Video Feedback in Education and Training: Putting Learning in the Picture

**Ruben Georges Fukkink · Noortje Trienekens ·
Lisa J. C. Kramer**

Abstract This meta-analysis demonstrates that the video feedback method has a statistically significant effect on the interaction skills of professionals in a range of contact professions. The aggregate effect, calculated on the basis of 217 experimental comparisons from 33 experimental studies involving a total of 1,058 people, was 0.40 standard deviation (SE=0.07). The effects of training were greater for programs working with a standard observation form of target skills that were central to the program. Results were more positive for outcome measures that measured positive skills rather than negative ones. In addition, molar outcome measures, which were obtained by means of an assessment scale, showed larger effects than micromeasures, which were scored using event sampling. Finally, recommendations are made for video feedback design and for future research.

Not long after Charlie Ginsburg invented the video recorder in 1951 (Ginsburg 1956), it was used for instructional purposes and professional development. The "chance marriage of the videotape recorder and a program of microteaching" (Hargie *et al*. 1983: 153) at Stanford University in 1963 led to the very first educational application of video feedback (Allen 1966, 1967), an example that was soon followed by Ivey and colleagues at Colorado State University in 1968 (Ivey and Authier 1978). At both these universities, course participants were filmed as they interacted with clients, after which a trainer discussed the tapes with the participants. Thus the new medium of video was exploited not only to make in-depth studies of the behavior of professionals, but also to modify that behavior. Following this pioneering work, various researchers and trainers have explored the new educational opportunities offered by the medium (see also Borg *et al*. 1970; Cooper and Allen 1970; McKnight 1980; MacLeod 1987 for the historical context of video feedback). Video feedback, which Kurtz *et al*. (2005) describe as "the gold standard of communication

R. G. Fukkink (✉) · N. Trienekens · L. J. C. Kramer
Graduate School of Child Development and Education, University of Amsterdam,
Nieuwe Prinsengracht 130, 1018 VZ Amsterdam, The Netherlands
e-mail: R.G.Fukkink@uva.nl

teaching" (p. 83), is used in various higher professional education and training courses to improve the communication skills of a broad group of "interpersonal professionals" (Hargie et al. 1983), including teachers, psychologists, social workers, doctors and nurses, for whom effective communication plays a vital role in their work (see Baker et al. 1990; Beckman and Frankel 1994; Gask 1992; Hill and Lent 2006; Hulsman et al. 1999; Huhra et al. 2008; Perlberg 1983; Quigley and Nyquist 1992; Romans et al. 1995; Schoonover et al. 1983; Sherin 2004; Silverman et al. 2005).

Feedback in general plays a vital role in skills teaching (Kluger and DeNisi 1996; Shute 2008). What makes the feedback in the video feedback method unique is that it allows course participants to look at themselves "from a distance" and with space for reflection, thereby giving them a realistic picture of their own skills, or self-image (Fuller and Manning 1973; Hargie et al. 1983; Hosford 1980). Through repeated playing of the videotape, this method also allows a detailed analysis of a person's behavior. Different studies within the microtraining paradigm (see Allen 1967; Borg et al. 1970; Ivey and Authier 1978) have focused on specific microbehaviors. Microbehaviors are concrete behaviors with a relatively brief duration, which are usually studied with behavorial counts (e.g., head nodding, hand gestures, the number of open questions that trainees ask or how often they look at the other person). Other training studies have examined broader, more holistic skills, which are related to behavorial dimensions like sensitivity, warmth or kindness and which are usually measured with rating scales. In some studies, instruction aims at certain core skills or essential qualities for effective interaction (e.g., empathy), for example. Other studies with a focus on more holistic skills have trained students in the application of a communication model that comprises different stages (e.g., initiating the session, gathering information and closing the session). More added value of the video medium lies in the comprehensive application possibilities it offers for focusing on verbal aspects (i.e., the content of what is being said), paralingual aspects (i.e., intonation, speaking pace, and volume) and non-verbal aspects (e.g., body posture, eye contact, use of gestures; see Hargie and Dickson 2004). Attention to each of these aspects is important because they all play a significant role in the various communication skills required in professional practice, including receptive skills (e.g., asking open questions, looking at the other person, use of silences), informative skills (e.g., explaining things in a comprehensible way, speaking calmly) and relational skills (e.g., asking about the other's experiences and displaying empathy) (see Duffy et al. 2004; Hulsman et al. 1999).

Variations of Video Feedback

Fifty years of innovation and empirical research have produced different variations of video feedback (hereafter abbreviated to VF). Studies from the early days of the method often used video images in "unstructured video replay" (Dowrick 1983). The strength of this "video self-confrontation" approach was believed to lie in viewing oneself, with other instructional elements playing only a minor role. The emphasis in today's training programs is no longer on confronting participants with images of themselves. Instead, "positive self-modeling" approaches focus solely or primarily on successful interactions by the participant in order to reinforce the desired target behavior and to give the participant a positive self-image (see Dowrick 1983; Hosford 1980). Tying in with Bandura's social learning theory (Bandura 1969, 1978, 1997), Hosford (1980) and Dowrick (1983, 1999) have emphasized that positive empowerment is pedagogically preferable for VF training programs because this kind of feedback boosts self-efficacy and leads to the behavior being displayed more

frequently. Separate from the specific VF context, educational studies into the effects of feedback in general have also shown that feedback that could erode someone's sense of self-worth is not very effective (Hattie and Timperley 2007; Kluger and DeNisi 1996). Equally important, feedback should be specific, because the specificity of the feedback helps trainees to discover the key elements of their behavior and to evaluate their performance.

Most of the current VF interventions go beyond the fairly isolated use of videotape, which according to Hosford and Mills (1983) is less effective than VF in combination with additional instructions. The emphasis in VF studies has thus shifted quite quickly away from the autonomous use of videotape for self-confrontation to more comprehensive interventions in which videotape, although an essential part of the intervention, is always supplemented by other forms of instruction. In a number of studies, for example, VF is accompanied by an explanation of effective professional behavior, by modeling or by viewing instructional videos (discrimination training), by practicing the skills in role plays with fellow course participants or in real practical situations and/or by guidance from a supervisor (see Hargie *et al*. 1983). The role of targeted feedback, sometimes referred to as "cueing" or "behavior coding", is given particular emphasis when viewing videotapes. Some training courses assign a key role to observation lists that provide an overview of the specific target behaviors as they can guide participants during the actual VF sessions (Borg *et al*. 1970; Fuller and Manning 1973; Hargie *et al*. 1983; Huhra *et al*. 2008). Borg and his colleagues (Borg 1972; Borg *et al*. 1970) were presumably the first researchers who combined instruction, modeling, practice, and video feedback with a structured evaluation form. Their "minicourses", which were at that time "probably the most comprehensive development of microteaching for in-service training" according to Cooper and Allen (1970: 8), are therefore characterized by a coherent instructional sequence. This is important because instruction, practice, and feedback are intrinsically linked in this format. The instruction operationally defines a specific skill and shows participants precisely what the target behavior is in a concrete, practical situation. The detailed and specific feedback on the target behavior subsequently helps participants to evaluate their performance in a structured manner.

The various VF interventions also differ in the way they apply the technical possibilities offered by video, which incidentally are not explicitly related to principles of learning theory. In this context, Hosford and Mills (1983) cite the accelerated and slow motion replay of videotapes, the use of the pause button (freeze frame), showing images without sound (picture-only feedback), or conversely, playing the sound without pictures (sound-only feedback). More sophisticated applications are the split-screen technique showing the professional and the client at the same time so the viewer can see in a single picture the effect that a person's behavior has on the other, and serial viewing, in which the trainer edits recordings of sessions held at different times into a single video to show a person's development over time (see also Dowrick 1991, 1983).

Effects of Video Feedback on Professionals

The very first review study into the effects of VF on the interaction skills of professionals dates back to 1973. In that review, Fuller and Manning came to the following conclusion: "Practitioners have good reasons for their optimism about self-confrontation, and researchers have good grounds for skepticism" (p. 511). A similar paradox appears in the final sentence of their article: "Self-confrontation now seems to us more promising than we

had dared to hope and more dangerous than we knew to fear" (p. 512). While stylistically satisfying, these findings from the early days of the VF method left unanswered the question of whether or not the method was effective. Following the first review, a range of experimental studies further explored the effectiveness of VF for professionals in various settings. The VF method continued to develop at that time thanks to innovative methods for improving the communication skills of participants with the aid of videotape. Later publications sketch a less equivocal and more positive picture of the effects of VF on the communication skills of professionals than Fuller and Manning (see for example Baker and Daniels 1989; Ford 1979; Hargie *et al.* 1983; Baker *et al.* 1990; Hargie 2006), although it should be noted that these studies relate not just to the VF method, but also to various other approaches to the teaching of communication skills. A systematic description of effect studies into VF is still lacking in this domain. Also lacking is a precise quantification of the effect of VF, as only narrative reviews have been conducted thus far. Related to this, another gap in the research literature is that we do not know enough about which pedagogical and methodological characteristics of the studies are associated with outcomes of VF. This is critical because reviews of the VF method in various domains have emphasized the need for future research to identify the key variables that correlate with the effectiveness of VF (see Fukkink 2008; Hargie *et al.* 1983; Hill and Lent 2006; Hosford 1980; Hung and Rosenthal 1981; MacLeod 1987).

Research Questions and Hypotheses

This study looks at two main questions.

1.	What is the effect of VF interventions on the interaction skills of professionals?
2.	Which methodological and pedagogical characteristics correlate systematically with the results of experimental studies into VF?

We investigate these questions in a meta-analysis of the results of experimental VF studies published between 1973 and 2009, the period following the publication of Fuller and Manning's classic review study. To answer the first question, in addition to the overall effect of VF, we identify the learning effects for verbal, non-verbal, and paralingual behavior (Hargie and Dickson 2004) and for receptive, informative, and relational skills (Huhra *et al.* 2008).

Three hypotheses were derived from the literature that make predictions about the effectiveness of VF. Firstly, various authors have pointed out that the learning outcomes are greater if the training program supplements video recordings with additional instruction (see for example, Hargie *et al.* 1983; Hosford and Mills 1983). From an educational perspective as well, researchers stress that feedback is more effective in general if it relates to instruction (Hattie and Timperley 2007). In line with the literature, the first hypothesis is as follows:

H1:	VF interventions combined with additional instruction are more effective than VF interventions with no additional instruction.
	We test this hypothesis by examining whether there is a difference between VF interventions with no additional instruction and VF interventions with instruction (hypothesis 1a). We also investigate whether the effectiveness of training increases with the number of additional instructional components, such as an oral or written explanation of the target behavior, modeling and discrimination training or practice (hypothesis 1b).

Various publications looking at the design of VF training for professionals emphasize that the feedback must be clearly focused and must relate to the specific instructed skills that the participants are expected to master. Focused feedback is assumed to be more effective because it is specific and is systematically linked to instruction (Borg *et al*. 1970; Brinko 1993; Fuller and Manning 1973; Hargie *et al*. 1983; Huhra *et al*. 2008; Star 1979; Thelen and Lasoski 1980). From this perspective, we test the hypothesis that approaches involving a detailed observation form listing specific target behavior are more effective than approaches with no such form.

H2: VF that incorporates a structured observation form is more effective than VF with no such form.

Lastly, we test whether learning outcomes correlate with the course participants' level of development.

H3: The experimental effects of VF are smaller for trainees with more experience compared with less-experienced trainees.

A number of publications highlights the fact that VF training effects are greatest in the early stage of training, are reduced in the subsequent stage of professional training and are even more modest for refresher training of qualified professionals (Huhra *et al*. 2008; Kruijver *et al*. 2000). Baker *et al*. (1990) suggested in their review study that undergraduate students may show more progress than graduate students. However, another meta-analysis by Hill and Lent (2006) did not find significant differences between these two groups. More research is therefore needed into this area.

Methods

Literature search

In order to find experimental studies into the effects of VF on the interaction skills of professionals, we searched the electronic databases of the Social Sciences Citation Index, ERIC and PsychInfo using a broad search profile that combined different search terms (video*; self-model*, self-confrontation, self-observation*, playback, feedback, self-confrontation, videotape-recorded playback; interaction*, communication*, skill*, performance*, competence*). We then used the so-called snowball method to search the relevant studies for references to other studies. We also searched citation links in the SSCI using the forward method in order to trace later studies.

To qualify for inclusion in our meta-analysis, interventions had to make use of videotape recordings featuring the participants themselves, the hallmark of VF. Studies reporting on the effects of video *instruction* (self-instructional videotape, see for example, Shernoff and Kratochwill 2007) in which participants did not see themselves on video or studies in which trainees learn observation skills with the aid of videotapes (see Star and Strickland 2008), were not included. Evaluations of "video clubs" (see for example, Sherin and Van Es 2009; Tan and Towndrow 2009) were also omitted because we were unable to establish unequivocally the extent to which participants viewed recordings of themselves in these studies. For two evaluations of broad programs comprising several components (Fantuzzo *et al*. 1996, 1997), the VF component was judged too small to be able to attribute the effects unequivocally to VF. One recall method study was not included because the videotape was used primarily as a research tool rather than for instructional purposes (Berthelsen and Brownlee 2007).

Another criterion for inclusion relating to the outcome measure of this study was that the intervention effects on the professionals' interaction skills had to have been tested by means of an external evaluation of behavior involving an observation instrument. For this reason, studies involving self-evaluation (Zimmerman *et al.* 2003) or a cataloguing of client perceptions of professionals (Sliwa *et al.* 2002) were not included. Studies into the effects of VF on knowledge, attitudes or skill identification—instead of the independent application of a skill—were excluded from the analysis for the same reason (Cassata *et al.* 1976; Engel *et al.* 1976; Hays 1976; Hehr 1981; Kpanja 2001; Martin-Reynolds 1980). Studies that identified the effects of VF on clients rather than professionals (see White and Poppen 1979) were also not included. Further, the studies had to describe quantitative research and to report on the statistical data required to calculate an effect measure. A number of studies did not qualify for inclusion for this reason (Ajayi-Dopemu and Talabi 1986; Brown and Kameen 1975; Cassata and Clements 1978; Cassata *et al.* 1977; Fyffe and Oei 1979; Gask 1998; Gask *et al.* 1991, 1987; Hosford and Johnson 1983; Hougham 1992; Hulsman *et al.* 2009; Kern 1980; Levinson and Roter 1993; Marita *et al.* 1999; Napper-Owen and Phillips 1995; Schmidt and Messner 1977; Scott *et al.* 1983; Sollie and Scott 1983; Speidel and Tharp 1978; Star 1977; Vassilas and Ho 2000; Verby *et al.* 1979; Zick *et al.* 2007). Finally, two studies which were written in Japanese (Endo 2008; Tomita and Tagami 1999) were not included. A total of 33 studies were included in the meta-analysis.

Coding the studies

Three types of characteristics were coded for each study: the content of the intervention, the sample population and the methodological characteristics (see below). Each study was coded independently by two raters. The inter-rater reliability was determined using Cohen's kappa (κ) for nominal variables and the intraclass correlation coefficient (*ICC*) for interval variables. Variables coded with an inter-rater reliability of less than .70 were not included in the analysis (the values are reported in brackets below for each coded characteristic). If the raters failed to agree, they reassessed their codings individually and then consulted with one another in order to determine the final coding.

Intervention characteristics Interventions were coded according to whether they contained forms of instruction in addition to VF (see hypothesis 1), such as an oral and/or written explanation of the target skills (1), modeling by an expert or video (1), and exercises (.89). To test hypothesis 2, we coded each study for the presence or absence of an observation form (.94). To describe the VF, we also coded the length of the video recording (1), whether video excerpts were selected (.87), those present at the follow-up discussion (1), and the number of days that elapsed between filming and discussion (.99). The different ways of viewing videotapes were coded: playing at normal speed or in slow motion (1), freeze frame (.89), with or without sound (1), the split-screen technique (1), and serial viewing (1). We also coded the period in which the program was offered (.95), the number of sessions (.98), and whether the program included a follow-up session (1).

Course participant characteristics In line with the classification developed by Huhra *et al.* (2008), in order to test hypothesis 3 we coded according to whether the participants had less than 1 year's practical experience (level 1), more than 1 year including completion of their internship (level 2) or were working as professionals (level 3) (.78). We also coded for whether participants were undergoing training (1), and if so, whether they were

undergraduates or graduates (.85). The average age of participants (.99), the number of years' work experience (1) and the relevant professional sector (.95) were also coded.

Methodological characteristics We coded for the following methodological characteristics: the presence or absence of a control group (.93), random assignment to conditions (.93), the presence of an alternative intervention in the control group (.94) and the presence or absence of external evaluation (.93). The behavioral aspect was coded for each independent variable according to a distinction between verbal, non-verbal, and paralingual outcome measures (see Hargie and Dickson 2004). Using Hulsman *et al.'s* (1999) classification, we also coded for receptive, informative or interpersonal-affective skills (.96, .96, and .92, respectively); outcome measures can relate to more than one behavioral and/or skills domain. For each outcome measure, we coded whether it involved a microskill or a molar skill (.97). A microskill is defined as a highly specific skill scored by means of event sampling (i.e., such as the number of times a course participant looks at the pupil during a session or the number of questions asked). A molar skill is defined as a broader skill assessed by means of a rating scale, such as rating the participant's degree of empathy or responsiveness. We also coded for "negative" (e.g., nervousness or passivity) or "positive" variables (e.g., active listening, authenticity, and a focus on client statements; .95). Finally, the effect size was determined by two raters for both the pre- (.99) and the post-test (.99).

Analyses

The effect measure used is Hedges' *g*, which corrects for bias with small samples. For studies that did not report on means and standard deviations, the effect size was established on the basis of other data with the help of formulae from Borenstein (2009). Effect sizes for "negative" variables were consistently converted so that all positive results corresponded to a positive value. For experimental comparisons involving "within" designs, the standard error for the effect measures was determined using Becker's (1988) formula (see also Morris and DeShon 2002); correlations between the pre- and post-test were never reported, and a conservative estimate of .5 was used in the calculation. In all, 217 experimental results have been derived from 33 experimental studies.

The experimental effects were aggregated by means of a multi-level random effects model (Bryk and Raudenbush 2002; Raudenbusch 2009), which takes into account to the hierarchical structure of the data, in which the experimental comparisons are nested under interventions. A multi-level regression model was used to analyze whether results were moderated by the study characteristics. By means of hierarchical regression analysis, we first checked for any statistically significant correlation between methodological variables and study results. After including these methodological characteristics in the regression model, we then tested whether intervention-related characteristics could explain additional variance in study results. The model was determined using the restricted maximum likelihood method (Hox 2002).

Results

Description of the video feedback programs

VF was investigated for participants engaged in initial vocational training (58%) and refresher courses (42%). Participants who were already qualified had on average 3.3 years'

work experience (SD=6.2), ranging from 0 to 17.5 years. This means that all the skill levels distinguished by Huhra *et al.* (2008) were covered in the experimental studies, with 39%, 12%, and 49% for participants at levels 1, 2, and 3, respectively (see Table 1 for an overview). Only twelve studies reported the participants' ages, which ranged from 20 to 45, with an average of 30 years (SD=9.1).

Table 1 Overview of intervention characteristics

Study	Instruction	No. sessions	VF with…	Form	Student	Level
Aalsvoort and Gossé (2007): I	–	4	T + P	Yes	No	3
Aalsvoort and Gossé (2007): II	–	4	T + P	Yes	No	3
Bailey *et al.* (1977)	P	2	T + P	No	Yes	1
Bolger (2008)	E + P	10	T + P + C	Yes	Yes	1
Caris-Verhallen *et al.* (2000): I	E + P	9	T + P + C	No	No	3
Caris-Verhallen *et al.* (2000): II	E + P	9	T + P + C	No	No	3
Dawson *et al.* (1975)	–	1	T + P	No	Yes	1
Elicker *et al.* (2008)	E	3	T + P	Yes	No	3
Ellet and Smith (1975)	E	7	P	Yes	No	3
Fukkink and Tavecchio (2010)	–	4	T + P	No	No	3
Gask *et al.* (1989)	E + M + P	7	T + P + C	Yes	Yes	3
Girolametto *et al.* (2003)	E + M + P	6	T + P	Yes	No	3
Girolametto *et al.* (2004)	E + M + P	3	T + P	Yes	No	3
Goldberg *et al.* (1980)	E + P	4	T + P	No	Yes	2
Kaye *et al.* (2000)	–	4	T + P	Yes	No	3
Kimbrough *et al.* (2008)	E	2	T + P	Yes	Yes	1
Klein (1999)	E + M	6	T + P + C	No	Yes	1
Maguire *et al.* (1978)	E + P	4	T + P	Yes	Yes	1
Maguire *et al.* (1986)	E + P	4	T + P	Yes	Yes	2
McCormick *et al.* (1993)	E	2	T + P	Yes	Yes	1
Razavi *et al.* (1993)	E + P	8	T + P + C	No	No	3
Roter *et al.* (2004)	E + P	3	T + P	Yes	Yes	2
Scheidt *et al.* (1986): external evaluator	E	4	T + P	Yes	Yes	1
Scheidt *et al.* (1986): self-critique	E	4	P	Yes	Yes	1
Schlundt *et al.* (1994)	E + M + P	4	T + P + C	Yes	Yes	2
Schmitz (1975)	E + M	1	T + P	Yes	Yes	1
Sorenson and Pickett (1986)	E + P	4	–	No	Yes	1
Van Dulmen and Holl (2000)	E + P	5	T + P + C	No	No	3
Van Dulmen and van Weert (2001)	E + P	3	T + P + C	No	No	3
Verby *et al.* (1980)	E	6	T + P	No	Yes	3
Williams (1975)	E + M + P	3	T + P + Con	Yes	No	3
Wilson (1975)	–	6	P	No	Yes	1
Yusuf (2006)	E	1	T + P + C	No	Yes	1

Explanation goals = explanation of goals of training; instruction (in addition toVF) = explanation (E), modelling (M) and/or practice (P) or no additional instruction (–); VF with = VF session with trainer (T), practitioner (P), colleague(s) (C) and/or special consultant (Con) or individual (–); chart = structured observation chart; student = student population (yes) or no; level = level 1 (less than 1 year practical experience), level 2 (>1 year practical experience, completion internship), level 3 (working as professional)

The VF interventions, all of which focused on effective communication in professional practice, had an average duration of about 10 weeks, with an average of 4.4 sessions (SD= 2.3; *min–max*, 1–10). There was no follow-up. Many programs included an explanation of the skills (76%) and exercises being trained (52%). Modeling of the target behavior through videotapes or having the teacher demonstrate the behavior was infrequently part of the program (21%). Participants were filmed on average for 20 min per recording session (SD= 22.5; *k*=25). Most studies did not report whether selection of excerpts took place (61%), although 13 studies (39%) did explicitly report such a selection. Three of these 13 referred explicitly to "positive self-modeling," while the other programs selected critical events that were deemed worthy of a second look. The videotapes were viewed on average 1 week later by the participant and trainer (55%), together with other participants (30%) and once with a special consultant (3%). In some instances, participants viewed the videos alone (10%). For the VF session, 19 studies (58%) included a structured observation form of the relevant interaction skills that were the training focus. Other studies involved no such form (42%). Not all studies reported in detail how the videotapes were viewed. The various technical possibilities that the medium offers (of this paper; Hosford and Mills 1983) were barely mentioned and, judging from the research reports, seem to have been of minor interest. This would suggest that the videotapes were played in the normal way.

Study designs

The most commonly used study design was a controlled design with a pre- and post-test. Random assignment to conditions occurred in half of the controlled studies. A detailed assessment using micromeasures occurred in 70% of cases, with assessment of molar skills occurring less frequently (30%). The majority of outcome measures were positive (88%). Outcome measures involved predominantly verbal skills (82%), and to a far lesser extent non-verbal (33%) and paralingual skills (17%). The outcome measures can be broken down into interpersonal-affective (54%), receptive (47%), and information skills (31%); the numbers do not add up to 100% because one outcome measure can cover more than one domain. An overview of the study design characteristics is presented in Table 2.

Analysis of experimental results

The aggregate effect of VF on professionals' interaction skills is ES=0.40, a medium effect size that is statistically significant (see Table 3). For the verbal, non-verbal, and paralingual domains, the effect sizes are 0.42, 0.35, and 0.39, respectively. Verbal behavior appears to be more easily influenced using the VF method than non-verbal and paralingual behavior, which show no statistically significant differences, although the differences between the three behavioral aspects are slight. The aggregate effect sizes for receptive, informative and relational skills are 0.44, 0.47, and 0.35, respectively. Receptive and productive skills can be more easily enhanced by VF than relational skills, although once again, as with behavioral aspects, the differences are slight.

Moderator analysis

The results are heterogeneous and there is statistically significant variance. In an additional moderator analysis, we therefore investigated which coded study characteristics were systematically associated with the study results. We first analyzed the influence of methodological variables, as it may have been necessary to control these before examining the hypotheses.

Table 2 Overview of methodological characteristics

Study	Pre-test	Design	Random	$N_{\text{EXP}} + N_{\text{CON}}$	Type var.	Pos/Neg var.
Aalsvoort and Gossé (2007): I	Yes	B	NR	2+2	Mo	+
Aalsvoort and Gossé (2007): II	Yes	B	NR	2	Mo	+
Bailey et al. (1977)	Yes	B	R	12+12	Mi	+/−
Bolger (2008)	Yes	W	–	11	Mi	+
Caris-Verhallen et al. (2000): I	Yes	B	NR	21+19	Mi + Mo	+/−
Caris-Verhallen et al. (2000): II	Yes	B	NR	21+19	Mi + Mo	+/−
Dawson et al. (1975)	Yes	W	–	2	Mi	+/−
Elicker et al. (2008)	Yes	W	–	4	Mo	+/−
Ellet and Smith (1975)	Yes	B	NR	20+20	Mo	+
Fukkink and Tavecchio (2010)	Yes	B	R	52+43	Mi + Mo	+/−
Gask et al. (1989)	Yes	W	–	20	Mo	+
Girolametto et al. (2003)	Yes	B	R	8+8	Mi	+
Girolametto et al. (2004)	Yes	B	R	8+9	Mi	+
Goldberg et al. (1980)	Yes	B	R	12+12	Mo	+
Kaye et al. (2000)	Yes	W	–	5	Mi	+
Kimbrough et al. (2008)	Yes	W	–	2+2	Mi	+
Klein (1999)	No	B	R	117+116	Mo	+
Maguire et al. (1978)	Yes	B	R	12+12	Mi + Mo	+
Maguire et al. (1986)	Yes	B	NR	18+18	Mo	+
McCormick et al. (1993)	Yes	W	–	14	Mo	+
Razavi et al. (1993)	No	B	R	36+36	Mi + Mo	+/−
Roter et al. (2004)	Yes	W	–	28	Mi	+
Scheidt et al. (1986): I	No	B	NR	39+33	Mi	+
Scheidt et al. (1986): II	No	B	NR	33+33	Mi	+
Schlundt et al. (1994)	Yes	W	–	60	Mo	+/−
Schmitz (1975)	Yes	B	R	8+8	Mi + Mo	+/−
Sorenson and Pickett (1986)	No	B	R	33	Mo	+
Van Dulmen and Holl (2000)	Yes	B	NR	10+11	Mi + Mo	+/−
Van Dulmen and van Weert (2001)	Yes	W	–	18	Mi	+/−
Verby et al. (1980)	Yes	B	NR	9+8	Mi	+
Williams (1975)	Yes	W	–	15	Mi	+/−
Wilson (1975)	Yes	B	R	5+5	Mi	+
Yusuf (2006)	Yes	B	NR	20+20	Mo	+

Design = "between"-design with control group (B) or "within"-design without control group (W); random = random assignment (R) of not random (NR) of not applicable in case of design without control condition (−); $N_{\text{EXP}} + N_{\text{CON}}$ = number of participants in experimental group + number in control group; type var = micromeasure (Mi) and/or molar measure as dependent variable (Mo); pos/neg = positive (+) and/or negative (−) outcome measure

Of the methodological variables, three were shown to have a statistically significant correlation with study results. The effect sizes are larger for positive than for negative outcome measures; in other words, the results are more positive for measures relating to the desired target behavior that a professional should display or should display more often. The effects are also larger if the independent variable is a molar outcome measure rather than a

Table 3 Aggregate effect of VF on professionals' communication skills

Parameters	ES	SE
Fixed effect		
Integrated effect	0.40*	0.07
Random effect		
Variance between studies	0.10*	0.04

*$p<0.05$, statistically significant results

micromeasure. A relationship was also found between pre- and post-test effect sizes. The influence of this characteristic on the aggregate effect seems modest, though, because the effect size in the pre-test in favor of the experimental group is small (mean ES=0.05, SD= 0.40). The results do not therefore seem to be strongly influenced by differences in the pre-test. The other methodological characteristics did not show any association with the results of the various studies (see Table 4).

Testing the hypotheses

Hypothesis 1, which predicts better results for VF programs that incorporate instruction, is not supported by the data from this meta-analysis. VF programs with no additional instruction were shown to be just as effective as programs with one or more additional forms of instruction (hypothesis 1a; β=0.13, SE=0.22). We should point out, however, that the majority of the programs we looked at contained one or more instruction components in addition to VF, which does not allow a strong test of this hypothesis. Nor is there a relationship between the number of instructional components over and above the actual VF and the size of the effects (hypothesis 1b; β=0.01, SE=0.08). However, the data does support the second hypothesis, which posits more favorable effects if a structured observation form is used. Training programs that include an observation form show significantly larger effects (ES=0.55) than programs with no such form (ES=0.21). The third and final hypothesis concerns the assumed relationship between a participant's learning stage and a decline in learning effects. Contrary to our prediction, the experimental effects did not decline for the more experienced participants (hypothesis 3). We also found no effect for a specific level compared with other levels (e.g., level 1 versus levels 2 and 3).

Table 4 Overview of moderator variables

	ES	Explained variance at study level (%)
Type of outcome measure		1%
Positive	0.41	
Negative	0.28	
Type of outcome measure		17%
Micro	0.32	
Molar	0.52	
Observation of target skills		43%
Without coding form	0.21	
With coding form	0.55	

An additional analysis of the number of years' work experience and the participant's age did not explain any variation in results either. Table 4 presents an overview of the statistically significant relationships; only the differences in the pre-test, as a non-dichotomous predictor, have not been included.

The moderators in this table were statistically significant in both a simple regression model with one predictor and a multiple regression model with all statistically significant moderators combined. A comparison of the simple and multiple regression models showed negligible difference between the β-values associated with each predictor variable, which is an indication of the robustness of the relationships found. For example, the positive correlation between effect sizes and the use of a standard observation form during the program is statistically significant both with and without correction for methodological characteristics ($\beta=0.34$, SE=.11 and $\beta=0.31$, SE=.11, respectively). The combined regression model predicts the largest experimental effect for training programs involving a standard observation form and in which researchers have opted for positive, molar outcome measures. In the absence of differences between the experimental and control groups in the pre-test, this model predicts an effect of ES=0.68. The explained variance of this model with four predictors is 48% at the level of study results; the remaining, non-explained variance in the random part of the model is still statistically significant (.047, SE=.020).

Discussion

Video feedback is a well-known instructional method that is applied in different training programs in order to improve the interaction skills of a broad group of professionals. This meta-analysis has shown that the video feedback method (VF) is effective for improving professionals' key interaction skills. By seeing themselves on video, professionals are able to improve their receptive, informative and relational skills. This study also shows that VF helps to improve verbal, non-verbal and paralingual aspects of communication in professional settings. VF is therefore an effective method that contributes to a wide range of key professional skills. However, expectations should be qualified slightly for the relational skills domain and for non-verbal aspects of interactional behavior, which seem more difficult to influence.

This meta-analysis also highlights a number of variables associated with the effects of VF, such as those found in experimental studies. The outcomes of VF are considerably greater if a standard evaluation form giving participants an overview of the desired target behavior forms part of the training program. A possible explanation for this finding is that such a form structures the observation, thereby focusing the participants' attention on the aspects of their own behavior that are central to the program. Both with and without previous instruction, participants who are given insufficient pointers about what to focus on may find it hard to concentrate on important, substantive aspects and may be distracted by superficial impressions or a one-sided focus. Structured observation forms enable participants—to use a metaphor borrowed from VF—to zoom in and focus on the professional target behavior that is practiced within the training program. No relationship was found with the presence of other forms of instruction supplementary to VF. The use of observation forms during the feedback sessions, which are very much at the heart of VF, therefore, emerge in this meta-analysis as more effective than other instructional

components, such as explaining, modeling, and practicing the target skills. Seen from a historical perspective, the outcome of this meta-analysis is, at least partially, related to the emerging understanding among researchers that it is not practice but feedback that is probably the most crucial dimension in terms of changing the trainee's behavior (see Cooper and Allen 1970). In addition, our study specifically suggests that an evaluation form enhances the power of feedback.

We found no significant relationship, assumed in some publications, between a participant's developmental level and the results of VF training. Individuals may indeed make the greatest progress in their professional development at the beginning of their training and thereafter progress in smaller steps, as Huhra and colleagues (2008) hypothesize. This meta-analysis shows, however, that training results expressed as an effect size (i.e., in relative terms, compared with the control group or the trainee's own initial level) are just as large for students at the start of their training as participants who are further ahead in their professional development. Formulated in positive terms, this means that VF is an effective method for a broad group of participants, from beginners through to professionals with some years' work experience.

This study also highlights three methodological variables as important moderators of experimental effects. Firstly, the results of the studies are a little larger for positive than for negative outcome measures. A possible explanation for this systematic difference in favor of positive outcome measures is that the VF training programs under review were primarily designed for the acquisition or improvement of target skills rather than for "unlearning" less effective behavior. This finding also raises the question of just how VF develops skills. Hosford and Mills (1983) state that by emphasizing positive behavior, VF gradually reinforces such behavior and "suppresses" other, less effective behavior. On the basis of the present study, this appears to be somewhat too optimistic. Our study suggests that although VF approaches do have a positive influence on positive behavior, they do not reduce or eliminate "negative" behavior to the same degree, and certainly not "automatically." This would imply that VF training programs for professionals should reinforce effective behavior (e.g., in a positive self-modeling approach) but should also work in a targeted fashion on reducing less effective behavior. To quote the refrain from the well-known song by Johnny Mercer, "You've got to accentuate the positive, eliminate the negative." Kluger and DeNisi's (1996) meta-analysis shows that both positive and negative feedback can promote learning, provided that the negative feedback is not directed at the person and does not erode their sense of self-worth or motivation to learn. It is also important for participants in training programs to be offered alternatives to less effective behavior (Hattie and Timperley 2007). Viewed in this light, it would be interesting in future research to develop VF training programs that incorporate both positive and negative feedback and to systematically check whether this combined approach works for both "positive" and "negative" outcome measures.

One result of this study, which may at first glance seem paradoxical, is that the experimental effects are systematically smaller for outcome measures at a microlevel. One could argue that evaluating an effective training course focusing on very specific behavior (e.g., asking open questions) should show large effects on a measure relating specifically to that particular skill. However, it is precisely molar measures that show larger effects. A possible explanation is that VF training leads to a fairly broad improvement in skills, broader than is operationalized by the micromeasures. This may make the molar research measures more sensitive to the fairly broad learning effect on participants than the micromeasures, which by definition are limited in scope. A further explanation concerns the quantitative character of the micromeasures. It is quite

conceivable that, following VF instruction, professionals will apply a particular interactional skill (e.g., looking at the client) more often but that once a certain minimum has been reached, they strike a ceiling where more frequent application of that skill would not contribute to the quality of communication. This suggests that an improvement in professional skills is a qualitative change rather than a quantitative one and molar measures may be better suited to capture this type of development.

Limitations of this study

Various authors have already highlighted the limitations of experimental studies into communication training courses in general, which has resulted in us knowing little about their effects. The limitations they have identified relate to the narrow scope of the training programs and to the small-scale nature of the study designs (Alberts and Edelstein 1990; Chant *et al*. 2002; Cegala and Broz 2002; Hill and Lent 2006; Hulsman *et al*. 1999; Kruijver *et al*. 2000; Kurtz *et al*. 1985). These shortcomings are equally present in the experimental VF research that is summarized in this study. A contributing factor here is the painstaking nature of VF studies. Supervision (often one-to-one) and the use of videotape in training are both labor-intensive. Just as painstaking is the work required of researchers when assessing the videotape, in particular when seeking to make a detailed evaluation of the effects of training on the basis of various micromeasures (see Derry *et al*. 2010).

Another limitation of this meta-analysis is that the lack of detailed information in the research reports makes it difficult to classify the content of the VF interventions. For some studies, we were unable to ascertain which videotape excerpts were viewed. It was not always clear whether participants primarily saw excerpts showing successful interactions, less successful interactions or a combination of the two. For this reason, we were unable to determine with any certainty whether it was a case of the positive self-modeling that Dowrick (1983) and Hosford (1980) have argued for. We therefore recommend that future research should indicate whether the videotape excerpts in which the participants view themselves involve a selection of successful or less successful interactions or a combination of the two. Given the concern about the potential effect of negative feedback on participants, it might be advisable to describe the affective aspect of feedback (Shute 2008). Other feedback characteristics from the educational literature (see for example, Hattie and Timperley 2007; Kluger and DeNisi 1996) might also be relevant for the classification of feedback.

Conclusion

Experimental studies have meant that VF is now regarded as a method for which we need not fear negative effects and may expect positive effects. The concerns about possible harmful consequences of VF have been pushed into the background through pedagogically sound approaches using videotape with an eye to the psychology of participants. In fact, confidence with regard to positive effects, which Fuller and Manning already expressed in their first review, has been confirmed in various studies. Future studies should offer further experimental support for the effectiveness of VF, but should above all clarify which approaches are more effective, thereby contributing to the optimum design of skills training for professionals.

References

*Aalsvoort, G. M., & Gossé, G. (2007). Effects of videorecorded interactions and counselling for teachers on their responses to preschoolers with intellectual impairments. *Intellectual and Developmental Disabilities, 45*(2), 103–115

Ajayi-Dopemu, Y., & Talabi, J. K. (1986). The effects of videotape recording on microteaching training techniques for education students. *Journal of Educational Television, 12*(1), 39–44.

Alberts, G., & Edelstein, B. (1990). Therapist training: A critical review of skill training studies. *Clinical Psychology Review, 10*, 497–511.

Allen, D. W. (1966). A new design for teacher education: The teacher intern program at Standford University. *The Journal of Teacher Education, 17*(3), 296–300.

Allen, D. W. (Ed.). (1967). *Microteaching: A description*. Palo Alto, CA: Stanford Teacher Education Program.

*Bailey, K. G., Deardorff, P., & Nay, W. R. (1977). Students play therapist: Relative effects of role playing, videotape feedback, and modeling in a simulated interview. *Journal of Consulting and Clinical Psychology, 45*(2), 257–266.

Baker, S. B., & Daniels, T. G. (1989). Integrating research on the microcounseling program: A meta-analysis. *Journal of Counseling Psychology, 36*(2), 213–222.

Baker, S. B., Daniels, T. G., & Greeley, A. T. (1990). Systematic training of graduate-level counselors: Narrative and meta-analytic reviews of three major programs. *The Counseling Psychologist, 18*(3), 355–421.

Bandura, A. (1969). *Principles of behavior modification*. New York: Holt, Rinehart &Winston.

Bandura, A. (1978). Self-efficacy: Toward a unifying theory of behavioural change. *Advances in Behavior Research and Therapy, 1*, 139–161.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Berthelsen, D.C. & Brownlee, J.M. (2007). Working with toddlers in child care. *Early Childhood Research Quarterly, 22*(3), 347–362.

*Bolger, J. (2008). Video feedback and its impact on the development of communication skills within social work education. Aberdeen, UK: Robert Gordon University.

Borenstein, M. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 221–235). New York: Russell Sage Foundation.

Borg, W. R. (1972). The minicourse as a vehicle for changing teacher behaviour: A three-year follow-up. *Journal of Educational Psychology, 63*(6), 572–579.

Borg, W. R., Kelley, M. L., Langer, P., & Gall, M. D. (1970). *The mini course; A microteaching approach to teacher education*. Beverly Hills, CA: Macmillan.

Brinko, K. T. (1993). The practice of giving feedback to improve teaching. *Journal of Higher Education, 64* (5), 574–593.

Brown, J. A., & Kameen, M. C. (1975). *Changing teacher competence perceptions by focused video-tape feedback*. (ERIC Document Reproduction Services No. ED110896).

Bryk, A. S., & Raudenbush, S. W. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Caris-Verhallen, W. M. C. M., Kerkstra, A., Bensing, J. M., & Grypdonck, M. H. F. (2000). Effects of video interaction analysis training on nurse-patient communication in the care of the elderly. *Patient Education and Counseling, 39*, 91–103.

Cassata, D. M., & Clements, P. W. (1978). Teaching communication skills trough videotape feedback: A rural health program. *Bioscience Communications, 4*, 39–50.

Cassata, D. M., Conroe, R. M., & Clements, P. W. (1977). A program for enhancing medical interviewing using video-tape feedback in the family practice residency. *The Journal of Family Practice, 4*(4), 673–677.

Cassata, D. M., Harris, I. B., Bland, C. J., & Ronning, G. F. (1976). A systematic approach to curriculum design in a medical school interview course. *Journal of Medical Education, 51*, 939–942.

Cegala, D. J., & Broz, S. L. (2002). Physician communication skills training: A review of theoretical backgrounds, objectives and skills. *Medical Education, 36*, 1004–1016.

Chant, S., Jenkinson, T., Randle, J., Rusell, G., & Webb, C. (2002). Communication skills training in the healthcare: A review of the literature. *Nurse Education Today, 22*, 189–202.

Cooper, J. M., & Allen, D. W. (1970). *Microteaching: History and present status*. (ERIC Document Reproduction Services No. ED036471).

*Dawson, P.J., Dawson, K.E., & Forness, S.R. (1975). Effect of video feedback on teacher behavior. *Journal of Educational Research*, *68*(5), 197–201.

Derry, S. J., Pea, R. D., Barron, B., Engel, R. A., Erickson, F., Goldman, R., et al. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *The Journal of the Learning Sciences, 19*, 3–53.

Dowrick, P. W. (1991). *Practical guide to using video in the behavorial sciences*. New York, NY: Wiley.

Dowrick, P. W. (1999). A review of self modelling and related interventions. *Applied & Prevention Psychology, 8*, 23–39.

Dowrick, P. (1983). Self-modeling. In P. Dowrick & S. Biggs (Eds.), *Using video: Psychological and social applications* (pp. 105–124). New York: Wiley.

Duffy, F., Gordon, G. H., Whelan, G., Cole-Kelley, K., & Frankel, R. (2004). Assessing competence in communication and interpersonal skills: The Kalamazoo II Report. *Academic Medicine, 79*(6), 495–507.

*Elicker, J., Georgescu, O., & Bartsch, E. (2008). Increasing the sensitivity of childcare providers. In F. Juffer, M.J. Bakermans-Kranenburg, & M.H. van IJzendoorn (Eds.), *Promoting Positive Parenting* (pp. 155–170). New York: Taylor & Francis Group.

*Ellet, L. E., & Smith, E. P. (1975). Improving performance of classroom teachers through videotaping and self-evaluation. *AV Communication Review, 23*(3), 277–287.

Endo, A. (2008). Intervention to change guidance provided by a teacher in a special education classroom: Feedback to avoid the teacher's resistance. *Japanese Journal of Educational Psychology, 56*(1), 116–126.

Engel, I. M., Resnick, P. J., & Levine, S. B. (1976). The use of programmed patients and videotape in teaching medical students to take a sexual history. *Journal of Medical Education, 51*, 425–427.

Fantuzzo, J., Childs, S., Hampton, V., Ginsburg-Block, M., Coolahan, K. C., & Debnam, D. (1997). Enhancing the quality of early childhood education: A follow-up of an experiential, collaborative training model for Head Start. *Early Childhood Research Quarterly, 12*, 425–437.

Fantuzzo, J., Childs, S., Stevenson, H., Coolahan, K. C., Ginsburg, M., Gay, K., et al. (1996). The Head Start Teaching Center: An evaluation of an experiential, collaborative training model for Head Start teachers and parent volunteers. *Early Childhood Research Quarterly, 11*, 79–99.

Ford, J. D. (1979). Research on training therapists and clinicians. *Review of Educational Research, 49*, 87–130.

Fukkink, R. (2008). Video feedback in widescreen; A meta-analysis of family programs. *Clinical Psychology Review, 28*(6), 904–916.

Fukkink, R. G., & Tavecchio, L. W. C. (2010). Effects of Video Interaction Guidance on early childhood teachers. *Teaching and Teacher Education, 26*, 1652–1659.

Fuller, F. F., & Manning, B. A. (1973). Self-confrontation reviewed: A conceptualization for video playback in teacher education. *Review of Educational Research, 43*(4), 469–528.

Fyffe, A. E., & Oei, T. P. S. (1979). Influence of modelling and feedback provided by the supervisors in a microskills training program for beginning counsellors. *Journal of Clinical Psychology, 35*(3), 651–656.

Gask, L. (1992). Training general practitioners to detect and manage emotional disorders. *International Review of Psychiatry, 4*, 293–300.

Gask, L. (1998). Small group interactive techniques utilizing videofeedback. *Journal of Psychiatry in Medicine, 28*, 97–113.

Gask, L., Goldberg, D., Boardman, J., Craig, T., Goddard, C., Jones, O., et al. (1991). Training general practitioners to teach psychiatric interviewing skills: An evaluation of group training. *Medical Education, 25*, 444–451.

*Gask, L., Goldberg, D., Porter, R., & Creed, F. (1989). The treatment of somatization: Evaluation of a teaching package with general practice trainees. *Journal of Psychosomatic Research*, *33*(6), 697–703.

Gask, L., McGrath, G., Goldberg, D., & Millar, T. (1987). Improving the psychiatric skills of established general practitioners: Evaluation of group teaching. *Medical Education, 21*, 362–368.

Ginsburg, C. (1956). A new magnetic video recording system. *Journal of Society of Motion Picture and Television, 65*, 302–304.

*Girolametto, L., Weitzman, E., & Greenberg, J. (2003). Training day care staff to facilitate children's language. *American Journal of Speech-Language Pathology, 12*, 299–311.

*Girolametto, L., Weitzman, E., & Greenberg, J. (2004). The effects of verbal support strategies on small-group peer interactions. *Language, Speech, and Hearing Services in Schools, 35*, 254–268.

*Goldberg, D. P., Steele, J. J., Smith, C., & Spivey, L. (1980). Training family doctors to recognise psychiatric illness with increased accuracy. *The Lancet, 316*, 521–523.

Hargie, O. (2006). Training in communication skills: Research, theory and practice. In O. Hargie (Ed.), *The handbook of communication skills* (3rd ed., pp. 551–565). London, UK: Routledge.

Hargie, O., & Dickson, D. (2004). *Skilled interpersonal communication; Research, theory and practice* (4th ed.). London, UK: Routledge.

Hargie, O., Saunders, C., & Dickson, D. (1983). *Social skills in interpersonal communication*. Hoboken, NJ: Wiley.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.

Hays, V. (1976). Extinguishing ineffective communication behavior: An innovative video-tape-feedback procedure. *Behavioral Engineering, 3*(3), 80.

Hehr, W. (1981). Fokussiertes Videofeedback und der Selbstwahrnehmungsprozeβ [Focused video feedback and the self-observation process]. *Zeitschrift für experimentelle und angewandte Psychologie, 28*(1), 130–159.

Hill, C. E., & Lent, R. W. (2006). A narrative and meta-analytic review of helping skills training: Time to review a dormant area of inquiry. *Psychotherapy: Theory, Research, Practice, Training, 43*, 154–172.

Hosford, R. E. (1980). Self-as-a-model: A cognitive, social-learning technique. *Counseling Psychology, 9*(1), 45–62.

Hosford, R. E., & Johnson, M. E. (1983). A comparison of self-observation, self-modeling, and practice without video-feedback for improving counselor interviewing behaviors. *Counselor Education and Supervision, 23*(1), 62–70.

*Hosford, R. E., & Mills, M. E. (1983). Video in social skills training. In P. W. Dowrick & S. J. Biggs (Eds.), *Using video; Psychological and social applications* (pp. 125–149). Hoboken: Wiley.

Hougham, P. (1992). *Improving student teachers' strategies for asking a range of both high and low level questions through video evaluations*. (ERIC Document Reproduction Services No. ED349309).

Hox, J. J. (2002). *Multilevel Analysis; Techniques and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Huhra, R. L., Yamokoski-Maynhart, C. A., & Prieto, L. R. (2008). Reviewing videotape in supervision: A developmental approach. *Journal of Counseling and Development, 86*, 412–418.

Hulsman, R. L., Harmsen, A. B., & Fabriek, M. (2009). Reflective teaching of medical communication skills with DiViDU: Assessing the level of student reflection on recorded consultations with simulated patients. *Patient Education and Counseling, 74*, 142–149.

Hulsman, R. L., Ros, W. J. G., Winnubst, J. A. M., & Bensing, J. M. (1999). Teaching clinically experienced physicians communication skills. *A review of evaluation studies. Medical Education, 33*, 655–668.

Hung, J. H. F., & Rosenthal, T. L. (1981). Therapeutic videotaped playback. In J. L. Fryrear & B. Fleshman (Eds.), *Videotherapy in mental health* (pp. 5–46). Springfield, IL: Charles C. Thomas.

Ivey, A. E., & Authier, R. J. (1978). *Microcounseling; Innovations in interviewing, counseling, psychotherapy and psycho-education*. Springfield, Il: Charles C. Thomas.

*Kaye, G., Fortsyth, P., & Simpson, R. (2000). Effective interaction in the classroom—towards a new viewpoint. *Educational and Child Psychology, 17*(4), 69–90.

Kern, H. J. (1980). Lehrertraining in der Schulpraxis: Erprobung van Trainingstechniken und Evaluations-verfahren [Teacher training in the classroom—Evaluating training-techniques and assessment procedures with the classroom teacher]. *Psychologie, Erziehung und Unterrichtung, 27*(1), 10–20.

*Kimbrough, S., Davis, J., & Wickersham, L. (2008). The use of video feedback and semi structured interviews for reflection among pre-service teachers. *Journal of Education and Human Development, 2*(2), 1–12.

*Klein, S. (1999). The effects of the participation of patients with cancer in teaching communication skills to medical undergraduates: A randomised study with follow-up after 2 years. *European Journal of Cancer, 35*(10), 1448–1456.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284.

Kpanja, E. (2001). A study of the effects of video tape recording in microteaching training. *Britisch Journal of Educational Technology, 32*(4), 483–386.

Kruijver, I. P. M., Kerkstra, A., Francke, A. L. F., Bensing, J. M., & van de Wiel, H. B. M. (2000). Evaluation of communication training programs in nursing care: A review of the literature. *Patient Education and Counseling, 39*, 129–145.

Kurtz, P. D., Marshall, E. K., & Banspach, S. W. (1985). Interpersonal skill-training research: A 12-year review and analysis. *Counselor Education and Supervision, 24*(3), 249–263.

Kurtz, S., Silverman, J., & Draper, J. (2005). *Teaching and learning communication skills in medicine*. Oxford, UK: Radcliffe Publishing.

Levinson, W., & Roter, D. (1993). The effects of two continuing medical education programs on communication skills of practicing primary care physicians. *Journal of General Internal Medicine, 8*, 318–324.

MacLeod, G. (1987). Microteaching: End of a research era? *International Journal of Educational Research, 11*(5), 531–541.

*Maguire, P., Fairbairn, S., & Fletcher, C. (1986). Consultation skills of young doctors: I—Benefits of feedback training in interviewing as students persist. *British Medical Journal, 292*, 1573–1576.

*Maguire, P., Roe, P., Goldberg, D., Jones, S., Hyde, C., & O'Dowd, T. (1978). The value of feedback in teaching interviewing skills to medical students. *Psychological Medicine, 8*, 695–704.

Marita, P., Leena, L., & Tarja, K. (1999). Nurses' self-reflection via videotaping to improve communication skills in health counseling. *Patient Education and Counseling, 36*, 3–11.

Martin-Reynolds, J. (1980). The effects of a self-evaluation model on the focus reaction of student-teachers during split-screen videotape-feedback. *Journal of Educational Research, 73*(6), 360–364.

*McCormick, D. P., Rassin, G. M., Stroup-Benham, C. A., Baldwin, C. D., Levine, H.G., Persaud, D. I., & Wolf, W. J. (1993). Use of videotaping to evaluate pediatric resident performance of health supervision examinations of infants. *Pediatrics, 92*(1), 116–120.

McKnight, P. C. (1980). Microteaching: Development from 1968 to 1978. *Journal of Education for Teaching, 6*(3), 214–227.

Morris, S. B., & DeShon, R. P. (2002). Combining effect sizes estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7*(1), 105–125.

Napper-Owen, G. E., & Phillips, D. A. (1995). A qualitative analysis of the impact of induction assistance on first-year physical educators. *Journal of Teaching in Physical Education, 14*, 305–327.

Perlberg, A. (1983). When professors confront themselves: Towards a theoretical conceptualization of video self-confrontation in higher education. *Higher Education, 12*(6), 633–663.

Quigley, B. L., & Nyquist, J. D. (1992). Using video technology to provide feedback to students in performance courses. *Communication Education, 41*, 324–334.

Raudenbusch, S. W. (2009). Analyzing effect sizes: Random-effects models. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 295–315). New York: Russell Sage Foundation.

*Razavi, D., Delvaux, N., Marchal, S., Bredart, A., Farvacques, C., & Paesmans, M. (1993). The effects of a 24-h psychological training program on attitudes, communication skills, and occupational stress in oncology: a randomised study. *European Journal of Cancer, 29A*(13), 1858–1863.

Romans, J. S. C., Boswell, D. L., Carlozzi, A. F., & Ferguson, D. B. (1995). Training supervision practices in clinical, counseling and school psychology programs. *Professional Psychology: Research and Practice, 26*(4), 407–412.

*Roter, D. L., Larson, S., Shinitzky, H., Chernoff, R., Serwint, J. R., Adamo, G., & Wissow, L. (2004). Use of innovative video feedback technique to enhance communication skills training. *Medical Education, 38*, 145–157.

*Scheidt, P. C., Lazoritz, S., Ebbeling, W. L., Figelman, A. R., Moessner, H. F., & Singer, J. E. (1986). Evaluation of system providing feedback to students on videotaped patient encounters. *Journal of Medical Education, 61*, 585–590.

*Schlundt, D. G., Quesenberry, L., Pichert, J. W., Lorenz, R. A., & Boswell, E. J. (1994). Evaluation of a training program for improving adherence promotion skills. *Patient Education and Counseling, 24*, 165–173.

Schmidt, D. D., & Messner, E. (1977). The use of videotape-techniques in the psychiatric training of family physicians. *The Journal of Family Practice, 5*(4), 585–588.

*Schmitz, R. A. (1975). Erprobung eines Lehrertrainings durch Micro-Teaching [Evaluation of a teacher training using micro-teaching]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 7*(3), 203–213.

Schoonover, S. C., Bassuk, E. L., Smith, R., & Gaskill, D. (1983). The use of videotape programs to teach interpersonal skills. *Journal of Medical Education, 58*, 804–810.

Scott, J. P., Sollie, D. L., & Duffey, N. (1983). Teaching communication skills with videomodeling and videofeedback in the university classroom. *Home Economics Research Journal, 12*(1), 17–24.

Sherin, M. G. (2004). New perspectives on the role of video on teacher education. In J. Brophy (Ed.), *Using video in teacher education* (pp. 1–27). Oxford, UK: Elsevier.

Sherin, M. G., & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education, 60*(1), 20–37.

Shernoff, E. S., & Kratochwill, T. R. (2007). Transporting an evidence-based classroom management program for preschoolers with disruptive behavior problems to a school: An analysis of implementation, outcomes, and contextual variables. *School Psychology Quarterly, 22*(3), 449–472.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189.

Silverman, J., Kurtz, S., & Draper, J. (2005). *Skills for communicating with patients* (2nd ed.). Abingdon, UK: Radcliffe Publishing.

Sliwa, J. A., Makoul, G., & Betts, H. (2002). Rehabilitation-specific communication skills training: Improving the physician patient relationship. *American Journal of Physical Medicine & Rehabilitation, 81*(2), 126–132.

Sollie, D. L., & Scott, J. P. (1983). Teaching communication skills: A comparison of videotape feedback methods. *Family Relations, 32*, 503–511.

Sorenson, R. L., & Pickett, T. A. (1986). A test of two teaching strategies designed to improve interview effectiveness: Rating behavior and videotaped feedback. *Communication Education, 35*, 13–22.

Speidel, G. E., & Tharp, R. G. (1978). Teacher-training workshop strategy: Instructions, discrimination training, modeling, guided practice, and video feedback. *Behavior Therapy, 9*, 735–739.

Star, B. (1977). The effects of videotape self-image confrontation on helping perceptions. *Journal of Education for Social Work, 13*(2), 114–119.

Star, B. (1979). Exploring the boundaries of videotape self-confrontation. *Journal of Education for Social Work, 15*(1), 87–94.

Star, J. R., & Strickland, S. K. (2008). Learning to observe; Using video to improve preservice mathematics teachers' ability to notice. *Journal of Mathematics Teacher Education, 11*, 107–125.

Tan, A. L., & Towndrow, P. A. (2009). Catalyzing student-teacher interactions and teacher learning in science practical formative assessment with digital video technology. *Teaching and Teacher Education, 25*, 61–67.

Thelen, M. H., & Lasoski, M. C. (1980). The separate and combined effects of focusing information and videotape self-confrontation feedback. *Journal of Behavior Therapy and Experimental Psychiatry, 11*, 173–178.

*Van Dulmen, M. A., & Holl, R. A. (2000). Effects of continuing pediatric education in interpersonal communication skills. *European Journal of Pediatrics, 159*, 489–495.

*Van Dulmen, M. A., & van Weert, J. C. M. (2001). Effects of gynaecological education on interpersonal communication skills. *British Journal of Obstetrics and Gynaecology, 108*, 485–491.

Vassilas, C., & Ho, L. (2000). Video for teaching purposes. *Advances in Psychiatric Treatment, 6*, 304–311.

*Verby, J., Davis, R. H., & Holden, P. (1980). A study of the interviewing skills of trainee assistants in general practice. *Patient Counseling and Health Education, 2*(2), 68–71.

Verby, J., Holden, P., & Davis, R. H. (1979). Peer review of consultations in primary care: The use of audiovisual recordings. *British Medical Journal, 1*, 1686–1688.

White, P., & Poppen, W. (1979). Videotape playback: Analysis of a new technique for counseling. *Psychotherapy: Theory, Research and Practice, 16*(1), 86–90.

*Williams, J. H. (1975). *Training in interaction analysis as a means of staff development for master teachers, classroom teachers and paraprofessionals*. (ERIC Document Reproduction Service No. ED113343).

*Wilson, S. H. (1975). *Determining of two differential self-modeling techniques on the acquisition of appropriate counseling behaviors of counselor trainees*. Knoxville: University of Tennessee

*Yusuf, M. O. (2006). Influence of video and audiotapes feedback modes on student teachers' performance. *Malaysian Online Journal of Instructional Technology, 3*(1), 29–35.

Zick, A., Granieri, M., & Makoul, G. (2007). First-year medical students' assessment if their own communication skills: A video-based, open-ended approach. *Patient Education and Counseling, 68*, 161–166.

Zimmerman, L., Amodeo, M., Fassler, I., Ellis, M., & Clay, C. (2003). Training team leaders in a child welfare setting using the SPIN leadership guidance model. *Children and Youth Services Review, 25*(11), 891–910.

References marked with an asterisk indicate studies included in the meta-analysis.