychology Boards. (2003). *Supervision guidelines*. Montgomery, AL: Author. Retrieved from http://www.asppb.net/resource/resmgr/guidelines/current_asppb_supervision_gu.pdf
- Association of State and Provincial Psychology Boards. (2009). *Guidelines on practicum experience for licensure*. Montgomery, AL: Author. Retrieved from http://www.appic.org/Portals/0/docs/ASPPB%20Memo_Prac%20Guidelines%201%2031%2009.doc

- Association of State and Provincial Psychology Boards. (2013, July). *Definitions of competence, Day 2*. 5th International Congress on Licensure, Certification and Credentialing in Psychology, Stockholm, Sweden.
- Australian Capital Territory Psychologists Board. (2009). *Standard statement 11: Guidelines for conditionally registered psychologists*. Brandon, Queensland, Australia: Author. Retrieved from <http://health.act.gov.au/c/health?a=da&did=10033491>
- Australian Psychological Society. (2003). *APS ethical guidelines: Guidelines on supervision*. Melbourne, Victoria, Australia: Author. Retrieved from http://www.psychology4change.com/forms/APS_ethical_guidelines_on_supervision.pdf
- Bahrnick, A. S., Russell, R. K., & Salmi, S. W. (1991). The effects of role induction on trainees' perceptions of supervision. *Journal of Counseling & Development, 69*, 434-438. doi:10.1002/j.1556-6676.1991.tb01540.x
- Barlow, D. H. (2010). Negative effects from psychological trauma. *American Psychologist, 65*, 13-20. doi:10.1037/a0017299
- Barnett-Queen, T., & Larrabee, M. J. (2000). Sexually oriented relationships between educators and students in mental-health education programs. *Journal of Mental Health Counseling, 22*, 68-84.
- Bartell, P. A., & Rubin, L. J. (1990). Dangerous liaisons: Sexual intimacies in supervision. *Professional Psychology: Research and Practice, 21*, 442-450. doi:10.1037/0735-7028.21.6.442
- Bernard, J. M., & Goodyear, R. K. (2014). *The fundamentals of clinical supervision* (5th ed.). New York, NY: Pearson.
- Borders, L. D. (1989). Credentialing supervisors: A commitment to professionalism. *Illinois Association for Counseling and Development Quarterly, 112*, 33-42.
- Borders, L. D., Bernard, J. M., Dye, H. A., Fond, M. L., Henderson, P., & Nance, D. W. (1991). Curriculum guide for training counseling supervisors: Rationale, development, and implementation. *Counselor Education and Supervision, 31*, 58-80. doi:10.1002/j.1556-6978.1991.tb00371.x
- British Association for Counseling. (1988). *Code of ethics and practice for the supervision of counselors*. Rugby, UK: Author.
- British Psychological Society. (2003). *Policy guidelines on supervision in the practice of clinical psychology*. Leicester, UK: Author.
- British Psychological Society. (2006). *Continued supervision*. Leicester, UK: Author.
- Burkard, A. W., Knox, S., Hess, S. A., & Schultz, J. (2009). Lesbian, gay, and bisexual supervisees' experiences of LGB-affirmative and nonaffirmative supervision. *Journal of Counseling Psychology, 56*, 176-188. doi:10.1037/0022-0167.56.1.176
- Canadian Psychological Association. (2009). *Ethical guidelines for supervision in psychology: Teaching, research, practice, and administration*. Ottawa, Ontario, Canada: Author.
- Castonguay, L. G., Boswell, J. F., Constantino, M. J., Goldfried, M. R., & Hill, C. E. (2010). Training implications of harmful effects of psychological treatments. *American Psychologist, 65*, 34-49. doi:10.1037/a0017330

- Celenza, A. (2007). *Sexual boundary violations: Therapeutic, supervisory, and academic contexts*. Lanham, MD: Jason Aronson.
- Center for Credentialing and Education. (2009). *Approved clinical supervisor code of ethics*. Greensboro, NC: Author. Retrieved from <http://www.cce-global.org/credentials-offered/acs>
- Cobia, D. C., & Boes, S. R. (2000). Professional disclosure statements and formal plans for supervision: Two strategies for minimizing the risk of ethical conflicts in postmaster's supervision. *Journal of Counseling & Development, 78*, 293-296. doi:10.1002/j.1556-6676.2000.tb01910.x
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Council for the Accreditation of Counseling and Related Educational Programs. (2009). *Accreditation procedures manual and standards*. Alexandria, VA: Author.
- Crespi, T. D., & Lopez, P. G. (1998). Practicum and internship supervision in the schools: Standards and considerations for school psychology supervisors. *Clinical Supervisor, 17*, 113-126. doi:10.1300/J001v17n02_07
- Dimidjian, S., & Hollon, S. D. (2010). How would we know if psychotherapy were harmful? *American Psychologist, 65*, 21-33. doi:10.1037/a001729
- Dressel, J. L., Consoli, A. J., Kim, B. S. K., & Atkinson, D. R. (2007). Successful and unsuccessful multicultural supervisory behaviors: A Delphi poll. *Journal of Multicultural Counseling and Development, 35*, 51-64. doi:10.1002/j.2161-1912.2007.tb00049.x
- Dye, H. A., & Borders, L. D. (1990). Counseling supervisors: Standards for preparation and practice. *Journal of Counseling & Development, 69*, 27-32. doi:10.1002/j.1556-6676.1990.tb01449.x
- Ellis, M. V. (2001). Harmful supervision, a cause for alarm: Commentary on Nelson & Friedlander (2001) and Gray et al. (2001). *Journal of Counseling Psychology, 48*, 401-406. doi:10.1037/0022-0167.48.4.401
- Ellis, M. V., & Ladany, N. (1997). Inferences concerning supervisees and clients in clinical supervision: An integrative review. In C. E. Watkins (Ed.), *Handbook of psychotherapy supervision* (pp. 447-507). New York, NY: Wiley.
- Falender, C. A., Collins, C. J., & Shafranske, E. P. (2009). "Impairment" and performance issues in clinical supervision: After the 2008 ADA Amendments Act. *Training and Education in Professional Psychology, 3*, 240-249. doi:10.1037/a0017153
- Falender, C. A., Cornish, J. A. E., Goodyear, R., Hatcher, R., Kaslow, N. J., Leventhal, G., . . . Grus, C. (2004). Defining competencies in psychology supervision: A consensus statement. *Journal of Clinical Psychology, 60*, 771-785. doi:10.1002/jclp.20013
- Falender, C. A., & Shafranske, E. P. (2004). *Clinical supervision: A competency-based approach*. Washington, DC: American Psychological Association.
- Fouad, A., Grus, C. L., Hatcher, R. L., Kaslow, N. J., Hutchings, P. S., Madson, M. B., . . . Crossman, R. E. (2009). Competency benchmarks: A model for understanding and measuring competence in professional psychology across training levels.

- Training and Education in Professional Psychology*, 3, S5-S26. doi:10.1037/a0015832
- Frazier, P., Anders, S., Perera, S., Tomich, P., Tennen, H., Park, C., & Tashiro, T. (2009). Traumatic events among undergraduate students: Prevalence and associated symptoms. *Journal of Counseling Psychology*, 56, 450-460. doi:10.1037/a0016412
- Giddings, M. M., Cleveland, P. H., & Smith, C. H. (2007). Responding to inadequate supervision: A model promoting integration for post-MSW practitioners. *Clinical Supervisor*, 25, 105-126. doi:10.1300/J001v25n01_08
- Goodyear, R. K., Bunch, K., & Claiborn, C. D. (2005). Current supervision scholarship in psychology: A five-year review. *Clinical Supervisor*, 24, 137-147. doi:10.1300/J001v24n01_07
- Gottlieb, M. C., Robinson, K., & Younggren, J. N. (2007). Multiple relations in supervision: Guidance for administrators, supervisors, and students. *Professional Psychology: Research and Practice*, 38, 241-247. doi:10.1037/0735-7028.38.3.241
- Grant, J., Crawford, S., & Schofield, M. J. (2012). Managing difficulties in supervision: Supervisors' perspectives. *Journal of Counseling Psychology*, 59, 528-541. doi:10.1037/a0030000
- Gray, L. A., Ladany, N., & Ancis, J. R. (2001). Psychotherapy trainees' experience of counterproductive events in supervision. *Journal of Counseling Psychology*, 48, 371-383. doi:10.1037/0022-0167.48.4.371
- Greer, J. A. (2002). Where to turn for help: Responses to inadequate clinical supervision. *Clinical Supervisor*, 21, 135-143. doi:10.1300/J001v21n01_11
- Guest, C. L. Jr., & Dooley, K. (1999). Supervisor malpractice: Liability to the supervisee in clinical supervision. *Counselor Education and Supervision*, 38, 269-279. doi:10.1002/j.1556-6978.1999.tb00577.x
- Haase, R. F., Ellis, M. V., & Ladany, N. (1989). Multiple criteria for evaluating the magnitude of experimental effects. *Journal of Counseling Psychology*, 36, 511-516. doi:10.1037/0022-0167.36.4.511
- Hall, J. E. (1988). Dual relationships in supervision. *Register Report*, 14(4), 3-4.
- Holland, B. S., & Copenhaver, M. D. (1988). Improved Bonferroni-type multiple testing procedures. *Psychological Bulletin*, 104, 145-149. doi:10.1037/0033-2909.104.1.145
- Huhra, R. L., Yamokoski-Maynhart, C. A., & Prieto, L. R. (2008). Reviewing videotape in supervision: A developmental approach. *Journal of Counseling & Development*, 86, 412-418. doi:10.1002/j.1556-6678.2008.tb00529.x
- Hutt, C. H., Scott, J., & King, M. (1983). A phenomenological study of supervisees' positive and negative experiences in supervision. *Psychotherapy: Theory, Research & Practice*, 20, 118-123. doi:10.1037/h0088471
- Jacobsen, C. H., & Tanggaard, L. (2009). Beginning therapists' experiences of what constitutes good and bad psychotherapy supervision: With a special focus on individual differences. *Nordic Psychology*, 61, 59-84. doi:10.1027/1901-2276.61.3.59

- Kaduvettoor, A., O'Shaughnessy, T., Mori, Y., Beverly, C., III, & Ladany, N. (2009). Helpful and hindering multicultural events in group supervision: Climate and multicultural competence. *The Counseling Psychologist, 37*, 786-820. doi:10.1177/0011000009333984
- Koenig, T. L., & Spano, R. N. (2003). Sex, supervision, and boundary violations: Pressing challenges and possible solutions. *Clinical Supervisor, 22*, 1-19. doi:10.1300/J001v21n01_01
- Ladany, N., Friedlander, M. L., & Nelson, M. L. (2005). *Critical events in psychotherapy supervision: An interpersonal approach*. Washington, DC: American Psychological Association.
- Ladany, N., Lehrman-Waterman, D., Molinaro, M., & Wolgast, B. (1999). Psychotherapy supervisor ethical practices: Adherence to guidelines, the supervisory working alliance, and supervisee satisfaction. *The Counseling Psychologist, 27*, 443-475. doi:10.1177/0011000099273008
- Ladany, N., Mori, Y., & Mehr, K. E. (2013). Effective and ineffective supervision. *The Counseling Psychologist, 41*, 28-47. doi:10.1177/0011000012442648
- Lamb, D. H., Catanzaro, S. J., & Moorman, A. S. (2003). Psychologists reflect on their sexual relationships with clients, supervisees, and students: Occurrence, impact, rationales and collegial intervention. *Professional Psychology: Research and Practice, 34*, 102-107. doi:10.1037/0735-7028.34.1.102
- Larrabee, M. J., & Miller, G. M. (1993). An examination of sexual intimacy in supervision. *Clinical Supervisor, 11*, 103-126. doi:10.1300/J001v11n02_09
- Lilienfeld, S. O. (2007). Psychological treatments that cause harm. *Perspectives on Psychological Science, 2*, 53-70. doi:10.1111/j.1745-6919.2007.00029.x
- Mays, D. T., & Frank, C. M. (1985). *Negative outcome in psychotherapy and what to do about it*. New York, NY: Springer.
- Milne, D., Aylott, H., Fitzpatrick, H., & Ellis, M. V. (2008). How does clinical supervision work? Using a "best evidence synthesis" approach to construct a basic model of supervision. *Clinical Supervisor, 27*, 170-190.
- Moskowitz, S. A., & Rupert, P. A. (1983). Conflict resolution within the supervisory relationship. *Professional Psychology: Research and Practice, 14*, 632-641. doi:10.1037/0735-7028.14.5.632
- National Association of Alcoholism and Drug Abuse Counselors. (2008). *NAADAC guide to certification*. Alexandria, VA: Author.
- National Association of Alcoholism and Drug Abuse Counselors. (2011). *NAADAC code of ethics*. Retrieved from <http://www.naadac.org/membership/code-of-ethics>
- National Association of Social Workers. (2008). *Code of ethics - revised*. Washington, DC: NASW Press. Retrieved from <http://www.socialworkers.org/pubs/code/code.asp>
- National Council on the Practice of Clinical Social Work. (2003). *Guidelines for clinical social work supervision*. Washington, DC: NASW Press.
- Nelson, M. L., Barnes, K. L., Evans, A. L., & Triggiano, P. J. (2008). Working with conflict in clinical supervision: Wise supervisor's perspectives. *Journal of Counseling Psychology, 55*, 172-184. doi:10.1037/0022-0167.55.2.172

- Nelson, M. L., & Friedlander, M. L. (2001). A close look at conflictual supervisory relationships: The trainee's perspective. *Journal of Counseling Psychology, 48*, 384-395. doi:10.1037/0022-0167.48.4.384
- New Zealand Psychologists Board. (2009a). *Core competencies for the practice of psychology in New Zealand*. Wellington, New Zealand: Author. Available from www.psychologistsboard.org.nz
- New Zealand Psychologists Board. (2009b). *Draft guidelines on supervision*. Wellington, New Zealand: Author. Available from www.psychologistsboard.org.nz
- New Zealand Psychologists Board. (2010). *The continuing competence programme for psychologists practising in Aotearoa New Zealand: A guide for participants*. Wellington, New Zealand: Author. Available from www.psychologistsboard.org.nz
- Noelle, M. (2002). Self-report in supervision: Positive and negative slants. *Clinical Supervisor, 21*, 125-134. doi:10.1300/J001v21n01_10
- Ramos-Sánchez, L., Esnil, E., Goodwin, A., Riggs, S., Touster, L., Wright, L.K., . . . Rodolfa, E. (2002). Negative supervisory events: Effects on supervision and supervisory alliance. *Professional Psychology: Research and Practice, 33*, 197-202. doi:10.1037/0735-7028.33.2.197
- Saccuzzo, D. P. (2002; November). Liability for the failure to supervise adequately: Let the master beware, Part 1. *The Psychologist's Legal Update, 13*, 1-14.
- Saccuzzo, D. P. (2003, May). Liability for the failure to supervise adequately: Let the master beware, Part 2: Ethical basis for standard of care in supervision. *The Psychologist's Legal Update, 13*, 1-20.
- Substance Abuse and Mental Health Services Administration. (2007). *Competencies for substance abuse treatment clinical supervisors: TAP 21-A*. Rockville, MD: Author.
- Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A. M. B., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: Implications for clinical practice. *American Psychologist, 62*, 271-286. doi:10.1037/0003-066X.62.4.271
- Sutter, E., McPherson, R. H., & Geeseman, R. (2002). Contracting for supervision. *Professional Psychology: Research and Practice, 33*, 495-498. doi:10.1037/0735-7028.33.5.495
- Thomas, J. T. (2007). A consent through contracting for supervision: Minimizing risks, enhancing benefits. *Professional Psychology: Research and Practice, 38*, 221-231. doi:10.1037/0735-7028.38.3.221
- Wampold, B. E., Davis, B., & Good, R. H., III. (1990). Hypothesis validity of clinical research. *Journal of Consulting and Clinical Psychology, 58*, 360-367. doi:10.1037/0022-006X.58.3.360
- Watkins, C. E., Jr. (1997). The ineffective psychotherapy supervisor: Some reflections about bad behaviors, poor process, and offensive outcomes. *Clinical Supervisor, 16*, 163-180. doi:10.1300/J001V16n01_09

Worthington, E. L., Jr. (1987). Changes in supervision as counselors and supervisors gain experience: A review. *Professional Psychology: Research and Practice*, 18, 189-208. doi:10.1037/0735-7028.18.3.189

Author Biographies

Michael V. Ellis, PhD, is a professor and director of the Division of Counseling Psychology in the Department of Educational and Counseling Psychology at the University at Albany. He is a fellow of APA (Division 17), an approved clinical supervisor, and recipient of the Supervision and Training Section's 2010 *Outstanding Professional Achievement Award*. His teaching and research interests include multicultural clinical supervision, supervisor training, clinical judgment, and research methodology, psychometrics, and statistics. He also maintains a private practice.

Lauren Berger, MS, is a doctoral candidate in counseling psychology at the University at Albany. She graduated with a bachelor's degree from Siena College and received a master's degree in mental health counseling from the University at Albany. Her research interests include clinical supervision, work/family balance, and social justice commitment.

Alexa E. Hanus is a doctoral candidate in counseling psychology at the University at Albany. She graduated with a bachelor's degree from Minnesota State University and received a master's degree in counseling psychology from Towson University. Her research interests include clinical supervision and training, as well as racial identity development and racism prevention.

Erin E. Ayala is a doctoral student in counseling psychology at the University of Albany. She received a bachelor's degree in psychology and women's studies from St. Catherine University. Her research interests include prevention and health promotion, multicultural counseling and training, research methodology, and the identity development of doctoral trainees.

Brett A. Swords is a doctoral candidate in counseling psychology at the University at Albany. He received a BS in psychology from York College of Pennsylvania. His research interests include clinical supervision and trainee anxiety. In particular, he is interested in investigating the detrimental effects of harmful or "bad" supervision.

Michael Siembor, PhD, is a staff psychologist and the alcohol and other drugs prevention coordinator at Hobart and William Smith Colleges' Center for Counseling and Student Wellness. He completed his undergraduate degree at the University of Notre Dame and received his PhD in counseling psychology from the University at Albany. His clinical and research interests include clinical supervision, alcohol and other drugs prevention, and higher education assessment.