UNIVERSITY OF OKLAHOMA

GRADUATE COLLLEGE

THE INTERACTION BETWEEN GENERAL COGNITIVE ABILITY AND TRAINING STRUCTURE: USING META-ANALYSIS TO INVESTIGATE APTITUDE-TREATMENT INTERACTIONS IN THE TRAINING LITERATURE

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

VANESSA KOWOLLIK
Norman, Oklahoma
2009

THE INTERACTION BETWEEN GENERAL COGNITIVE ABILITY AND TRAINING STRUCTURE: USING META-ANALYSIS TO INVESTIGATE APTITUDE-TREATMENT INTERACTIONS IN THE TRAINING LITERATURE

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF PSYCHOLOGY

BY

____
Dr. Eric A. Day, Chair

____
Dr. Dan O'Hair

____
Dr. Jorge Mendoza

____
Dr. Lori Anderson Snyder

____
Dr. Robert Terry

Acknowledgements

I would like to thank my major professor, Dr. Eric Day, for his constructive feedback throughout the entire process. He added valuable insights to the design of the project and the analysis of the data—not to forget the structure of the final document. I would also like to I would like to thank my committee, Dan O'Hair, Jorge Mendoza, Lori Anderson Snyder, and Robert Terry for their contributions and continued support throughout. Thanks to the people who helped collect data for this study: Xiaoqian Wang, Matthew Schuelke, Michael Hughes, and my undergraduate research assistants. Many, many thanks to my parents, Adelheid and Wilfried Kowollik, for believing in me and making this possible. And thank you, Yevgeny Sidelman, for loving and supporting me during the final stages of the process.

Table of Contents

List of Tables

List of Figures

Abstract

This study quantitatively reviewed aptitude-treatment interactions (ATIs) in the training literature to address inconclusive results with regard to their existence and nature. The aptitude investigated was trainee general cognitive ability and the treatment was training program structure. 15 studies were investigated with a primary approach that only included studies that directly examined ATIs. 51 studies were investigated with a secondary approach that involved coding the structure of training in studies that reported a correlation between general cognitive ability and training effectiveness. Correlations were meta-analyzed and compared as a function of structure and type of evaluation criterion. Results indicated that small ATIs exist but that the effects vary in size and nature depending on the criterion. For cognitive criteria structure exacerbated the effects of general cognitive ability. In contrast, for skill criteria structure attenuated the effects of general cognitive ability. However, the effects for skill were not robust—they were confounded with training content. The size and nature of the ATIs suggests it is not worthwhile to tailor the structure of training based on trainees' general cognitive ability.

The Interaction between General Cognitive Ability and Training Structure: Using Meta-Analysis to Investigate Aptitude-Treatment Interactions in the Training Literature

Training is big business in U.S. organizations. Over 90% of companies provide training for managerial personnel (Goldstein & Ford, 2002), and an estimated $55.3 billion is spent on formal training, as reported by the American Society for Training and Development (Bassi & Van Buren, 1998). Judging from these numbers, it is vital to organizational success that the implemented training programs are successful. However, training success is not only dependent on the design and implementation of training, which unquestionably most of the money is invested in, but also on *organizational* and *social* characteristics as well as on the *individual* characteristics of the trainees (Noe & Colquitt, 2002). Training researchers have begun to pay increased attention to those three factors, as they are able to "augment or negate the direct impact of training" (Kraiger, 2003, pp. 183). Research on the importance of continuous learning cultures (e.g., Tracey, Tannenbaum, & Kavanagh, 1995), as an example of an *organizational* characteristic, or the importance of supervisory support (e.g., Ford, Quinones, Sego, & Sorra, 1992), as an example of a *social* characteristic, are two lines of research that exemplify this increased interest. The focus of this document however, lies with the *individual differences* of the trainees (e.g., personality, age, attitudes, cognitive ability) because they have been regarded as a significant influence on training outcomes. For example, Noe and Colquitt's (2002) training effectiveness model depicts individual differences as distal predictors influencing training motivation, which in turn influences learning during training and ultimately transfer of training and job performance. This example shows that a consideration of individual differences in training is pivotal. On the one hand, main effects that individual differences have on training effectiveness are important and have practical implications. On the other hand, the extent to which individual differences interact with treatment (i.e., training) conditions indicates a need for more careful thought about matching particular individuals to the most appropriate training programs in a cost effective manner.

The focus of this document is on the latter point. More specifically, the focus is on the interaction between general cognitive ability and the structure of the training program. The reason for choosing general cognitive ability is that it is a frequently assessed individual difference in studies and thought to be the individual difference with the greatest influence on learning, training outcomes, and job performance (Hunter 1986; Ree, Carretta, & Teachout, 1995; Ree & Earles, 1991a; Ree & Earles, 1991b; Schmidt & Hunter, 1998). The reason for choosing structure as the treatment is that most discussions of interactions between aptitude and treatment include structure as the training design factor with which aptitudes are most likely to interact. The interaction between an individual difference and treatment variable is commonly referred to as an aptitude-treatment interaction (ATI).

Prior empirical research on ATIs has predominantly occurred in education, and although conflicting and inconsistent results have been produced in that literature (Bracht, 1970), many scholars and practitioners in education are optimistic that ATIs exist and have important implications (e.g., Cronbach & Snow, 1977). ATIs from the educational literature are frequently mentioned in reviews of organizational training, but currently the extent to which ATIs exist in organizational-based training is unclear.

The central premise of this study is that a meta-analytic approach would be useful to resolve the inconsistencies in the scientific literature and shed more light on the likelihood and nature (direction) of ATIs with respect to organizationally relevant training. That is, a meta-analysis focused on the training literature will better speak to the need for training scholars and practitioners to attend to ATIs when addressing industrial and organizational issues. To do so, the training literature was analyzed with two approaches. The first approach, called the *primary approach* from here on forward, constituted the inclusion and analysis of studies that actually investigated an interaction between general cognitive ability and training structure. In the second approach, called the *secondary approach* from here on forward, studies were included that reported correlations between general cognitive ability and training criteria and the structure of the training programs described was coded. Subsequently, it was then examined if the correlation between general cognitive ability and training effectiveness systematically varied as a function of training structure.

Aptitude-Treatment Interactions

Despite Goldstein's (1980) call for more empirical training research involving individual differences, and more specifically to match individual differences to various instructional strategies, Wexley (1984), a few years later, lamented the dearth of such research. However, new interest has arisen on the role that individual differences play in training effectiveness (Kraiger, 2003). Interestingly, educational psychologists have emphasized the importance of individual differences and their interaction with various types of treatments for decades (Cronbach & Snow, 1977) but only recently have training researchers emphasized their importance (Campbell & Kuncel, 2001).

Cronbach and Snow (1977) asserted thirty years ago that not all individuals are alike and that some individual differences that are correlated with learning interact with the nature of the instructional environment. This suggests that the relationship between an individual difference variable and training effectiveness is not the same across different training programs. Some people might benefit more from one type of training while others might benefit more from a second type of training. The major implication is not to administer the same training program to all trainees when the interest lies in maximizing aggregate gains across all trainees—one size does NOT fit all when it comes to training (Campbell & Kuncel, 2001).

The early focus on research into ATIs, mainly by educational psychologists, primarily relied on investigating aptitudes (i.e., potential), such as general cognitive ability. Although there may be different kinds of ATIs, the most frequently discussed one is the interaction between aptitude and structure of the training content, where structure refers to instructor guidance, detailed objectives for trainees, explicit specification of the content to be taught, and frequent instructor feedback (Campbell & Kuncel, 2001). It has been thought that trainees with lower aptitude tend to do better in more structured training programs and trainees with higher aptitude tend to do better in less structured training programs.

In an early review of the educational psychology literature, Snow and Lohman (1984) cited several ATI findings (e.g., Dansereau, McDonald, Collins, Garland, Holley, Diekhoff, & Evans, 1979; Rigney, Munro, & Crook, 1979; Sharps, 1974). For example, Dansereau et al. (1979) found a disordinal interaction between crystallized intelligence (termed *verbal crystallized ability*) and networking strategy training on retention of text passages. The networking strategy program taught students a technique for transforming text into conceptual networks (i.e., high structure) while the trainees in the control program did not receive such strategy training (i.e., low structure). Results showed that students with high crystallized intelligence had higher achievement scores without strategy training than those with strategy training, while students with low crystallized intelligence had higher achievement scores with strategy training than those without it. As another example of a disordinal interaction between crystallized intelligence and training, Sharps (1974) showed that students with high crystallized intelligence attained higher reading comprehension scores in a conventional teaching program (i.e., low structure) than in an individually prescribed instruction program (i.e., high structure), while those with lower crystallized intelligence showed higher reading comprehension scores in the individually prescribed instruction program than those in the conventional teaching program.

Although these studies, and several others (e.g., Peterson, Janicki, & Swing, 1979) cited by Snow and Lohman (1984) are labeled as investigations of ATIs, they all investigated crystallized intelligence. While crystallized intelligence reflects the attainment of knowledge (i.e., intelligence as a store of knowledge acquired over time), aptitude refers more purely to future potential. Crystallized intelligence is thus not the same as what is considered a general aptitude and has to be distinguished from it. Hence, much of the research in the educational literature on ATIs has not focused on general cognitive ability per se, although a few studies have investigated it (e.g., Kyllonen, Lohman, & Snow, 1984). The present investigation solely concentrates on general cognitive ability (i.e., potential) as the individual difference variable involved in ATIs.

As mentioned earlier, research on ATIs in the educational literature has often provided conflicting results. For example, in one of the first reviews on the existence of ATIs Bracht (1970) concluded that their existence is unlikely. A second opinion is that they exist but that they provide little to no value, meaning that administering different treatments to different individuals is impractical (Gehlbach, 1979). Finally, a third opinion put forth by Berliner & Cahen (1973) and Snow & Lohman (1984) is that ATIs exist and are of value and the only reason for the relatively unsupportive research is the unsound methodology of those studies.

Hence, a meta-analysis might provide a remedy to these conflicting ATI reviews, and focusing on the training literature brings the idea of administering different training programs to different trainees closer to the field of industrial and organizational psychology. However, mere knowledge about the existence of ATIs is not enough—it is also important to know the way that general cognitive ability and training program structure interact.

Types of Interactions

Aptitudes and treatments can interact in one of two major ways to influence training effectiveness (Bracht, 1970). The first major way is a disordinal interaction, in which one treatment yields high effectiveness for trainees who are at one end of the aptitude continuum while a different treatment yields high effectiveness for trainees who are at the other end of the continuum (Figure 1).

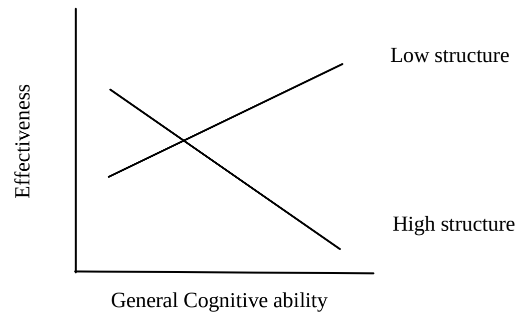

Figure 1. Disordinal Interaction

This interaction is represented by the classical “X” crossing of the regression lines and shows that to maximize training effectiveness it is best to administer different training programs to the individuals at each end of the aptitude continuum, in this case general cognitive ability. This type of interaction does not exhibit a main effect for either general cognitive ability or training program structure. Here, the correlation between general cognitive ability and training effectiveness is positive in training programs with low structure and negative in training programs with high structure. Although this kind of interaction is typically used to characterize ATIs, it contradicts the robust finding that aptitude is positively related with not only job performance (Hunter, 1986) but also training performance (Ree et al., 1995; Ree & Earles, 1991a). This type of interaction also contradicts the commonly held belief that structure and guidance are beneficial to learning (e.g., Kozlowski, Toney, Mullins, Weissbein, Brown, & Bell, 2001) and that open or discovery learning detracts from training effectiveness (e.g., Brown, 2001; Kirschner, Sweller, & Clark, 2006). Therefore, no support for this model in the present meta-analysis is expected to be found.

The second major way for aptitudes to interact with training program structure is an ordinal interaction, in which one treatment is superior, regardless of where on the aptitude continuum the trainee falls. This type of interaction has many variants. One of the variants of this interaction is represented by an “inverse spreading interaction” where one regression line is almost invariant across the aptitude continuum while the other regression line varies across the aptitude continuum in such a way that structure attenuates aptitude effects (Figure 2).

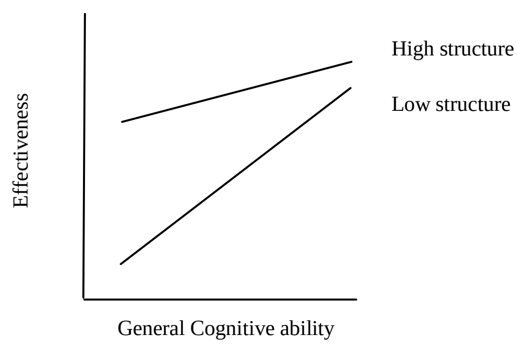

Figure 2. Ordinal Interaction – Inverse Spreading Interaction

Here, trainees on the high end of the aptitude continuum achieve the same or slightly more regardless of the training structure, but trainees on the low end of the aptitude continuum benefit much more from high training structure. Hence, it is important, when one is interested in only administering one training program, to administer the one that benefits trainees on both ends of the aptitude continuum. This kind of interaction exhibits main effects for training program structure and general cognitive ability. Here, the correlation between general cognitive ability and training effectiveness is larger for training programs with low structure than it is for those with high structure. With such an ATI, training program structure compensates for low general cognitive ability. Support for this model is more likely than for the first model given the main effects. The greater benefit of structure for individuals low on general cognitive ability is loosely consistent with the aforementioned disordinal model. The less there is a positive main effect for structure, the more likely it is that a meaningful disordinal (cross-over) interaction is occurring.

Another variant is represented by a "spreading interaction" where the regression lines do not intersect but rather fan out across the aptitude continuum in such a way that structure exacerbates aptitude effects (Figure 3).

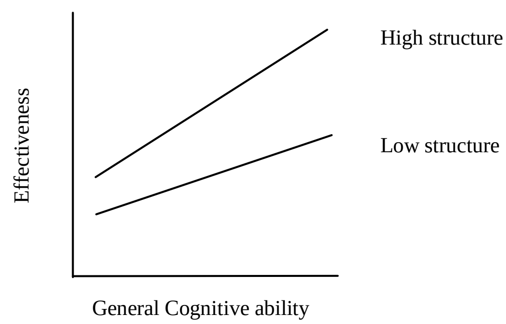

Figure 3. Ordinal Interaction – Spreading Interaction

This type of interaction shows that those high in general cognitive ability benefit more from structure than those low in general cognitive ability. Here, the correlation between general cognitive ability and training effectiveness is larger under high structure training programs than under low ones. This type of interaction also shows main effects for training program structure and general cognitive ability. This type of ATI has not been directly examined or explicitly mentioned in the training literature. However, research has shown that individuals high in aptitude benefit more from practice in general, compared to individuals lower in aptitude (Fleishman &

Mumford, 1989; Day, Bell, Ewards, Bennett, Mendoza, & Tubré, 2005). Relatedly, Ceci and Papierno (2005) noted that educational interventions that are implemented to decrease the gap between disadvantaged and advantaged youths, across various domains, often lead to not only benefitting the disadvantaged but also benefiting the advantaged more so, thereby widening the preintervention gap further. For example, Borkowski and Peck (1986) examined gifted and nongifted children in the training and transfer of strategies. A metamemory battery was universally administered prior to and after a training intervention on cognitive strategies. The gap of scores on metamemory between the gifted and nongifted children, which was statistically significant before training began, increased even further after training. This gap persisted on transfer tasks. Another example comes from Ruiz (1985) who investigated the academic performance of low and high socioeconomic status (SES) adolescents. After all adolescents were given a cognitive intervention program, they improved their performance. However, high SES adolescents showed greater gains than their low SES counterparts. Although the interactions shown in Figures 2 and 3 are both ordinal in nature, they contradict each other in terms of who benefits most from structure. Again, as mentioned above, the less there is a positive main effect for structure, the more likely it is that a meaningful disordinal (cross-over) interaction is occurring.

Alternatively, the strength of an ATI may be relatively small if not nonexistent. In this model (Figure 4), regardless of where on the aptitude continuum trainees fall, one training program is always equally more beneficial for all trainees, and aptitude effects are the same regardless of structure.

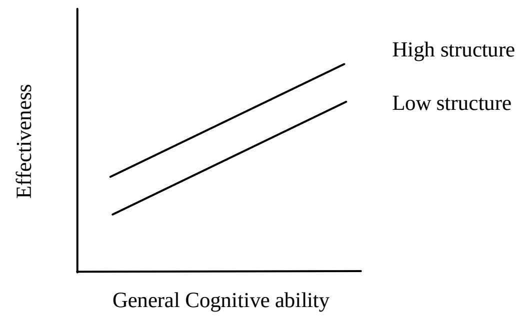

Figure 4. No Interaction

In other words, the correlation between general cognitive ability and training effectiveness is the same under low and high structure training programs. This type of model shows main effects for general cognitive ability and training program structure, but no ATI.

Current Meta-Analysis

This meta-analysis examined the interaction between general cognitive ability and the structure of the training program. In general, these main research questions were examined: (1) does a general ATI exist in the training research literature, (2) if so, what is the nature of this interaction, and (3) what is the strength of the interaction? In addition to addressing these questions, this meta-analysis is worthwhile for two other reasons. One, this meta-analysis examined if the presence of ATIs depends on other variables, namely the type of training criteria (e.g., skill-based, cognitive, or affective/attitudinal). Second, the correction for sampling error and measurement unreliability helps lessen the extent to which statistical artifacts mask the presence of ATIs (Whitener, 1989).

Whitener (1989) conducted, to my knowledge, the only meta-analysis on ATIs. The meta-analysis focused on studies published in the educational psychology literature. Whitener's meta-analysis examined the interaction between prior achievement (i.e., knowledge) and instructional support on learning achievement. To test for the nature and strength of the ATI she investigated if the difference between the regression slopes for higher and lower instructional support would be greater than zero. Collecting information on a total of eleven studies, she found a weighted average regression coefficient for the interaction term of .11 with a 90% confidence interval ranging between -.01 and .23, leading Whitener to suggest that these results show the existence of ATIs. In this particular case, the slope for higher instructional support was greater than the slope for lower instructional support, such that students with high prior achievement benefited more from increased instructional support than students with low prior achievement with regard to learning achievement. On the other hand, there was not such a big difference between students of high and low prior achievement when instructional support was low. Her results would support an interaction model that is similar to Figure 3 (i.e., spreading interaction).

Whitener's meta-analysis only consisted of eleven studies and the literature has since been greatly expanded. This expansion has also included the training literature, which has produced a considerable increase of studies investigating the interaction of trainee attributes and training program characteristics (e.g., Goska & Ackerman, 1996; Gully, Payne, Koles, & Whiteman, 2002; Kanfer & Ackerman, 1989; Towler & Dipboye, 2001). A meta-analysis is thus of great use to the training literature for shedding light not only on the existence of ATIs but also on their size and nature with regard to training effectiveness. The current meta-analysis comes at an important time because of the augmented interest of individual differences in the training literature. The size of the effect (as well as the type of the interaction) will help clarify if administering differing training programs to different trainees is advantageous. If the effect is negligible, spending valuable resources on developing and administering differing training programs would be inefficient. However, if the effect is strong, organizations will have much to gain by designing and providing differing training programs.

A general account of the number of studies that have investigated an ATI, as defined in terms of ability and not non-ability individual difference variables, was also conducted. Providing a simple count of how many studies investigated ATIs contrasted with a count of how many studies could have but did not to examine ATI effects will provide an account reflecting the general interest in conventional ATIs among applied researchers. This should be especially interesting in light of how much increased attention has been paid to the importance of individual differences in general in the training literature in recent years. If this meta-analysis shows that ATIs exist and have a sizeable effect, this overview of the state of ATI research would be even more interesting and hopefully spur future research investigating these interactions. Additionally, the reviewed studies that investigated ATIs were also meta-analyzed to provide a point of comparison against the other meta-analytic results involving studies coded for structure, which reported a correlation between general cognitive ability and a training criterion, but did not specifically examine an ATI.

Method

Literature Search

An extensive literature search was conducted to identify empirical studies that involved an evaluation of an organizationally relevant training program in which the cognitive ability of trainees was part of the investigation. For this search the published and unpublished training literature from 1959 to 2007 was reviewed. The literature search encompassed studies published in journals and books, book chapters, conference papers and presentations, and dissertations that are related to the evaluation of an organizational training program or to the empirical evaluation of an organizational training method or approach.

The process started with a search of eight computer databases (*Academic Search Elite, Business Source Elite, Dissertation Abstracts, Econlit, Educational Research Information Center, Government Printing Office, PsychINFO, and SocINDEX*) using the following keywords: *training, training effectiveness, training evaluation, individual differences, attribute treatment interaction, aptitude treatment interaction, aptitude*, and *cognitive ability*. The electronic search was supplemented by a manual search of the reference lists of recent reviews of the training literature (Alliger & Janak, 1997; Alliger, Tannenbaum, Bennett, Traver, & Shotland, 1997; Arthur, Bennett, Edens, & Bell, 2003; Arthur, Bennett, Stanush, & McNelly, 1998; Colquitt, LePine, & Noe, 2000), a search of the Department of Defense database (i.e., Defense Technical Information Center), and emails to relevant authors requesting studies. Approximately 5,800 citations were obtained as a result of this search. Abstracts obtained through this initial search were read to determine appropriate content. Only citations in English and those that actually investigated an organizationally relevant training program were retained. This search reduced the original list to 426 citations. Each of those citations was then reviewed for inclusion in the meta-analysis.

Inclusion Criteria

A number of decision rules were used to determine which articles would be included in the meta-analysis. First, to be included in the meta-analysis, a study must have investigated an evaluation of an organizational training program with interest in how the cognitive ability of the trainees was related to training criterion scores. This can either have occurred in a laboratory or a field setting. Laboratory studies were included if it was judged that the author sought to generalize the findings to organizationally relevant contexts. The focus of this meta-analysis was specifically on the general cognitive ability (i.e., general aptitude) of trainees. This means that studies that investigated specific abilities (e.g., verbal ability, arithmetic reasoning) or knowledge (e.g., language skills, math knowledge) were excluded from the meta-analysis. Second, to be included, studies had to report sample sizes along with a statistic (e.g., r) reflecting the observed relationship between general cognitive ability and a training criterion assessed after completion of training. If a study reported an outcome statistic such as F, t, d, or χ^2 the statistic was converted into an r using the appropriate conversion formulas. Studies that described their results at the group level rather than at the individual level were also excluded. Also, the article must have described enough information to make reasonable judgments about the structure of the training that was implemented. Studies published in educational journals were excluded unless the researchers specifically sought to generalize their findings to organizational settings. Consequently, studies with children as participants were also excluded.

These decision rules reduced the list of eligible studies to 77 (18% retention). The reasons for excluding some studies, listed in the order of majority of occurrence, were as follows: nonorganizational training program, general cognitive ability was not investigated as an individual difference variable, insufficient statistical information, not enough information was provided to make reasonable judgments about the structure of the training program, nonempirical study, and unable to locate or obtain a copy of the study. The final dataset was comprised of the following sources: 44 studies from journal articles, 17 from dissertations, 12 from technical reports, 2 from book chapters, and 2 from conference papers and presentations.

Primary Approach to Analyzing ATIs

Data Set. An initial 77 studies were identified that investigated the general cognitive ability of trainees. Of those, 15 studies investigated the interaction between the general cognitive ability of trainees and the structure of the training program—that is, an ATI was directly investigated.

Nonindependence. The remaining 15 studies produced a dataset of 29 datapoints (*r*s). However, many of the datapoints were considered nonindependent. Datapoints are considered nonindependent if they are computed from the same group of participants. In the current study, many of the datapoints were nonindependent because they involved the same sample with more than one correlation for more than one training criterion. For any analysis that examined overall effects, datapoints were averaged if they came from the same group of participants (i.e., the same study) because they were considered to be nonindependent. For analyses that examined how the effects differed by criterion type, datapoints were considered independent and disaggregated if they came from the same group of participants but were grouped into different criterion categories (e.g., skill versus cognitive criteria) because they were considered to be independent. However, if the datapoints came from the same group of participants and the same criterion category (e.g., skill), the datapoints were considered nonindependent and averaged.

There are multiple reasons why nonindependence is an important consideration in meta-analyses (Arthur, Bennett, & Huffcutt, 2001). The first reason is that nonindependence reduces the observed variability of the computed correlations (i.e., effect size). The second reason is that nonindependence artificially increases the participant sample sizes to make them appear larger than they actually are. The third and final reason is that nonindependence overvalues the contribution of studies so that they contribute to multiple, yet nonindependent, datapoints. To combat these effects it is common practice, when conducting meta-analyses, to aggregate the nonindependent datapoints by calculating their average. Aggregating the nonindependent datapoints resulted in 18 independent datapoints from 15 studies.

Computation of Effect Sizes. Conducting a meta-analysis on the interaction between attributes and treatments (i.e., the actual ATIs provided by 15 of the 77 studies) necessitated a slightly altered approach to the traditional meta-analytic procedures described later. Donovan and Radosevich (1998) have conducted one of the very few meta-analyses on interaction effects and the effect size computations in the current meta-analysis were modeled after their approach. Donovan and Radosevich (1998) used a four-step approach to the computation of effect sizes. First, they obtained the R^2 value associated with the interaction for each individual study. Second, the square root of the R^2 values was taken to obtain semipartial correlations. Third, the sample-weighted semipartial correlations were aggregated and an overall semipartial correlation value was obtained. Finally, the obtained overall semipartial correlation value was squared again to return to an aggregated R^2 value. These same steps were followed in the present investigation.

Secondary Approach to Analyzing ATIs

Data Set. Of the 77 studies that were initially identified, 13 did not report a correlation between general cognitive ability and a training criterion. This left a total of 64 studies that were included in the analysis using the secondary approach.

Nonindependence. The remaining 64 studies produced a dataset of 277 datapoints (*r*s). However, many of the datapoints were considered nonindependent. Aggregating the nonindependent datapoints resulted in 119 independent datapoints from 64 studies.

Outliers. The detection of outliers in meta-analyses is an important practice because failure to do so can increase the residual variability and might therefore distort the mean effect size. Some researchers have noted that essentially all data sets are likely to contain some outliers (e.g., Gulliksen, 1986; Wollins, 1962), leading to the conclusion that also meta-analytic data sets suffer from the presence of outliers. The detection of an outlier in a meta-analytic data set would include inspecting the dataset for any study effect size that appears to be inconsistent with the other study effect sizes.

Reasons for this inconsistency might be methodological (e.g., choice of participants, study design) or errors in the data collection and/or computation. In addition to visually inspecting the data, a procedure developed by Huffcutt and Arthur (1995) was applied to the current dataset to detect potential outliers. Traditional procedures for detecting outliers (e.g., box plot analysis) are not proper procedures for detecting outliers in a meta-analysis because they do not take sample size into account. However, the Sample-Adjusted Meta-Analytic Deviancy (SAMD) procedure developed by Huffcutt and Arthur (1995) adjusts for differing sample sizes. The SAMD procedure "compares the value of each study correlation to the mean sample-weighted correlation without that correlation in the analysis, then adjusts that difference for the sample size of the study" (Arthur, Bennett, & Huffcutt, 2001, p. 119). A SAMD statistic is calculated for each data point in the dataset and each of the absolute SAMD values are rank ordered from highest to lowest. A cut off is then set to investigate those datapoints that are extreme and might be potential outliers.

A SAMD statistic was calculated for each of the 119 independent data points. A separate SAMD statistic was also calculated for each criterion type (i.e., skill, cognitive, and attitudinal as well as declarative knowledge, other cognitive, adaptive, and proceduralized criteria). The mean SAMD values by criterion type ranged from 0.54 (*SD* = 0.40) to 2.21 (*SD* = 3.39). The overall mean SAMD value was 1.04 (*SD* = 1.26). Based on a visual inspection of the data, the decision was made to eliminate any studies with a SAMD value of 6.00 or above. This decision resulted in the elimination of 1 study (1 data point), which had a SAMD value of 8.27 (Oakes, Ferris, Martocchio, Buckley, & Broach, 2001). A review of this outlier suggested that the difference in the

SAMD value of this study was due to a highly increased sample size compared to the sample size of other datapoints included in the dataset (9038 vs. an average of 142). This one outlier constituted about 1% of the 119 *r*s in the dataset. Dropping this outlier resulted in a final dataset of 63 studies and 118 datapoints.

Computation of Effect Sizes. In meta-analysis it is required that a common outcome metric is used. This means that the effects across studies need to be converted to this common outcome metric (Hunter & Schmidt, 2004). The current study used the correlation coefficient (*r*) as the common outcome metric. To analyze the main effects of structure, the effect size Cohen's *d* was calculated and then converted to an *r* if studies provided actual means and standard deviations statistics. For those studies that provided other statistics (i.e., *t*, or univariate two-group *F*) the appropriate conversion formulas were used to calculate the *r* (Arthur et al., 2001). To analyze the relationship between general cognitive ability and training effectiveness by differing degrees of structure and feedback (i.e., moderator analysis) the *r*s, the correlations between general cognitive ability and criteria, provided by each study were separated by their structure score according to the trichotomous schemes mentioned in section "Training Structure" and then meta-analyzed.

Cumulating Effect Sizes Across Studies. For computations for the overall meta-analysis, Arthur et al. (2001) SAS PROC MEANS meta-analysis program was used. Mean sample-size-weighted correlations were calculated using the following formula:

$$\bar{r} = \frac{\Sigma(\mathrm{N}_i * r_i)}{\Sigma N_i}$$

where $\bar{r}$ is the mean sample-size-weighted correlation; N_i the sample size for each study; and r_i the observed correlation. The advantage of sample weighting is that it assigns studies with larger sample sizes more weight than studies with smaller samples.

Correction for Unreliability. As is general practice in any meta-analysis, corrections for predictor and criterion unreliability were made as they both attenuate the observed correlations. To correct for unreliability, estimates of internal consistency were collected (i.e., split-half, coefficient alpha, KR-20) for both the criterion and predictor (i.e., cognitive ability) and a list of attenuating artifact values was compiled from each of the primary studies that provided such information. The values in each distribution of attenuating artifacts were then converted to the appropriate psychometric form by taking the square root of the reported reliability coefficients with the following formulas:

$$a = \sqrt{r_{xx}}$$

$$b = \sqrt{r_{yy}}$$

where r_{xx} is the measurement error in the predictor and r_{yy} measurement error in the criterion. The resulting values were then used to calculate the compound attenuation factor. After the corrections for unreliability, the corrected r (ρ) and standard deviation of the population correlations (SD ρ) were obtained with the following formulas:

$$\rho = \frac{\bar{r}}{\text{AA}}$$

$$SD\ \rho = \sqrt{Var}\,(\rho)$$

where ρ represents the fully corrected correlation; $\bar{r}$ the mean sample-size-weighted correlation; and AA the compound attenuation factor. The percent of the observed variance that is accounted for by sampling error and unreliability was also calculated.

Description of Variables

Training E*valuation Criteria*. To examine how the effects might differ based on the nature of the training criteria, a combination of Kirkpatrick's (1959, 1976, 1996) model (i.e., reaction and results) and Kraiger, Ford, and Salas' (1993) model (i.e., cognitive, skill-based, and attitudinal outcomes) was used to hierarchically code the criterion measures used.

Based on the aforementioned models, an initial criteria hierarchy was formed. The hierarchy consisted of results criteria, cognitive criteria, skill criteria, and attitudinal criteria. The cognitive, skill, and attitudinal criteria were further broken down, such that the cognitive criteria were divided into declarative knowledge, knowledge organization, and cognitive strategies criteria. The skill criteria could be coded along two dimensions: (1) proceduralized versus adaptive criteria and (2) on-the-job and not on-the-job criteria. The skill criteria were coded as adaptive if they were described as a test of transfer (not analogical transfer), adaptability, or generalization, otherwise they were coded as proceduralized. The attitudinal criteria were divided into reaction, attitudinal, and motivational criteria. Furthermore, the reaction criteria were divided into affective and utility criteria, the attitudinal criteria into job-related and organization-related criteria, and the motivational criteria into self-efficacy, goal-setting/commitment, and continuous learning criteria.

Unfortunately, not enough studies were found to warrant the hierarchical criterion break down just described. First, there were no studies found that used a results criterion and thus this type of criterion was eliminated from the analyses. Second, the cognitive criterion breakdown into declarative knowledge, knowledge organization, and cognitive strategies did not yield enough studies to separately analyze these criteria. Because there were enough studies for the declarative knowledge criterion but not for the others, it was decided to break down the cognitive criteria into declarative knowledge and other cognitive criteria. Third, there were not enough studies involving on-the-job performance criteria; thus skill criteria were only coded as either proceduralized or adaptive. Fourth, the attitudinal breakdown into reaction, attitudinal, and motivational criteria, and the further breakdown of these, did not yield enough studies to warrant the breakdown. It was therefore decided to only use the overall, attitudinal criterion and to not break it down any further. Based on these decisions, the final evaluation criterion hierarchy consisted of the following criteria: cognitive (including declarative knowledge and other cognitive criteria), skill (including proceduralized and adaptive criteria), and attitudinal criteria.

Training Structure. The specific methods, procedures, and instructional strategies used to deliver training were coded. Specific attention was paid to the amount of structure involved in the training. The amount of structure present in a training program was coded with seven structure variables, which were developed based on how structure has been defined previously (i.e., Campbell & Kuncel, 2001; Snow, 1989). The variables used to code structure, along with the rating scale used, are shown in Table 1.

Table 1
Structure Variables with Rating Scale

Structure Variable	Rating Scale
Instructor controlled activities	0 – 2
Instructional messages	0 – 2
Clarification of material	0 – 2
Personal assistance from the instructor	0 – 2
Break-down of training into modules	0 – 2
Detailed objectives for trainees	0 – 2
Cognitive strategy intervention	No/Yes*

Note. Structure rating scale: low = 0 (less than 10% of the training), medium = 1 (more than 10% but less than 50% of the training), high = 2 (more than 50% of the training).
*No/Yes rating scale was converted to a 0/2 rating scale.
Cognitive strategy intervention examples: metacognition, elaboration

To arrive at a structure score for each data point, the scores of each variable that measured structure were summed. Possible ratings for each variable ranged from 0 (*not at all* involved in training) to 2 (involved in the *majority* training). The structure variables included (1) instructor controlled activities, (2) frequency of instructional messages, (3) frequency of clarification of material, (4) personal assistance from the instructor, (5) break-down of training into modules, (6) provision of detailed objectives for trainees, and (7) provision of a cognitive strategy intervention. The overall structure scores ranged from 0 to 11. The distribution of structure scores can be seen in Table 2.

Table 2
Structure Variable Frequencies

Structure Score	Frequency
0	7
1	8
2	13
3	18
4	38
5	12
6	30
7	11
8	2
9	7
11	2

After examining the frequency distributions of the structure scores, it was decided to trichotomize the structure variable for the moderator analyses in the following way: low structure = 0 – 2, medium structure = 3 – 5, and high structure = 6 – 11. Following the categorization of structure scores, the studies to be included in the moderator analyses were determined. The decisions to include studies were as follows: (1) studies with one training condition that reported a correlation between general cognitive ability and training effectiveness were included, (2) studies that investigated at least two training conditions but only reported one overall correlation between general cognitive ability and training effectiveness were excluded, (3) studies that investigated at least two training conditions were only included if they reported at least two correlations between general cognitive ability and training effectiveness and these correlations were grouped into the same trichotomized structure category. If a study reported data for at least two training conditions but these training conditions were grouped into a different trichotomized structure category, the study was not included in the moderator analysis. For example, if a study investigated two training conditions and one of these training conditions had an overall structure score of 2 (i.e., low structure) and the other training condition had an overall structure score of 9 (i.e., high structure), the study would have been excluded from the moderator analyses. However, if instead the second training condition had an overall structure score of 1 (i.e., low structure), the study would have been included in the moderator analyses.

Of the initial 63 studies, a total of 51 studies were retained and included in the moderator analyses. Table 3 shows the trichotomized structure frequency distribution of the studies included in the moderator analysis.

Table 3
Structure Category Frequencies of Studies Included in the Moderator Analysis

Structure Category	Frequency of Studies Included
Low	13
Medium	28
High	10

Specific attention was also paid to another method used to deliver training: the amount of feedback involved in the training. However, due to several limitations, details on the categorization of the feedback variable and various results related to this variable (e.g., main effects) are presented in the Appendix. The limitations were three-fold: (1) studies only mentioned feedback if feedback was the focus of the document—it could therefore be rarely coded—which led to (2) the distribution of feedback scores not being normal (see Table 11 in the Appendix), and (3) the number of studies found for an analysis of the feedback main effect was rather small.

Training Condition Comparisons. To determine the main effects for structure, 17 studies were identified to be included in the analysis. To be included a study had to report data (i.e., *M*s and *SD*s, or *d*s, or *t*s) on at least two training conditions that each had a structure score that was placed into a separate trichotomized structure category (i.e., either low, medium, or high structure).

After the identification of studies, the effect size Cohen's *d* was calculated for each study and any study that had a *d* greater than 2 was eliminated because it was considered extreme. This resulted in the elimination of one study, which had a *d* of 3.76 (Simon & Werner, 1996). This reduced the number of studies for the analysis of the structure main effect to 16. The *d*s were then converted into the common outcome metric *r* and meta-analyzed. The main effect analysis was further broken down by cognitive and skill evaluation criteria. A continued breakdown (e.g., into declarative knowledge) was not warranted because of the number of studies available.

Coding Reliability

Four graduate students coded the data reported in this meta-analysis. Training for the coders began by each receiving a copy of a training manual that was developed by the author of this document. The manual was used by the coders to code three practice articles on their own. After practice the group of coders reconvened to discuss issues that arose with either the coding sheet and/or the coding manual. Accordingly, changes were made to fine-tune the coding sheet and manual and reduce coding inconsistencies. About half of the studies included in the meta-analysis were coded individually and meetings with the coders were set to discuss discrepancies and disagreements. At this point, each of the four coders was assigned a common set of 10 articles to determine the degree of interrater agreement. Interrater agreement was determined by comparing the values that each coder assigned to each of the variables of interest.

The agreement between raters was generally high. If discrepancies existed, they were resolved through consensus meetings. Across the common set of 10 articles, the evaluation criteria were coded with 89% agreement. Ratings for the structure variable yielded an ICC of 0.82.

Results

The results section of this document is divided into four parts. The first part discusses the results with respect to the state of ATI research in the training literature. The second part addresses the question if ATIs exist in the training literature. The third part investigates the nature of these ATIs. The fourth part discusses the size of the ATIs. Finally, the last part describes ancillary analyses that were conducted.

General Account of ATI Research in the Training Literature

The first issue that was investigated was the state of the training literature with regard to the amount of ATI research that is performed. A count was performed of studies that investigated an ATI directly and also a count of the number of studies that could have investigated an ATI. Studies that could have investigated an ATI were considered as such if they investigated the relevant variables, such as the general cognitive ability of trainees, described a training program, and had at least two training conditions. Almost 50% of the studies that could have investigated an ATI actually did. More specifically, of the 31 studies that would have been able to investigate an ATI, 15 did investigate an ATI while 16 did not.

Do ATIs Exist in the Training Research Literature?

The first research objective was to determine if ATIs exist in the training research literature—that is, does the correlation between general cognitive ability and training effectiveness differ as a function of structure? To determine the existence of ATIs two approaches were chosen.

Primary Approach. As mentioned before, the primary approach constituted a meta-analysis of studies that reported an interaction between general cognitive ability and training program structure. As previously mentioned, a total of 15 studies were found that investigated an ATI directly. These studies either reported the ATI as a ΔR^2 or an F, both of which were converted to the common outcome metric r with the appropriated formulas. Following the approach of Donovan and Radosevich (1998) these rs were then meta-analyzed and finally reconverted to an ΔR^2. As can be seen in Table 4, across different criteria, a weak ATI was present (ΔR^2 = 0.012).

Table 4
ATI Results—Overall and by Cognitive and Skill Criteria

Structure	k	N	ΔR^2	Range of ΔR^2	r	Range of rs
Overall	15	2371	0.012	0.000 – 0.123	0.11	0.00 – 0.35
Cognitive	10	1083	0.029	0.000 – 0.123	0.17	0.01 – 0.35
Declarative knowledge	8	624	0.058	0.010 – 0.123	0.24	0.10 – 0.35
Other	2	459	0.008	0.001 – 0.010	0.09	0.01 – 0.10
Skill	7	1463	0.005	0.000 – 0.048	0.07	0.00 – 0.22
Proceduralized	6	1354	0.005	0.000 – 0.048	0.07	0.00 – 0.22

Note. k = number of observations (i.e., studies). N = number of participants. ΔR^2 = sample-weighted mean squared correlation coefficient indicating the value associated with the additional variance explained by the interaction term. r = square root of ΔR^2.

When examining the ATI effect by cognitive and skill criteria, it can be observed that the interaction effect is larger for cognitive criteria ($\Delta R^2 = 0.029$) than for skill criteria ($\Delta R^2 = 0.005$). The largest effect was yielded for declarative knowledge criteria ($\Delta R^2 = 0.058$). According to this approach, ATIs with regard to cognitive criteria and skill criteria exist in the training literature but their size is small and varies depending on the criteria with which training effectiveness is evaluated. Unfortunately, no studies were found to warrant an analysis with regard to attitude criteria.

Secondary Approach. The secondary approach consisted of coding the structure of training programs described in studies that reported correlations between general cognitive ability and training evaluation criterion scores and subsequently comparing the meta-analytic effects across differing levels of structure. As previously mentioned, a total of 63 studies were found that provided a correlation between the general cognitive ability of trainees and the effectiveness of the training program. After grouping each of these studies into the trichotomized structure categories (i.e., low, medium, and high), and determining which studies would meet the inclusion criteria for the moderator analyses (i.e., different training conditions had to be grouped into the same structure category), a total of 51 studies were included in this approach.

Two statistics were of special importance when determining if the correlations between general cognitive ability and training effectiveness differ as a function of structure. The first one was the sample-weighted mean *r* (*swmr*), an aggregated correlation coefficient, which is an average of all correlation coefficients adjusted for sample size. It was decided to rather examine the *swmr* instead of the corrected sample-weighted mean *r* (ρ) because the *swmr* is a better reflection of the correlation statistic that primary studies report. That is, primary studies do not usually correct the reported correlations for unreliability, which is what ρ is adjusting for. To determine if the correlation between general cognitive ability and training effectiveness varies as a function of structure, the differences in the *swmr*s across the different structure categories (i.e., low, medium, high) were examined. The greater the difference between the correlations, the more likely a meaningful difference exists between the correlations as a function of structure.

The second statistic that was used to determine if a meaningful difference exists between the correlations as a function of structure was the confidence interval. In this study a confidence interval of 90% was chosen. The confidence interval assesses the accuracy of the estimate of the *swmr*—it estimates to which degree sampling error remains in the *swmr* estimate. The confidence interval provides a range of values in which the *swmr* is likely to fall if other sets of studies were to be meta-analyzed (Arthur et al., 2001). The smaller the confidence interval the less sampling error the estimate contains; the larger the confidence interval the more sampling error it contains. If the confidence interval includes zero, the *swmr* estimate is not considered to be reliable—if it does not include zero, the *swmr* estimate is considered a reliable estimate. For the current meta-analysis it was decided that if the confidence intervals of the *swmr*s overlap no meaningful difference would exist between them, because the *swmr*s might not be as different as it would be suggested if one only examined the *swmr*s without taking the confidence interval into account. However, if the confidence intervals do not overlap a meaningful difference between the correlations was said to exist. For example, if the confidence intervals for the correlation between general cognitive ability and training effectiveness for high structure and low structure were 0.22 – 0.30 and 0.24 – 0.33 respectively, they would overlap and a meaningful difference between the two correlations would not exist. If however, the confidence intervals were 0.22 – 0.30 and 0.31 – 0.39 respectively, they would not overlap and a meaningful difference between the correlations would exist. For this study, two *swmr*s were considered meaningfully different (i.e., differences in effects are reliable) if the upper bound of the 90% confidence interval for the smaller *swmr* was equal to or less than the lower bound of the 90% confidence interval for the larger *swmr*.

Another consideration to examining the existence of ATIs is to determine the presence of moderators. To determine the existence and operation of moderators, Schmidt and Hunter's (1977) 75% rule was applied. This rules states that if 75% or more of the variance can be explained by the corrections made (i.e., sampling error and/or unreliability), then it can be concluded that all of the variance is due to artifacts because the remaining percentage is most likely due to uncorrected artifacts (e.g., computational errors). If however, less than 75% of the variance is explained by the corrections made, moderator variables are operating. In the present study, the moderators examined were training structure and evaluation criteria.

Table 5 shows the meta-analytic results for the relationship between general cognitive ability and training effectiveness by structure nested with criteria. The minimum number of studies to meaningfully interpret results was set at 5; however, for the sake of completeness, results are shown in Table 5 even if the number of studies fell below 5.

Table 5
Relationship between General Cognitive Ability and Training Effectiveness by Structure Nested within Criteria

				90% CI					90% CV	
Criteria/structure	*k*	*N*	*r*	L	U	ρ	*SD* ρ	% Var. explained	L	U
Overall	51	7559	0.28	0.26	0.30	0.37	0.14	35.41	0.14	0.59
Structure										
High	10	1439	0.31	0.27	0.35	0.40	0.14	33.87	0.17	0.63
Medium	28	4946	0.32	0.30	0.34	0.42	0.25	12.46	0.01	0.84
Low	13	1174	0.27	0.23	0.32	0.33	0.13	46.60	0.13	0.54
Cognitive										
High structure	7	1091	0.34	0.29	0.38	0.42	0.10	46.56	0.26	0.58
Medium structure	14	2806	0.29	0.26	0.32	0.38	0.17	24.37	0.11	0.65
Low structure	5	540	0.21	0.14	0.27	0.25	0.00	100.00	0.25	0.25
Declarative knowledge										
High structure	6	1034	0.36	0.36	0.36	0.40	0.00	100.00	0.40	0.40
Medium structure	12	2670	0.28	0.25	0.31	0.39	0.16	22.07	0.12	0.65
Low structure	4	488	0.20	0.13	0.27	0.24	0.00	100.00	0.24	0.24
Other										
High structure	2	438	0.34	0.27	0.41	0.42	0.00	100.00	0.42	0.42
Medium structure	4	884	0.34	0.29	0.39	0.38	0.30	10.62	0.06	0.70
Skill										
High structure	7	1145	0.24	0.19	0.29	0.34	0.20	22.55	0.02	0.66
Medium structure	22	3824	0.27	0.25	0.30	0.37	0.16	27.34	0.10	0.63
Low structure	10	807	0.34	0.29	0.39	0.42	0.15	41.00	0.18	0.66
Proceduralized										
High structure	5	1029	0.21	0.16	0.26	0.31	0.17	24.61	0.03	0.59
Medium structure	18	3063	0.25	0.22	0.28	0.33	0.16	27.56	0.07	0.59
Low structure	7	593	0.32	0.26	0.38	0.37	0.12	46.15	0.17	0.56
Adaptive										
High structure	2	116	0.50	0.50	0.50	0.50[1]	0.00	100.00	0.50	0.50
Medium structure	7	1037	0.35	0.31	0.40	0.46	0.09	50.60	0.31	0.62
Low structure	3	214	0.40	0.40	0.40	0.48	0.00	100.00	0.48	0.48

Table 5 Continued
Relationship between Cognitive Ability and Training Effectiveness by Structure Nested within Criteria

				90% CI					90% CV	
Criteria/structure	*k*	*N*	*r*	L	U	ρ	*SD* ρ	% Var. explained	L	U
Attitude										
Medium structure	6	488	0.11	0.04	0.18	0.15	0.00	100.00	0.15	0.15
Low structure	3	272	0.08	0.08	0.08	0.09	0.00	100.00	0.09	0.09

Note. *k* = number of observations (i.e., studies). *N* = number of participants. *r* = sample-weighted mean correlation. CI = confidence interval (L = lower, U = upper). ρ = estimated true validity after correcting for criterion and predictor unreliability. *SD* ρ = standard deviation of the estimated true validity. % Var. explained = % variance explained by sampling error and criterion unreliability. CV = credibility value (L = lower, U = upper). [1] = ρ and *r* are the same because the reliability estimates for both the criterion and predictor were not reported.

When examining Table 5, it is notable that the relationship between general cognitive ability and training effectiveness is positive (*swmr* = 0.28; ρ = 0.37), signifying that the higher the general cognitive ability of trainees, the better their performance across multiple training evaluation criteria. It is also noteworthy that all credibility values are positive and do not include zero, indicating that the relationship between general cognitive ability and training effectiveness could not be negative, as is suggested by the model in Figure 1, which shows a negative correlation between general cognitive ability and training effectiveness for training programs with high structure. Based on the credibility values shown in Table 5, it can be said that it is rare to find a negative correlation between general cognitive ability and training effectiveness in a single primary study—the positive correlation can therefore be generalized across training contents and contexts.

When examining the relationship between general cognitive ability and training effectiveness in Table 5, it is noticeable that the percent variance explained suggests the existence of moderators, such that in some contexts the relationship is more positive than in other contexts. This justifies examining structure as a moderator to determine if the overall positive relationship is different depending on the amount of structure in a training program. Consequently, the amount of structure was investigated as a potential moderator. As can be seen in Table 5, small differences exist in the *swmr*s between high, medium, and low structure (0.31, 0.32, and 0.27, respectively). Moreover, it can also be seen that the confidence intervals of those *swmr*s overlap to a great degree, suggesting that the correlations between general cognitive ability and training program effectiveness do not meaningfully vary as a function of structure. However, when investigating the percent variance explained, it is apparent that additional potential moderators might be operating. This justifies the examination of an additional moderator. Consequently, the additional moderator that was investigated was the type of criterion with which training effectiveness was measured.

Examining the results for cognitive criteria in Table 5, it can be seen that the *swmr*s have larger differences (0.34, 0.29, and 0.21, respectively), with the greatest difference between the high and low structure *swmr*s (0.34 and 0.21, respectively). Furthermore, the confidence intervals of the high and low *swmr*s do not overlap (high structure = 0.29 – 0.38 and low structure: 0.14 – 0. 27), suggesting a meaningful difference between the two correlations. These results suggest that structure increases the effects of general cognitive ability differences that existed between trainees at the outset of the training program. These results are consistent with the model presented in

Figure 3. Yet, when examining the percent variance explained, it is noticeable that they still suggest the operation of additional moderators (although the percent variance explained for the low structure variable has already increased to 100%). To examine this, the cognitive criteria were further divided into declarative knowledge criteria and other cognitive criteria (the latter did not yield enough studies to warrant a meaningful interpretation). The results for the declarative knowledge criteria indicate the same trend found for the cognitive criteria: trainees with higher general cognitive ability appear to benefit more from training programs with high structure than trainees with lower general cognitive ability. Additionally, it is noticeable that the percent variance explained increased substantially when declarative knowledge was investigated as a moderator. The percent variance explained for the high and low structure relationship reached 100%, indicating that no additional moderators are operating. Taken together, these results suggest small but robust differences between general cognitive ability and cognitive criteria as function of high versus low training structure. In other words, these results reflect an ATI with respect to cognitive criteria.

Examining Table 5 with regard to skill criteria, the *swmr*s for the various amounts of structure also exhibit rather large differences between them (0.24, 0.27, and 0.34, respectively). The largest difference, similar to cognitive criteria, is again that between high and low structure (0.24 and 0.34, respectively). However, for the skill criteria, the relationship between general cognitive ability and training effectiveness is larger for low structured training programs than for high structured training programs—this result stands in complete opposition to that found for the cognitive criteria. The confidence intervals of the *swmr*s do not overlap (high structure = 0.19 –

0.29 and low structure = 0.29 – 0. 39), implying a meaningful difference between the correlations. While results for the cognitive criteria were consistent with the model presented in Figure 3, the results for the skill criteria are consistent with the model presented in Figure 2. These results suggest that structure decreases the effects of general cognitive ability differences that existed between trainees at the outset of the training program. Similarly to the cognitive criteria, the percent variance explained for the skill criteria, also suggests the presence and operation of further moderators. Consequently, the skill criteria were further divided into proceduralized criteria and adaptive criteria (the latter did not produce enough studies to warrant a meaningful interpretation). When examining Table 5, it can be seen that the results for the proceduralized criteria are consistent with the trend observed for the skill criteria. Additionally, it is noticeable that the percent variance explained barely increased when proceduralized skill was investigated as a moderator. The percent variance explained is still far below the 75% that was suggested by Schmidt and Hunter (1977) as a minimum for excluding the existence of moderators. This suggests that the relationship between general cognitive ability and proceduralized skill criteria varies as function of some unknown moderator.

Although the number of studies for the attitude criteria is not large enough to warrant a meaningful interpretation, the results are a shortly mentioned. The results for the attitude criteria, as presented in Table 5, are less clear than those for the cognitive and skill criteria. Even though there is a difference between the two *swmr*s, the confidence intervals overlap to a great degree, suggesting no meaningful difference between the two correlations. Furthermore, it is worth noting that the *swmr*s are smaller than those for the cognitive and skill criteria.

To summarize, from the presented results, it appears as if ATIs exist in organizationally-relevant training but that their presence depends on the criteria with which training effectiveness is evaluated. That is, with regard to attitudinal criteria there do not appear to exist meaningful ATIs but with regard to cognitive and skill criteria their existence is more apparent.

What is the Nature of the Interaction?

Primary Approach. In order to determine the nature of ATIs with the primary approach it was necessary to examine the nature of the interaction of the studies that directly investigated ATIs. Unfortunately, only about 2/3 of the studies that directly investigated an ATI provided the necessary variables/figures to determine the nature of the interaction. Of those studies that provided the necessary information, more than half found an ATI that was in support of the model in Figure 3. That is, the correlation between general cognitive ability and training effectiveness is greater for training programs with high structure than for those with low structure—trainees with high general cognitive ability tend to benefit more from structured training programs than trainees with low general cognitive ability. A little less than half of the studies supported the model in Figure 2 (i.e., trainees with low general cognitive ability tend to benefit more from highly structured training programs than trainees with high general cognitive ability).

Secondary Approach. To determine the nature of the interactions, it was first necessary to examine the main effect of structure to determine if providing structure in a training program is helpful to trainees in general. Combining the main effects of structure with the differential effects yielded by general cognitive ability as a function of structure as previously shown (Table 5) provides the information necessary to fully describe and plot the nature of the interactions occurring. Table 6 shows the main effect results for the structure variable. Because not enough studies were found to warrant analyses for all possible structure combinations (i.e., high vs. low structure, high vs. medium structure, and medium vs. low structure), some combinations are not displayed.

Table 6
Main Effect Results of Structure by Cognitive and Skill Criteria

				90% CI					90% CV	
Structure	*k*	*N*	*r*	L	U	ρ	*SD* ρ	% Var. explained	L	U
Overall	16	1094	0.13	0.08	0.18	0.17	0.21	37.28	-0.17	0.51
High vs. Medium	6	516	0.08	0.01	0.15	0.11	0.24	28.01	-0.29	0.49
Medium vs. Low	10	578	0.17	0.10	0.24	0.23	0.17	51.16	-0.05	0.51
Cognitive	11	610	0.11	0.04	0.18	0.17	0.44	18.67	-0.55	0.88
High vs. Medium	4	188	0.12	0.00	0.24	0.19	0.54	16.57	-0.69	1.00
Medium vs. Low	7	422	0.10	0.03	0.18	0.15	0.37	19.91	-0.46	0.76
Skill	8	604	0.11	0.11	0.11	0.12	0.00	100.00	0.12	0.12
High vs. Medium	3	375	0.09	0.09	0.09	0.09[1]	0.00	100.00	0.09	0.09
Medium vs. Low	5	229	0.14	0.14	0.14	0.16	0.00	100.00	0.16	0.16

Note. *k* = number of observations (i.e., studies). *N* = number of participants. *r* = sample-weighted mean correlation. CI = confidence interval (L = lower, U = upper). ρ = estimated true validity after correcting for criterion and predictor unreliability. *SD* ρ = standard deviation of the estimated true validity. % Var. explained = % variance explained by sampling error and criterion unreliability. CV = credibility value (L = lower, U = upper). [1] = ρ and 90% CV are the same as *r* because the reliability estimates for both the criterion and predictor were not reported.

Table 6 shows that the effect of structure on training effectiveness is small but positive (*swmr* = 0.13, ρ = 0.17), suggesting that increased structure benefits trainees in general. However, because the credibility values include negative numbers, a high confidence in that higher structure always benefits trainees is not justified. The fact that the credibility values include negative values suggests that the effect of structure on training effectiveness could occasionally be negative, such that trainees benefit more from less structured training programs.

When examining the effect of structure on training effectiveness in Table 6, it can be seen that the percent variance explained indicates the operation of potential moderators. A potential moderator that was examined, was the different types of structure (i.e., high vs. medium and medium vs. low structure). These results are also consistent with the result that increased structure benefits trainees: the correlations are small but positive (*swmr* = 0.08 for the high vs. medium comparison and *swmr* = 0.17 for the medium v. low structure comparison). Yet, the existence and operation of moderators is still suggested by the percent variance explained, and so a further moderator was examined—the type of criterion with which training effectiveness was measured.

Table 6 shows that with regard to cognitive criteria, the *swmr*s for the various amounts of structure also support the benefit of increased structure for trainees. The effects are small but positive (*swmr* = 0.12 for the high vs. medium comparison and *swmr* = 0.10 for the medium v. low structure comparison). The same result is displayed for the skill criteria (*swmr* = 0.09 for the high vs. medium comparison and *swmr* = 0.14 for the medium v. low structure comparison). The percent variance explained, however, differs for the cognitive and skill criteria. For the cognitive criteria the percent variance explained is very low, indicating the operation of additional moderators while for skill criteria the percent variance explained is 100%, indicating that no additional moderators influence the effect of structure on training effectiveness. Additionally, it is noticeable that the credibility intervals for the cognitive criteria include negative numbers. This suggests that the effect of structure on training effectiveness could also be negative, such that low structure benefits trainees when the evaluation criteria are cognitive in nature. On the other hand, the credibility intervals for the skill criteria do no include zero, suggesting that the effect of structure on training effectiveness is always positive when training effectiveness is measured with skill criteria.

To determine the nature of the different ATIs found for the cognitive-based and skill-based evaluation criteria figures were created such that the main effects of structure were plotted with the different slopes for low versus high structure. Unfortunately, the attitude criteria did not yield enough studies for either the analysis of the main effect or the graphing of a figure. In the figures, the regression lines were based on the *swmr*s for the high and low structure results. To graph the main effects the *swmr*s for cognitive criteria (*swmr* = 0.11) and skill criteria (*swmr* = 0.11) were converted to Cohen's *d*s and then plotted at the middle of the x-axis representing the general cognitive ability distribution. Next, the slope of the effects for high and low structure, as reflected by the *swmr*s in Table 4, was graphed for each type of criterion (i.e., cognitive and skill). Although the graphs show datapoints at –2 standard deviations and +2 standard deviations, the full range of the general cognitive ability distribution in the population is not represented due to the nature of participants investigated in the primary studies included in this meta-analysis: it is very doubtful that the studies meta-analyzed here included participants at the extreme low end of the general cognitive ability distribution (i.e., mental retardation). Thus, the figures should be regarded as hypothetical in nature.

Figure 5 shows that high structure generally benefits trainees when their training effectiveness is measured with cognitive criteria. Table 6 mirrors those results: the effect of structure on training effectiveness is positive for cognitive criteria (*swmr* = 0.11).

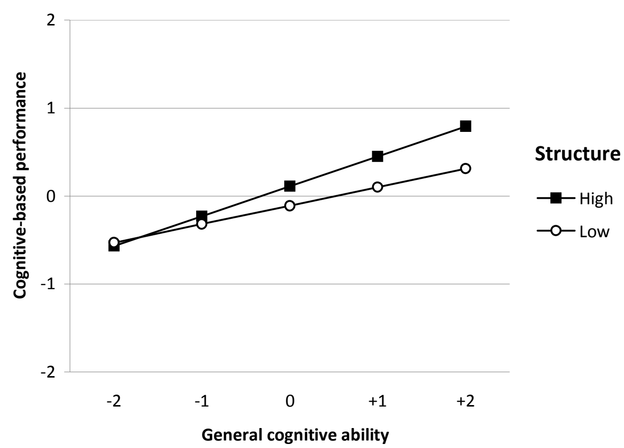

Figure 5. ATI for Cognitive Criteria. The values on the axes reflect standard deviations.

However, Figure 5 also shows that the effects of structure are stronger for trainees who are at the higher ends of the general cognitive ability distribution, lending support to the model presented in Figure 3. As can be seen in Figure 5, the effects become stronger at one standard deviation above the mean of general cognitive ability (*d* = 0.45) and are strongest at two standard deviations (*d* = 0.79). This means that trainees high in general cognitive ability benefit from a training program high in structure while trainees low in general cognitive ability do not. The performance of trainees with low general cognitive ability does not increase to a great degree, if at all, with the provision of more structure in the training program. In general, the benefit of structure is observed for trainees who are in the top half of the general cognitive ability distribution while less benefit is observed for trainees on the lower half of the general cognitive ability distribution when training is evaluated with cognitive criteria.

Figure 6 also shows that high structure generally benefits trainees when their training effectiveness is measured with skill criteria and Table 6 mirrors this result with a small, but positive effect of structure on training effectiveness (*swmr* = 0.11).

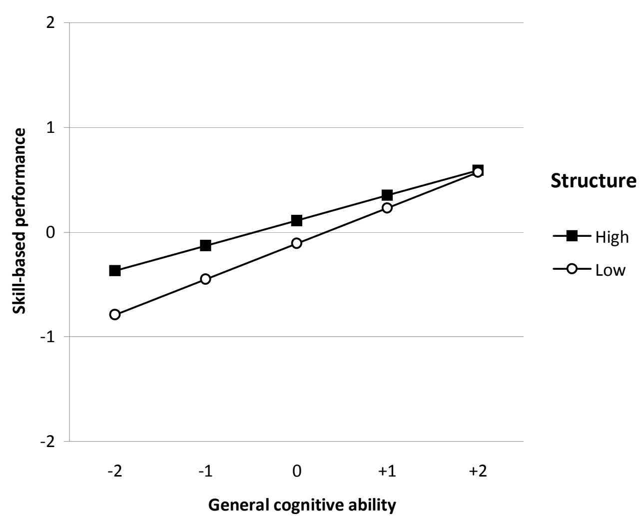

Figure 6. ATI for Skill Criteria. The values on the axes reflect standard deviations.

However, when training effectiveness is measured with skill criteria, the nature of the interaction is reversed. Figure 6 shows that the effects of structure are stronger for trainees who are at the lower ends of the general cognitive ability distribution, lending support to the model presented in Figure 2. The effects become stronger at one standard deviation below the mean of general cognitive ability (d = -0.13) and are strongest at two standard deviations (d = -0.37). At the extreme high end of the general cognitive ability continuum, trainees perform equally well in training programs with either high or low structure, while trainees with a general cognitive ability of one standard deviation above the mean still benefit, to a small degree, from training programs with high structure. Trainees with low general cognitive ability on the other hand, genuinely benefit from a training program high in structure as compared to a training program low in structure. Thus, the benefits of structure are observed for trainees in the lower half of the general cognitive ability distribution while less benefits are observed for trainees in the top half of the general cognitive ability distribution when training is evaluated with skill criteria.

In summary, increasing the structure in a training program generally appears to benefit trainees. However, the type of criterion with which training effectiveness is measured and the general cognitive ability of trainees alters who benefits from high structured training programs. When trainees' performance in a training program is measured with cognitive criteria, trainees high in general cognitive ability seem to benefit from training programs high in structure while for trainees low in general cognitive ability the structure of the training program does not matter. However, when trainees' performance is measured with skill criteria, trainees with low general cognitive ability seem to benefit from training programs high in structure, while trainees high in general cognitive ability also benefit, but to a much smaller degree.

What is the Size of the Interaction Effect?

The third research question focused on the size of the interaction effects. Generally, small ATIs were found across the two different approaches described. These small effects can be observed in Figures 5 and 6, and Table 4 (i.e., $\Delta R^2 = 0.012$), which shows the results for studies that directly investigated ATIs. To further determine the size of the effects, and also to differentiate the interaction effect sizes by the different types of evaluation criteria, the *q* statistic was computed (Cohen, 1992). The *q* statistic was used to solely investigate the interaction effect sizes for the results of the secondary approach. The *q* statistic calculates the size of the effect, in this case by examining the difference between the correlations for high and low structured training programs (except for the attitudinal criterion, where the difference was calculated between the medium and low structured training program). To calculate the *q* statistic, the various correlations were first converted to Fisher *z*s and then the absolute difference between the *z*s was determined. Just as for an *r*, small, medium, and large effects are represented by a *q* statistic of 0.1, 0.3, and 0.5, respectively. Table 7 shows the size of the effects for the overall result and the size of the effects divided by the three types of evaluation criteria. The cognitive and skill criteria were further divided into declarative knowledge and proceduralized criteria, to reflect the results reported previously.

Table 7
Size of ATIs

	ATI Analysis Approach	
Criteria/structure	Primary Approach *r*	Secondary Approach *q*
Structure Overall	0.11	0.04
Cognitive	0.17	0.14
Declarative knowledge	0.24	0.17
Skill	0.07	0.11
Proceduralized	0.07	0.12
Attitude	N/A	0.03

Table 7 shows that the ATI effects were small, as already suggested by Figures 5 and 6 and Table 4. Moreover, the overall size of the ATIs in Table 6 is comparable to those ATIs presented in Table 4 and the tendency that stronger effects are found for cognitive than for skill criteria (Table 5) is also consistent with the results shown in Table 4. This lends support to analyzing the existence of ATIs with the secondary approach used in this study. As previously mentioned, the interaction effect sizes differed depending on the type of evaluation criteria. The stronger interactions were found for the cognitive evaluation criteria ($r = 0.17$; $q = 0.14$)—especially for the declarative knowledge criteria ($r = 0.24$; $q = 0.17$). The interactions for the skill evaluation criteria were lower ($r = 0.07$; $q = 0.11$).

Ancillary Analyses: Does the Content of Training Matter?

Do the moderators that were investigated covary with other extraneous variables? With this question in mind, the found effects were reviewed for other variables that might drive the effects other than the structure of the training program and the evaluation criteria. The variable that was investigated was the content (i.e., the issues of focus and/or tasks to be learned) of the training program because it was thought to covary in particular with the evaluation criteria.

To address this possibility, the content of the training program was coded. More specifically, the training content was coded for its amount of cognitive, physical, and interpersonal demands with the following scale: 0 = not applicable, 1 = somewhat applicable, and 2 = very applicable. The effects of the three types of training content were investigated separately and as their sum (CPI), the latter forming a score of complexity of the training content.

Table 8 shows the means and standard deviations of each training content demand by the different types of training evaluation criteria and levels of training structure.

Table 8
Training Content by Training Evaluation Criteria and Training Structure

	Training content scores							
	Cognitive		Physical		Interpersonal		CPI	
Criteria/structure	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Cognitive	1.73	0.45	0.54	0.86	0.38	0.70	2.65	1.06
High structure	1.57	0.53	0.43	0.79	0.29	0.76	2.29	0.76
Medium structure	1.79	0.43	0.71	0.99	0.43	0.65	2.93	1.27
Low structure	1.80	0.45	0.20	0.45	0.40	0.89	2.14	0.55
Skill	1.74	0.44	0.69	0.89	0.33	0.66	2.87	1.20
High structure	1.29	0.49	0.71	0.95	0.86	1.07	2.86	1.57
Medium structure	1.86	0.35	0.77	0.97	0.27	0.55	3.09	1.15
Low structure	1.80	0.42	0.50	0.71	0.10	0.32	2.40	0.97

Note. Training content scoring scale: 0 = not applicable, 1 = somewhat applicable, 2 = very applicable. CPI = sum of cognitive, physical, and interpersonal scores.

As can be seen in Table 8, there are some small differences in the means of the cognitive training content scores as well as the physical training content scores across criteria and structure. However, it can be seen that for the cognitive evaluation criteria and high structure the mean cognitive training content score is different from that for the skill evaluation criteria and high structure whereby the mean cognitive evaluation criteria and high structure score is higher than the mean skill training evaluation criteria and high structure score (d = 0.55). For the cognitive evaluation criteria and high structure the mean physical training content score however is smaller than the mean score for the physical training content for the skill evaluation criteria and high structure (d = -0.32). This same trend is observable for the cognitive evaluation criteria and low structure and the skill evaluation criteria and low structure (d = -0.51) for the physical training content. Differences are also noticeable for the interpersonal training content and the CPI scores. Table 8 shows that the mean interpersonal training score for the cognitive evaluation criteria and high structure is lower than the mean interpersonal training score for the skill evaluation criteria and high structure (*d* = -0.61) while for the cognitive evaluation criteria and low structure it is higher than the mean interpersonal training score for the skill evaluation criteria and low structure (*d* = 0.45). Finally, the mean CPI score for the cognitive evaluation criteria and high structure is smaller than the mean CPI score for the skill evaluation criteria and high structure (*d* = -0.46) and the mean CPI score for the cognitive evaluation criteria and low structure is smaller than the mean CPI score for the skill evaluation criteria and low structure (*d* = -0.33).

These results suggest that the evaluation criteria and the training content variables indeed covary. Therefore, a moderator analysis of the training content was conducted to determine if the relationship between general cognitive ability and training effectiveness varies as a function of the training content. Table 9 shows the results of this analysis.

Table 9
Relationship between General Cognitive Ability and Training Effectiveness by Training Content

Training Content	*k*	*N*	*r*	90% CI L	90% CI U	ρ	*SD* ρ	% Var. explained	90% CV L	90% CV U
Cognitive										
High	50	7111	0.28	0.26	0.29	0.35	0.12	42.46	0.15	0.55
Medium	13	1969	0.28	0.25	0.31	0.37	0.18	25.28	0.08	0.67
Physical										
High	13	1986	0.26	0.12	0.30	0.36	0.16	30.40	0.10	0.63
Medium	9	1700	0.29	0.26	0.33	0.33	0.10	36.80	0.27	0.49
Low	41	5394	0.28	0.26	0.30	0.31	0.12	36.24	0.11	0.51
Interpersonal										
High	6	1249	0.16	0.16	0.16	0.17	0.00	100.00	0.16	0.16
Medium	7	1442	0.25	0.21	0.29	0.33	0.07	62.82	0.22	0.44
Low	50	6489	0.30	0.28	0.32	0.39	0.15	36.29	0.15	0.63
CPI										
High	3	564	0.17	0.17	0.17	0.17[1]	0.17	100.00	0.17	0.17
Medium	27	4681	0.27	0.25	0.29	0.33	0.12	34.15	0.13	0.53
Low	33	3755	0.30	0.27	0.32	0.40	0.15	38.83	0.15	0.64

Note. Training content scoring scale: low = 0 (not applicable), medium = 1 (somewhat applicable), high = 2 (very applicable). CPI = sum of cognitive, physical, and interpersonal scores. CPI scoring scale: low = 1 – 2, medium = 3 – 4, high = 5 – 6. *k* = number of observations (i.e., studies). *N* = number of participants. *r* = sample-weighted mean correlation. CI = confidence interval (L = lower, U = upper). ρ = estimated true validity after correcting for criterion and predictor unreliability. *SD* ρ = standard deviation of the estimated true validity. % Var. explained = % variance explained by sampling error and criterion unreliability. CV = credibility value (L = lower, U = upper). [1] = ρ and *r* are the same because the reliability estimates for both the criterion and predictor were not reported.

When examining Table 9 it can be seen that the *swmr*s do not differ by either the cognitive or the physical training content and the size of the *swmr*s is comparable across the cognitive and physical training contents. However, the *swmr*s meaningfully differ as a function of interpersonal training content as well as CPI scores. This indicates that the relationship between general cognitive ability and training effectiveness differs as a function of the training content, at least for those types of trainings that are interpersonal in nature or have a higher complexity. This supports the contention of training content having an extraneous influence and therefore the moderator analyses presented in Table 5 were re-conducted with the six studies excluded (Baldwin, Magjuka, & Loher, 1991; Bramble, 1993; Dean, Conte, & Blankenhorn, 2006; Knerr, Harris, O'Brien, Sticha, & Goldberg, 1984; Lievens, Harris, Van Keer, & Bisqueret, 2003; Tziner & Dolan, 1982) which were shown in Table 8 to influence the relationship between general cognitive ability and training effectiveness. The studies that were identified with a high CPI score were part of the six studies that were identified with a high interpersonal training content score. Therefore, the CPI differences shown in Table 9 were solely driven by the interpersonal training content.

Table 10 shows the relationship between general cognitive ability and training effectiveness by structure nested within criteria removed from the influence of training content.

Table 10
Relationship between General Cognitive Ability and Training Effectiveness by Structure Nested within Criteria removed from the Influence of Training Content

				90% CI					90% CV	
Criteria/structure	*k*	*N*	*r*	L	U	ρ	*SD* ρ	% Var. explained	L	U
Overall	45	6410	0.30	0.28	0.32	0.40	0.14	35.78	0.17	0.62
Structure										
High	7	998	0.38	0.34	0.42	0.49	0.07	62.74	0.37	0.61
Medium	27	4546	0.33	0.31	0.35	0.44	0.26	12.30	0.01	0.86
Low	12	967	0.31	0.26	0.35	0.38	0.12	53.51	0.18	0.57

Table 10 continued
Relationship between General Cognitive Ability and Training Effectiveness by Structure Nested within Criteria removed from the Influence of Training Content

				90% CI					90% CV	
Criteria/structure	*k*	*N*	*r*	L	U	ρ	*SD* ρ	% Var. explained	L	U
Cognitive										
High structure	6	791	0.37	0.32	0.42	0.46	0.09	50.99	0.31	0.61
Medium structure	13	2406	0.31	0.28	0.34	0.40	0.17	24.33	0.12	0.68
Low structure	4	333	0.24	0.15	0.32	0.29	0.00	100.00	0.29	0.29
Declarative knowledge										
High structure	5	734	0.39	0.39	0.39	0.45	0.00	100.00	0.45	0.45
Medium structure	11	2270	0.30	0.27	0.33	0.41	0.17	25.60	0.14	0.69
Low structure	3	281	0.23	0.13	0.32	0.30	0.00	100.00	0.30	0.30
Other										
High structure	2	438	0.34	0.27	0.41	0.42	0.00	100.00	0.42	0.42
Medium structure	4	884	0.34	0.29	0.39	0.38	0.30	10.62	0.06	0.70
Skill										
High structure	4	704	0.32	0.26	0.37	0.45	0.19	20.00	0.14	0.77
Medium structure	21	3723	0.27	0.25	0.30	0.37	0.16	27.02	0.11	0.63
Low structure	10	807	0.34	0.29	0.39	0.42	0.15	41.00	0.18	0.66
Proceduralized										
High structure	2	588	0.28	0.22	0.35	0.42	0.19	15.07	0.11	0.72
Medium structure	17	2962	0.25	0.23	0.28	0.34	0.16	26.86	0.08	0.60
Low structure	7	593	0.32	0.26	0.38	0.37	0.12	46.15	0.17	0.56
Adaptive										
High structure	2	116	0.50	0.50	0.50	0.50[1]	0.00	100.00	0.50	0.50
Medium structure	7	1037	0.35	0.31	0.40	0.46	0.09	50.60	0.31	0.62
Low structure	3	214	0.40	0.40	0.40	0.48	0.00	100.00	0.48	0.48
Attitude										
Medium structure	6	488	0.11	0.04	0.18	0.15	0.00	100.00	0.15	0.15
Low structure	3	272	0.08	0.08	0.08	0.09	0.00	100.00	0.09	0.09

Note. *k* = number of observations (i.e., studies). *N* = number of participants. *r* = sample-weighted mean correlation. CI = confidence interval (L = lower, U = upper). ρ = estimated true validity after correcting for criterion and predictor unreliability. *SD* ρ = standard deviation of the estimated true validity. % Var. explained = % variance explained by sampling error and criterion unreliability. CV = credibility value (L = lower, U = upper). [1] = ρ and *r* are the same because the reliability estimates for both the criterion and predictor were not reported.

Comparing Table 5, which shows the relationship between general cognitive ability and training effectiveness by structure nested within criteria without removing the influence of training content, to Table 10 it is noticeable that the influence of structure on the relationship between general cognitive ability and training effectiveness is stronger (Table 9 *swmr*s = 0.38, 0.33, and 0.31 for high, medium, and low structure, respectively)
when the influence of training content is removed from the analyses (Table 5 *swmr*s = 0.31, 0.32, and 0.27 for high, medium, and low structure, respectively). This is shown by the *swmr*s for high versus medium versus low structure exhibiting greater differences and the confidence intervals not overlapping as much. In Table 5 these differences were smaller and the confidence intervals overlapped more. Even so, the percent of variance explained in Table 10 still suggests the operation of additional moderators, which still supports the investigation of evaluation criteria as a possible moderator. When comparing Table 5 with Table 10 with regard to the evaluation criteria, it can be seen that the cognitive criteria results are comparable in size—that is, the ATIs found with regard to cognitive criteria and declarative knowledge are still present even if the influence of training content is removed (Table 5 *swmr*s for cognitive criteria = 0.34, 0.29, and 0.21 and *swmr*s for declarative knowledge = 0.36, 0.28, and 0.20 for high, medium, and low structure, respectively versus Table 10 *swmr*s for cognitive criteria = 0.37, 0.31, and 0.24 and *swmr*s for declarative knowledge = 0.39, 0.30, and 0.23 for high, medium, and low structure, respectively). However, when examining the *swmr*s for the skill criteria and proceduralized criteria in Table 10 (*swmr*s for skill criteria = 0.32, 0. 27, and 0.34 and *swmr*s for proceduralized criteria =

0.28, 0. 25, and 0.32 for high, medium, and low structure, respectively) it is noticeable that the meaningful differences that existed between high and low structure in Table 5 (*swmr*s for skill criteria = 0.24, 0. 27, and 0.34 and *swmr*s for proceduralized criteria = 0.21, 0. 25, and 0.32 for high, medium, and low structure, respectively) are no longer evident in Table 10.

In summary, these results suggest that ATIs exist in the training research literature but they also suggest that their size tends to small. An ATI for cognitive criteria was found whereby trainees high in general cognitive ability seem to benefit from training programs high in structure while for trainees low in general cognitive ability the structure of the training program is not of importance. An ATI was also found for skill criteria whereby trainees with low general cognitive ability seem to benefit from training programs high in structure, while trainees high in general cognitive ability also benefit, but to a much smaller degree. Generally, the found effects were larger for the cognitive-based criteria than for the skill-based criteria.

However, the ATIs found for skill-based criteria appeared to be confounded by the content of the training program, specifically by interpersonal training content demands. This suggests that the ATIs found for cognitive-based evaluation criteria are stronger and more robust than the ATIs found for skill-based criteria. In the Discussion section, possible reasons for the small effect size are explored and discussed. The Discussion section also addresses the practical implications of the ATIs observed.

Discussion

Educational research (e.g., Snow & Lohman, 1984) has shown that not all individuals are alike and that some individual differences interact with treatment variables, such that some trainees benefit to a greater degree from one type of training program while other trainees benefit to a greater degree from a second type of training program. These interactions have been termed aptitude-treatment interactions (ATIs). The purpose of this study was to investigate the size and nature of these ATIs in the training literature. The aptitude that was investigated in this study was the general cognitive ability of trainees because it is the individual difference variable with the greatest influence on learning, training outcomes, and job performance (e.g., Hunter 1986; Ree, Carretta, & Teachout, 1995). The treatment that was investigated was the structure of the training program (e.g., instructor guidance, provision of objectives) because structure is the treatment variable that most ATI research has examined (Campbell & Kuncel, 2001). However, past research has produced inconsistent results with regard to the existence of ATIs—some research has suggested they exist (Snow & Lohman, 1984) while other research suggested that ATIs are not likely to exist (Bracht, 1970). Moreover, research has not been able to clearly show the nature of the interaction between trainee aptitude and the structure of the training program (Goldstein & Ford, 2002). Based on these inconclusive results it is understandable that ATIs are questioned to be worthwhile to consider.

This meta-analysis was able to shed some light on the existence and nature of ATIs in the training literature. The results supported the existence of ATIs in the training research literature, which stands in contrast to the inconsistent results found by prior research with regard to their existence. There are two possible reasons for this: (1) the size of the ATIs found in this meta-analysis were small, which suggests that primary research is unlikely to find ATIs because of small power and (2) moderators exist that influence the relationship of general cognitive ability and training effectiveness, which are not investigated by primary research, making if difficult to find evidence for the existence of ATIs.

The results also showed that higher structure in training programs generally benefits trainees, although given that the credibility intervals contained negative numbers, less structure might at times be beneficial as well. Research by Bell and Kozlowski (2002) on adaptive guidance is loosely consistent with the finding of increased structure being beneficial. These researchers have found that adaptive guidance positively impacts an individual's learning and performance. Adaptive guidance refers to information that helps trainees interpret the meaning of their past performance and helps them determine what they should focus on to improve their performance in the future. Although adaptive guidance is not identical to structure they share the concept of provision of additional information that individuals can use to attain higher levels of learning and performance.

An important result of this meta-analysis was that the interaction between general cognitive ability and the structure of the training program was moderated by the type of evaluation criterion and the content of the training program. An ATI was found for cognitive criteria showing that training structure exacerbates general cognitive ability effects, such that structure tends to facilitate cognitive learning for trainees high in general cognitive ability more so than for trainees low in general cognitive ability (in support of the model shown in Figure 3). Another ATI was found for skill criteria showing that training structure attenuates general cognitive ability effects, such that structure tends to facilitate skill-based learning for trainees low in general cognitive ability more so than for trainees high in general cognitive ability (in support of the model shown in Figure 2). However, the content of the training program, rather than the type of evaluation criterion, was found to be the possible driver behind the ATI that was found for the skill-based criterion. When the effect of training content was removed from the analyses, the ATI that was found for cognitive-based evaluation criteria was still present, while the ATI found for skill-based evaluation criteria was not evident anymore. This suggests that a robust ATI exists when trainees are evaluated with cognitive-based criteria while ATIs for skill criteria are less robust. That is, the confounding variable training content makes conclusions for the ATIs found for skill-based evaluation criteria tenuous.

Why do trainees with high general cognitive ability increase in cognitive learning more in a training program with high structure than trainees with low general cognitive ability? To answer this question, specific attention is paid to the robust ATI that was found with declarative knowledge. The learning of factual information (i.e., declarative knowledge) is the first stage of skill acquisition (Anderson, 1982). During this declarative knowledge stage individuals are focused on understanding the facts and fundamental basics needed to perform the task at hand. It is also during this first stage of skill acquisition where substantial cognitive resource demands are imposed on an individual as the individual invests a large amount of attention to understanding the task (Anderson, 1982). During this early stage of skill acquisition individuals primarily rely on their general cognitive ability to make sense of the novel task and its requirements. Individuals higher in aptitude will have more resource capacity for learning a task than individuals lower in aptitude. Ackerman's (1986a) research showed that general cognitive ability plays a strong role in the initial stages of skill acquisition when the basic task information is learned. With further practice, more specific abilities (e.g., perceptual speed ability) become important. However, during early skill acquisition it is crucial that cognitive resources be devoted to learning the task at hand—any information (e.g., a secondary task, off-task activities) that draws an individual's attention away from this focus, will draw the individual's cognitive recourses away from task issues and impair learning (Kanfer & Ackerman, 1989).

Based on these findings, I speculate that structure acts like a secondary task in some respects, inadvertently imposing a greater demand on the cognitive resources of a trainee, and diverting the resources that were already devoted to learning the task. The additional provision of structure causes trainees low in general cognitive ability to pay attention not only to the task information but also to the information on structure and the information that structure provides. In other words, trainees may become overloaded with information. Because trainees low in general cognitive ability have fewer cognitive resources to make sense of the task content and the structure provided, they are not able to benefit from the additional information that structure provides. Trainees high in general cognitive ability on the other hand, have a greater resource capacity and are able to better devote their cognitive resources to both information on the task and information that structure provides and are thus able to use this information to benefit their learning of declarative knowledge.

Although the results for the ATIs found for cognitive criteria are consistent with a previous meta-analysis conducted on ATIs in education (Whitener, 1988), these results stand in contrast to primary research that found increases in variables similar to structure (e.g., proceduralized versus exploratory training) to increase the learning of low aptitude individuals more than the learning of high aptitude individuals (e.g., Bell & Kozlowski, 2008). This dissimilarity might be due to structure being operationalized somewhat counter in this study to some other authors' operationalization of structure. That is, some authors have previously suggested that structure provides additional information that is easily useable by trainees low in general cognitive ability while trainees high in general cognitive ability do not need this additional information because they already possess the knowledge they need to learn the task (e.g., Cronbach & Snow, 1969, Snow, 1989). In contrast, in this study high structure involves trainees having to understand not only the task issues but also how the training works—information that is too much for trainees with low general cognitive ability to comprehend and use to their advantage. In previous studies, structure has been operationalized in fairly simple terms (i.e., one or two instructional elements), whereas in this investigation structure was operationalized in terms of a combination of instructional elements. Hence, the differences in results with regard to the benefits of increased structure might be due to the way structure is operationalized. Although the results of this meta-analysis with regard to the ATI with cognitive criteria are counter to previous primary study results, they are consistent with the idea that trainees high in general cognitive ability have the capabilities and knowledge to capitalize on and benefit from the additional structure that is provided (Whitener, 1989) and/or are more willing to take advantage of the additional information that structure provides them (Ceci & Paperieno, 2005).

It is also important to note the ATI found with skill criteria, although the effects were not robust due to a possible confounding with training content. This ATI showed that trainees low in general cognitive ability tended to benefit from training programs high in structure while trainees high in general cognitive ability neither benefited nor suffered from high structure. Whereas the ATI found with cognitive-based criteria is less consistent with past research, the ATI with skill-based criteria is more consistent with past research that showed structure to help those lower in aptitude to gain greater levels of skill while those high in aptitude are neither hurt nor benefited by high structure (e.g., Snow, 1986). Kanfer and Ackerman's (1989) research is also consistent with this: they found that individuals low in general cognitive ability were able to make greater use of goal-setting interventions than trainees high in general cognitive ability during later stages of skill acquisition, when the focus has moved from understanding the task to performing the task.

A final result of this meta-analysis is that the ATIs found are rather small in size. A possible explanation for ATIs generally not being found in primary studies is that the sample size of these studies is rather small. The average sample size of a study was approximately 142, therefore the sample size in each training condition present in the training program was at most half, approximately 71 trainees. Cohen (1992) was able to show that to detect a small effect-size-difference between two correlations of two training conditions with a power of 0.80 and an alpha of 0.05, each training condition has to have at least 1,573 trainees. This number is more than ten times larger than the total number of trainees included in an average study included in this meta-analysis. Thus, one important reason for primary studies rarely detecting ATIs is lack of statistical power.

Practical Implications

The results of this meta-analysis imply that ATIs are not practically relevant. This is due to two reasons: (1) the effect of ATIs found was very small and (2) the nature of the ATIs that were found are not what is commonly thought of when ATIs are described. ATIs are often described as being disordinal in nature (see Figure 1), which is represented by the classical "X" crossing of the regression lines suggesting that to maximize training effectiveness it is best to administer different training programs to trainees at each end of the aptitude continuum. The ATIs found in this study were ordinal in nature (see Figures 2 and 3), one spreading ATI with regard to cognitive-based criteria and one inverse ordinal interaction with regard to skill-based criteria. These types of interactions indicate that assigning trainees to differing training programs based on their general cognitive ability is not worthwhile because structure generally benefits all trainees—structure was not found to be disadvantageous to one group of trainees, as is suggested by a disordinal interaction. The implication of ATIs not being practically relevant, when the variables of interest are trainees' general cognitive ability and the structure of training, can be advantageous to organizations in that the assignment of different groups of trainees to different types of training programs is not necessary and so the investment in resources for designing and implementing different types of training programs can be avoided. The results however suggest that an increase in structure generally benefits trainees and so organizations should invest in offering trainees structured training programs, regardless of the general cognitive ability level of trainees or the criterion type with which training effectiveness is measured.

Limitations

As with other research, this meta-analysis also has several limitations. First, for some variables not enough studies could be found to either warrant an analysis (e.g., a division of the attitude criteria) or a meaningful interpretation (e.g., adaptive skill criteria). It might be worthwhile to examine the cognitive criteria even further (e.g., knowledge organization criteria) or to receive additional clarification on the ATIs with regard to the attitude criteria. Particularly interesting to practitioners would be findings on ATIs involving on-the-job skill criteria; however, not enough studies were located to be able to conduct an analysis on this criterion. Second, the standard deviations and percent of variance explained of the *swmr*s often still suggested that additional moderators might be operating. However, the decision was made not to seek out additional moderators due to two reasons. The first one was that theory did not suggest an additional investigation and the second one was that an investigation of additional moderators would have reduced the size of studies even further. Third, apart from structure, the moderator that was included in the analysis was the type of evaluation criterion; however, this does not directly speak to the content of the training program, which might be another driver of the nature of the interaction, as suggested by Campbell and Kuncel (2001). Although this possibility was briefly investigated, additional research into the training content would provide further insight. Finally, this meta-analysis is able to only speak to proceduralized skill criteria as one of the moderators of the ATIs investigated. Unfortunately, not enough studies were located to warrant an interpretation of the adaptive skill criteria, although tentative results are presented in Table 4. Additional results on adaptive skill would be interesting because this would give further insight into issues related to the transfer of training.

Future Research

As the interest in the influence of individual differences increases in the training literature, future research on ATIs might investigate the inclusion of different attributes of trainees: either other, more specific aptitude variables (e.g., specific abilities, such as verbal ability) or other individual difference variables that are non-aptitudes and are more personality and motivational based (e.g., self-efficacy, goal orientation/achievement motivation, need for cognition). The reasoning behind this suggestion is that a growing literature shows ATIs with non-aptitude individual difference variables (e.g., Bell & Kozlowski, 2008; Cullen, 2004; Gully et al., 2002; McInerney, McInerney, & Marsh, 1997; Towler & Dipboye, 2001). Due to the finding that the ATIs found for skill criteria might actually be driven by the content of the training program rather than by the type of evaluation criteria used, an additional suggestion for future research would be to manipulate the training content and the structure of the training program to gain a better understanding of the moderator that is operating in the relationship between general cognitive ability and training effectiveness. The resource allocation model (Kanfer & Ackerman, 1989) was discussed earlier as a possible explanation of the ATI found with declarative knowledge. It might be warranted for future research to investigate the resource allocation model as a viable explanation of why trainees high in general cognitive ability benefit from training programs high in structure while trainees low in general cognitive ability do not. Primary studies investigating the extent to which the introduction of various forms and amounts of structure differentially divert the cognitive resources of trainees lower or higher in general cognitive ability might be able to provide more insight into the given explanation. A concern that arose while collecting data for this meta-analysis was that many studies that provided pertinent variables did not provide the necessary statistics (e.g., *F*, *t*, *r*, *M*, *SD*) for inclusion. Future research should make an effort to report these kinds of statistics to facilitate the inclusion of more studies in future meta-analyses. This would also allow for more stable estimates and provide the possibility to investigate additional moderators.

References

References marked with a symbol indicate studies included in the meta-analysis:

† References marked with this symbol indicate studies included in the primary approach to analyzing ATIs

* References marked with this symbol indicate studies included in the secondary approach to analyzing ATIs

Ackerman, P. L. (1986a). Indivdidual differences in information processing: An investigation of intellectual abilities and task performance during practice. *Intelligence, 10*, 101-139.

*Ackerman, P. L., & Kanfer, R. (1993). Integrating laboratory and field study for improving selection: Development of a battery for predicting air traffic controller success. *Journal of Applied Psychology, 78*, 413-432.

*Ackerman, P. L., Kanfer, R., & Goff, M. (1995). Cognitive and noncognitive determinants and consequences of complex skill acquisition. *Journal of Experimental Psychology: Applied, 1*, 270-304.

Alliger, G. M., & Janak, E. A. (1997). Kirpatrick's levels of training criteria: Thirty years later. *Personnel Psychology, 42*, 331-342.

Alliger, G. M., Tannenbaum, S. I., Bennett, W. Jr., Traver, H., & Shotland, A. (1997). A meta-analysis of the relations among training criteria. *Personnel Psychology, 50*, 341-358.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369-406.

Arthur, W. Jr., Bennett, W. Jr., Edens, P. S., & Bell, S. T. (2003). Effectiveness of training in organizations: A meta-analysis of design and evaluation features. *Journal of Applied Psychology, 88*, 234-245.

Arthur, W. Jr., Bennett, W., Jr., & Huffcutt, A. I. (2001). *Conducting meta-analysis using SAS*. Mahwah, NJ: Lawrence Erlbaum Associates.

Arthur, W. Jr., Bennett, W. Jr., Stanush, P. L., & McNelly, T. L. (1998). Factors that influence skill decay a retention: A quantitative review and analysis. *Human Performance, 11*, 234-245.

*Baldwin, R. D., Cliborn, R. E., & Foskett, R. J. (1976). *The acquisition and retention of visual aircraft recognition skills*, ARI TR-76-A4, U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA.

*Baldwin, T. T., & Karl, K. A. (1987). *The development and empirical test of a measure for assessing motivation to learn in management education*. Paper presented at the 47th Annual Meeting of the Academy of Management, New Orleans, LA.

*Baldwin, T. T., Magjuka, R. J., & Loher, B. T. (1991). The perils of participation: Effects of choice of training on trainee motivation and learning. *Personnel Psychology, 44*, 51-65.

Bassi, L. J., & Van Buren, M. E. (1998). The 1998 ASTD state of the industry report. *Training & Development, 52*, 21-35.

†*Bell, B. (2002). An examination of the instructional, motivational, and emotional elements of error training. (Doctoral dissertation, Michigan State University, 2002). *Dissertation Abstracts International: Section B: The Sciences and Engineering, 63* (9-B).

Bell, B. S., & Kozlowski, S. W. J. (2002). Adaptive guidance: Enhancing self-regulation, knowledge, and performance in technology-based training. *Personnel Psychology, 55*, 267-306.

Bell, B. S., & Kozlowski, S. W. J. (2008). Active learning: Effects of core training design elements on self-regulatory processes, learning, and adaptability. *Journal of Applied Psychology, 93*, 296-316.

*Beier, M. E., & Ackerman, P. L. (2005). Age, ability, and the role of prior knowledge and the acquisition of new domain knowledge: Promising results in a real-world learning environment. *Psychology and Ageing, 20*. 341-355.

Berliner, D. C., & Cahen, L. S. (1973). Trait-treatment interaction and learning. *Review of Research in Education, 1*, 58-94.

Borowski, J. C., & Peck, V. A. (1986). Causes and consequences of metamemory in gifted children. In R. J. Sternberg & J. E. Davidson (Eds.), *Concepts of giftedness* (pp. 182-200). Cambridge, England: Cambridge University Press.

Bracht, G. H. (1970). Experimental factors related to aptitude-treatment interactions. *Review of Educational Research, 40*, 627-645.

*Bramble, W. J. Jr. (1993). Stability of performance-based ability tests during simulated flight training. (Doctoral dissertation, University of Central Florida, 1993). *Dissertation Abstracts International: Section B: The Sciences and Engineering, 60* (7-B).

Brown, K. G. (2001). Using computers to deliver training: Which employees learn and why? *Personnel Psychology, 54*, 271-296.

Campbell, J. P., & Kuncel, N. R. (2001). Individual and team training. In N. Anderson, D. S. Ones, H. K. Sinangil, Handan Kepir; C. Viswesvaran (Eds.) *Handbook of industrial, work and organizational psychology* (Vol. 1: Personnel psychology; pp. 278-312). Thousand Oaks: Sage Publications.

†*Campbell, M. (2007). Error management training from a resource allocation perspective: An investigation of individual differences and the training components that contribute to transfer. (Doctoral dissertation, Rice University, 2007). *Dissertation Abstracts International: Section B: The Sciences and Engineering, 68* (3-B).

*Campbell, S. K., & Catano, V. M. (2004). Using measures of specific abilities to predict training performance in Canadian Forces operator occupations. *Military Psychology, 16*, 183-201.

*Carter, S. D. (2002). Matching training methods and factors of cognitive ability: A means to improve training outcomes. *Human Resource Development Quarterly, 13*, 71-87.

Ceci, S. J., & Papierno, P. B. (2005). The rhetoric and reality of gap closing: When the “have-nots” gain but the “haves” gain even more. *American Psychologist, 60*, 149-160.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology, 85*, 678-707.

*Connelly, S., Allen, M. T., & Waples, E. (2007). The impact of content and structure on a case-based approach to developing leadership skills. *International Journal of Learning and Change, 2*, 218-249.

†*Crawford, J. (1978). Interactions of learner characteristics with the difficulty level of instruction. *Journal of Educational Psychology, 4*, 523-531.

Cronbach, L. J., & Snow, R. E. (1969). Individual differences in learning ability as a function of instructional variables. (Technical Report). Stanford: Stanford University, Office of Education, U.S. Department of Health, Education, and Welfare.

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. Oxford, England: Irvington.

Cullen, M. J. (2004). Training interpersonal skills: An aptitude-treatment interaction approach. (Doctoral dissertation, University of Minnesota, 2007). *Dissertation Abstracts International, 65* (08-B).

Dansereau, D. F., McDonald, B. A., Collins, D. W., Garland, J., Holley, C. D., Diekhoff, G. M., & Evans, S. H. (1979). Evaluation of a teaching strategy system. In H. F. O'Neil, Jr., & C. D. Spielberger (Eds.), *Cognitive and affective learning strategies* (pp. 3-43). New York: Academic Press.

Day, E. A., Arthur, W. Jr., Bell, S. T., Edwards, B. D., Bennett, W. Jr., Mendoza, J., & Tubré, T. C. (2005). Ability-based pairing strategies in the team-based training of a complex skill: Does the intelligence of your training partner matter? *Intelligence, 33*, 39-65.

*Day, E. A., Arthur, W. Jr., Gettman, D. (2001). Knowledge structures and the acquisition of a complex skill. *Journal of Applied Psychology, 86,* 1022-1033.

*Day, E. A., Arthur, W. Jr., Shebilske, W. L. (1997). Ability determinants of complex skill acquisition: Effects of training protocol. *Acta Psychologica, 9*7, 145-165.

*Day, E. A., Espejo, J., Kowollik, V., Boatman, P. R., & McEntire, L. E. (2007). Modeling the links between need for cognition and the acquisition of a complex skill. *Personality and Individual Differences, 42*, 201-212.

*Dean, M. A., Conte, J. M., & Blankenhorn, T. R. (2006). Examination of the predictive validity of Big Five personality dimensions across training performance criteria. *Personality and Individual Differences, 41*, 1229-1239.

*Donnelly, A. E. (1996). The effects of metacognitive skills training on hands-on learning from science objects. (Doctoral dissertation, University of Florida, 1996). *Dissertation Abstracts International, 58* (7-A).

Donovan, J. J., & Radosevich, D. J. (1998). The moderating role of goal commitment on the difficulty-performance relationship: A meta-analytic review and critical reanalysis. *Journal of Applied Psychology, 83*, 308-315.

*Eastman, P. M. (1972). The interaction of spatial visualization and general reasoning abilities with instructional treatment in quadratic inequalities: A follow-up study. (Doctoral dissertation, University of Texas at Austin, 1972). *Dissertation Abstracts International, 33* (9-A).

†*Ellefsen, E. R. (1996). Teaching a strategy to promote self-regulated learning in computer-based instruction. (Doctoral dissertation, University of Northern Colorado, 1996). *Dissertation Abstracts International, 57* (4-A).

*Espejo, J., Day, E. A., & Scott, G. (2005). Performance evaluations, need for cognition, and the acquisition of a complex skill: An attribute-treatment interaction. *Personality and Individual Differences, 38*, 1867-1877.

*Eyring, J. D., Johnson, D. S., & Francis, D. J. (1993). A cross-level units-of-analysis approach to individual differences in skill acquisition. *Journal of Applied Psychology, 78*, 805-814.

*Federico, P.-A., & Landis, D. B. (1979). *Discriminating between failures and graduates in a computer-managed course using measures of cognitive styles, abilities, and aptitudes,* NPRDC TR-79-21, Navy Personnel Research and Development Center, San Diego, CA.

*Fisher, S. L., & Ford, J. K. (1998). Differential effects of learner effort and goal orientation on two learning outcomes. *Personnel Psychology, 51*, 397-420.

Fleishman, E. A., & Mumford, M. D. (1989). Individual attributes and training performance. In I. L. Goldstein (Ed.), *Training and development in organizations* (pp. 183-255). San Francisco, CA: Jossey-Bass.

*Ford, J. K., Quinones, M. A., Sego, D. J., & Sorra, J. S. (1992). Factors affecting the opportunity to perform trained tasks on the job. *Personnel Psychology, 45*, 511-527.

*Furgang, S. R. (1980). The use of individual difference measures to affect the assignment of Army trainees to media options. (Doctoral dissertation, University of Southern California, 1980). *Dissertation Abstracts International, 41* (5-A).

Gehlbach, R. D. (1979). Individual differences: Implications for instructional theory, research, and innovation. *Educational Researcher, 8*, 8-14.

Goldstein, I. L. (1980). Training in work organizations. *Annual Review of Psychology, 31*, 229-272.

Goldstein, I. L., & Ford, J. K. (2002). *Training in organizations: Needs assessment, development, and evaluation* (4th ed.). Belmont, CA, US: Wadsworth/Thomson Learning.

*Gordon, M. E., & Kleiman, L. S. (1976). The prediction of trainability using a work sample test and an aptitude test: A direct comparison. *Personnel Psychology, 29*, 243-253.

*Goska, R. E., & Ackerman, P. L. (1996). An aptitude-treatment interaction approach to transfer within training. *Journal of Educational Psychology, 88*, 249-259.

Gulliksen, H. (1986). The increasing importance of mathematics in psychological research (Part 3). *The Score, 9*, 1-5.

Gully, S. M., Payne, S. C., Koles, K. L. K., & Whiteman, J.-A. K. (2002). The impact of error training and individual differences on training outcomes: An attribute-treatment interaction perspective. *Journal of Applied Psychology*, *87*, 143-155.

*Hertenstein, E. J. (2001). Goal orientation and practice condition as predictors of training results. *Human Resource Development Quarterly, 12*, 403-419.

*Holladay, C. L., Quiñones, M. A. (2003). Practice variability and transfer of training: The role of self-efficacy generality. *Journal of Applied Psychology*, *88*, 1094-1103.

Huffcutt, A. I., & Arthur, W., Jr. (1995). Development of a new outlier statistic for meta-analytic data. *Journal of Applied Psychology, 80*, 327-334.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior, 29*, 340-362.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, CA: Sage Publications.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74*, 657-690.

†*Karl, K. A. (1991). Effects of videotaped modeling and videotaped feedback on performance in training: An examination of personality by treatment interactions. (Doctoral dissertation, Michigan State University, 1991). *Dissertation Abstracts International*, *52* (7-A).

†*Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74*, 657-690.

*Kanfer, R., Ackerman, P. L., & Heggestad, E. D. (1996). Motivational skills and self-regulation for learning: A trait perspective. *Learning and Individual Differences*, *8, 185-209.*

Kirkpatrick (1959). Techniques for evaluation training programs. *Journal of the American Society for Training Directors*, *13*, 3-9.

Kirkpatrick, D. L. (1976). Evaluation of training. In R. L. Craig (Ed.), *Training and development handbook: A guide to human resource development* (2nd ed., pp. 301-319). New York: McGraw-Hill.

Kirkpatrick, D. L. (1996). Invited reaction: Reaction to Holton article. *Human Resource Development Quarterly, 7*, 23-25.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-base, experiential, inquiry-based teaching. *Educational Psychologist, 41*, 75-86.

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

*Knerr, C. M., Harris, J. H., O'Brien, B. K., Sticha, P. J., & Goldberg, S. L. (1984). *Armor procedural skills: Learning and retention,* ARI TR-621, U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA.

Kozlowski, S. W. J., Toney, R. J., Mullins, M. E., Weissbein, D. A., Brown, K. G., & Bell, B. S. (2001). Developing adaptability: A theory for the design of integrated-embedded training systems. In E. Salas (Ed.), *Advances in human performance and cognitive engineering research*, Vol. 1 (pp. 59-123). Amsterdam: Elsevier.

*Kozlowski, S. W. J., & Bell, B. S. (2006). Disentangling achievement orientation and goal setting: Effects on self-regulatory processes. *Journal of Applied Psychology, 91*, 900-916.

Kraiger, K. (2003). Perspectives on training and development. In W. C. Borman, D. Ilgen, R. J. Klimoski, & I. B. Weimer (Eds.), *Handbook of psychology* (pp. 171-192). Hoboken, NJ: John Wiley & Sons, Inc.

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology, 78,* 311-328.

Kyllonen, P. C., Lohman, D. F., & Snow, R. E. (1984). Effects of aptitudes, strategy training, and task facets on spatial task performance. *Journal of Educational Psychology, 76,* 130-146.

†Lemke, E. A., Randle, K., & Robertshaw, C. S. (1969). Effects of degree of initial acquisition, group size, and general mental ability on concept learning and transfer. *Journal of Educational Psychology, 60,* 76-78.

*Lievens, P., Harris, M. M., Van Keer, E., & Bisqueret, C. (2003). Predicting cross-cultural training performance: The validity of personality, cognitive ability, and dimensions measured by assessment center and a behavior description interview. *Journal of Applied Psychology, 88*, 476-489.

*Loh, V., Andrews, S., Griffin, B., & Hesketh, B. (2007). *Individual differences in the effects of error training: Who learns from their mistakes?* Paper presented at the 22nd Annual Conference of the Society for Industrial and Organizational Psychology, New York, NY.

*Lyons, T. J., Bayless, J. A., & Park, R. K. (2001). Relationship of cognitive, biographical, and personality measures with the training and job performance of detention enforcement officers in a federal government agency. *Applied HRM Research, 6*, 67-70.

†Mayer, R. E., Stichl, C. C., & Greeno, J. G. (1975). Acquisition of understanding and skill in relation to subjects' preparation and meaningfulness of instruction. *Journal of Educational Psychology, 67*, 331-350.

McInerney, V., McInerney, D. M. & Marsh, H. W. (1997). Effects of meta-cognitive strategy training within a cooperative group learning context on computer-achievement and anxiety: An aptitude-treatment interaction study. *Journal of Educational Psychology*, 89, 686-695.

*McLeod, D. B., & Adams, V. M. (1980). Aptitude-treatment interaction in mathematics instruction using expository and discovery methods. *Journal for Research in Mathematical Education, 11*, 225-234.

†*McLeod, D. B., & Briggs, J. T. (1980). Interactions of field independence and general reasoning with inductive instruction in mathematics. *Journal for Research in Mathematical Education, 11*, 94-103.

*Moffie, D., Calhoon, R., O'Brien, J. K. (1964). Evaluation of a management development program. *Personnel Psychology, 17*, 431-440.

*Myers, C. (1992). Core skills and transfer in the youth training schemes: A field study of trainee motor mechanics. *Journal of Organizational Behavior, 13*, 625-632.

*Nair, S. N., Czaja, S. J., & Sharit, J. (2007). A multilevel modeling approach to examining individual differences in skill acquisition for a computer-based task. *Journals of Gerontology, 62B*, 85-96.

Noe, R. A., & Colquitt, J. A. (2002). Planning for training impact: Principles of training effectiveness. In K. Kraiger (Ed.), *Creating, implementing, and maintaining effective training and developments: State-of-the-art lessons for practice* (pp. 53-79). San Francisco: Jossey-Bass.

Oakes, D. W., Ferris, G. P., Martocchio, J. J., Buckley, M. R., & Broach, D. (2001). Cognitive ability and personality predictors of training program skill acquisition and job performance. *Journal of Business and Psychology, 15,* 523-548.

*Pak, R., Roger, W. A., & Fisk, A. D. (2006). Spatial ability subfactors and their influence on computer-based information search task. *Human Factors*, *48*, 154-165.

Peterson, P. L., Janicki, T. C., & Swing, S. R. (1979). *Aptitude-treatment interaction effects of three teaching approaches: Lecture-recitation, inquiry, and public issues discussion* (Tech. Rep. No. 517). Madison: University of Wisconsin, Wisconsin Research and Development Center for Individualized Schooling.

Ree, M. J., Carretta, T. R., & Teachout, M. S. (1995). Role of ability and prior job knowledge in complex training performance. *Journal of Applied Psychology, 80,* 721-730.

Ree, M. J., & Earles, J. A. (1991a). Predicting training success: Not much more than *g*. *Personnel Psychology, 44,* 321-332.

Ree, M. J., & Earles, J. A. (1991b). Predicting job performance: Not much more than *g*. *Journal of Applied Psychology, 79,* 518-524.

*Rigney, J. W., Munro, A., & Crook, D. E. (1979). Teaching task-oriented selective reading: A learning strategy. In H. F. O'Neil, Jr., & C. D. Spielberger (Eds.), *Cognitive and affective learning strategies* (pp. 177-206). New York: Academic Press.

Ruiz, C. J. (1985). *Cognitive modifiability and irreversibility* (Pub. No. 4). Guayana, Venezuela: University of Guayana.

*Scarfo, S. J. (2002). Relationship between police academy performance and cadet level of education and cognitive ability. *Applied HRM Research, 7,* 39-40.

Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology, 65,* 643-661.

Schmidt, F. L, & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262-274.

*Sein, M. K., & Bostrom, R. P. (1989). Individual differences and conceptual models in training novice users. *Human Computer Interaction, 4,* 197-229.

Sharps, R. (1974). A study of interactions between fluid and crystallized abilities and two methods of teaching reading and arithmetic. (Doctoral dissertation, Pennsylvania State University, 1973). *Dissertation Abstracts International*, *35* (1432-A).

*Simon, S. J., Grover, V., Teng, J. T. C., & Whitcomb, K. (1996). The relationship of information system training methods and cognitive ability to end-user satisfaction, comprehension, and skill transfer: A longitudinal field study. *Information Systems Research, 4*, 466-490.

*Simon, S. J., & Werner, J. M. (1996). Computer training through behavior modeling, self-paced, and instructional approaches: A field experiment. *Journal of Applied Psychology, 81*, 648-659.

Snow, R. E. (1986). Individual differences in the design of educational programs. *American Psychologist, 41*, 1029-1039.

Snow, R. E. (1989). Aptitude treatment interaction as a framework for research on individual differences in learning. In P. L. Ackerman, R. J. Sternberg, & R. G. Glasser (Eds.), *Learning and individual differences* (pp. 13-59). New York: Freedman.

Snow, R. E., & Lohman, D. F. (1984). Toward a theory of cognitive aptitude for learning from instruction. *Journal of Educational Psychology*, *76*, 347-376.

†Stindard, T. A., & Dolphin, W. D. (1981). Which students from self-paced mastery instruction and why. *Journal of Educational Psychology*, *73*, 754-763.

†*Tobias, S. (1973). Sequence, familiarity, and attribute by treatment interactions in programmed instruction. *Journal of Educational Psychology*, *64*, 133-141.

Towler, A. J., & Dipboye, R. L. (2001). Effects of trainer expressiveness, organization, and trainee goal orientation on training outcomes. *Journal of Applied Psychology*, 86, 664-673.

Tracey, J. B., Tannenbaum, S. I., & Kavanagh, M. J. (1995). Applying trained skills on the job: The importance of the work environment. *Journal of Applied Psychology*, *80*, 239-252.

*Trites, D.K., & Cobb, B. B. (1963). *Problems in air traffic management IV. comparisons of pre-employment, job-related experience with aptitude tests as predictors of training and job performance of air traffic control specialists*, FAA TR-63-31, Federal Aviation Agency Civil Aeromedical Research Institute Aeronautical Center, Oklahoma City, OK.

*Tziner, A., & Dolan, S. (1982). Evaluation of a traditional selection system in predicting success of females in officer training. *Journal of Occupational Psychology, 55*, 269-275.

*Webb, L. F. (1971). Interaction effects between selected cognitive abilities and instructional treatment in algebra. (Doctoral dissertation, University of Texas at Austin, 1971). *Dissertation Abstracts International, 32* (10-A).

Wexley, K. N. (1984). Personnel training*. Annual Review of Psychology, 35*, 519-551.

†*Whitener, E. M. (1988). The utility of the interaction of aptitude with instructional support in training. (Doctoral dissertation, Michigan State University, 1989). *Dissertation Abstracts International, 49* (11-A).

Whitener, E. M. (1989). A meta-analytic review of the effect on learning of the interaction between prior achievement and instructional support. *Review of Educational Research, 59*, 65-86.

Wollins, L. (1962). Responsibility of raw data. *American Psychologist, 17*, 657-658.

*Young, W. C., Broach, D., & Farmer, W. L. (1997). *The effects of video game experience on computer-based air traffic controller specialist, air traffic scenario test scores*, FAA TR-AM-97/4, Federal Aviation Administration Civil Aeromedical Research Institute, Oklahoma City, OK.

Appendix A

ATI with Feedback as the Treatment Variable

Due to several limitations with the feedback variable (e.g., not normal distribution of feedback scores), details with regard to the Method and Results sections of this variable are presented in this Appendix.

Method

Training Feedback. To arrive at a feedback score for each data point, the scores of each variable that measured feedback were summed. Possible ratings for each variable ranged from 0 (involved in training) to 1 (not involved in the training). The feedback variables included: (1) general feedback, (2) specific feedback, (3) informative feedback, (4) normative feedback, (5) criterion feedback, (6) feedback with guidance, (7) positive feedback, (8) negative feedback, (9) immediate feedback, (10) delayed feedback, (11) feedback from a peer, (12) feedback from a supervisor, (13) feedback from a lab assistant/trainer, and (14) feedback from a machine/computer. The overall feedback scores ranged from 0 to 17. The distribution of feedback scores can be seen in Table 11.

Table 11
Feedback Variable Frequencies

Structure Score	Frequency
0	47
5	1
6	4
7	30
8	11
13	43
14	3
15	8
17	1

After examining the frequency distributions of the feedback scores, it was decided to trichotomize the structure variable for the moderator analyses in the following way: low feedback = 0 – 4, medium feedback = 5 – 13, and high feedback = 14 – 17. Following the categorization of feedback scores, the studies to be included in the moderator analyses were determined in the same way as they were determined for the structure variable. Of the initial 63 studies, a total of 60 studies were retained and included in the moderator analyses. Table 12 shows the trichotomized feedback frequency distribution of the studies included in the moderator analysis.

Table 12
Feedback Category Frequencies of Studies Included in the Moderator Analysis

Feedback Category	Frequency of Studies Included
Low	27
Medium	27
High	6

Training Condition Comparisons. To determine the main effects for feedback, 7 studies were identified to be included in the main effect analysis. The inclusion criteria of studies in the main effect analysis were the same as for the structure main effect analysis.

Coding Reliability. Across the common set of 10 articles, an ICC of 0.84 was obtained for the ratings of the feedback variable.

Results

Results for the feedback treatment variable are presented briefly. The focus is on the moderator and main effect analyses.

Moderator Analysis. Table 13 shows the meta-analytic results for the relationship between general cognitive ability and training effectiveness by feedback nested with criteria.

Table 13
Relationship between Cognitive Ability and Training Effectiveness by Feedback Nested within Criteria

Criteria/feedback	*k*	*N*	*r*	90% CI L	90% CI U	ρ	*SD* ρ	% Var. explained	90% CV L	90% CV U
Overall	60	8565	0.29	0.27	0.30	0.37	0.13	38.79	0.15	0.58
Feedback										
High	6	1364	0.18	0.14	0.23	0.24	0.07	57.31	0.12	0.35
Medium	27	4739	0.30	0.28	0.32	0.37	0.10	42.42	0.20	0.54
Low	27	2462	0.30	0.27	0.33	0.28	0.20	28.34	0.05	0.70
Cognitive										
High feedback	4	1152	0.29	0.24	0.33	0.37	0.08	41.99	0.24	0.50
Medium feedback	13	2508	0.34	0.31	0.36	0.40	0.12	28.54	0.20	0.60
Low feedback	17	1675	0.27	0.23	0.31	0.35	0.19	33.21	0.04	0.66
Declarative knowledge										
High feedback	4	1152	0.29	0.24	0.33	0.37	0.08	41.99	0.24	0.50
Medium feedback	12	2424	0.34	0.31	0.37	0.41	0.11	33.59	0.23	0.59
Low feedback	14	1514	0.28	0.24	0.32	0.37	0.20	31.40	0.05	0.70
Other										
Medium feedback	3	976	0.27	0.22	0.32	0.30	0.27	4.26	-0.14	0.75
Low feedback	5	446	0.13	0.05	0.21	0.14	0.01	99.80	0.13	0.15
Skill										
High feedback	6	1364	0.18	0.14	0.23	0.24	0.07	57.31	0.12	0.35
Medium feedback	22	3887	0.32	0.32	0.32	0.37	0.00	100.00	0.37	0.37
Low feedback	16	1213	0.39	0.35	0.43	0.49	0.15	43.53	0.25	0.73
Proceduralized										
High feedback	5	1275	0.13	0.09	0.18	0.19	0.03	89.30	0.14	0.24
Medium feedback	21	3380	0.30	0.30	0.30	0.35	0.00	100.00	0.35	0.35
Low feedback	9	723	0.38	0.33	0.44	0.45	0.13	42.84	0.23	0.66
Adaptive										
Medium feedback	5	1068	0.39	0.39	0.39	0.44	0.00	100.00	0.44	0.44
Low feedback	7	490	0.40	0.40	0.40	0.48	0.00	100.00	0.48	0.48

Table 13 continued
Relationship between Cognitive Ability and Training Effectiveness by Feedback Nested within Criteria

				90% CI					90% CV	
Criteria/feedback	*k*	*N*	*r*	L	U	ρ	*SD* ρ	% Var. explained	L	U
Attitude										
Medium feedback	3	620	0.16	0.09	0.22	0.17	0.04	79.69	0.11	0.23
Low feedback	10	771	0.08	0.02	0.14	0.09	0.00	100.00	0.09	0.09

Note. *k* = number of observations (i.e., studies). *N* = number of participants. *r* = sample-weighted mean correlation. CI = confidence interval (L = lower, U = upper). ρ = estimated true validity after correcting for criterion and predictor unreliability. *SD* ρ = standard deviation of the estimated true validity. % Var. explained = % variance explained by sampling error and criterion unreliability. CV = credibility value (L = lower, U = upper).

In contrast to the results with regard to structure, it can be seen in Table 13 that the feedback variable moderates the correlation between general cognitive ability and training effectiveness: rather large differences exist in the *swmr*s between high and low feedback (0.18 and 0.30, respectively) and the confidence intervals do not overlap. However, the operation of additional moderators is suggested by the percent of variance explained and justifies the analysis of evaluation criteria type as a moderator. The trend for the cognitive evaluation criteria for structure (i.e., a stronger relationship between trainee general cognitive ability and training effectiveness for high as opposed to low structured training programs) is present for the feedback variable as well but to such a small degree that the differences cannot be said to be meaningful. The same result of no meaningful differences between the low and high feedback training programs is obtained for the declarative knowledge evaluation criteria. It is to note though, that the cognitive criteria and high feedback variable only included four studies. Furthermore, it is also notable that the difference between the medium and low feedback for the cognitive criteria is large (*swm*r for medium feedback = 0.34; *swm*r for low feedback = 0.27) and the non-overlapping confidence intervals suggest this to be a meaningful difference.

Similarly to the results with regard to structure and skill evaluation criteria, the results for feedback also show smaller correlations for training programs high in feedback compared to training programs low in feedback, with a strong trend for proceduralized skill criteria.

As with the moderator analysis with regard to the structure variable, the moderator analysis for the feedback variable was re-conducted with the six studies high in interpersonal training content excluded. The results of this analysis are shown in Table 14.

Table 14
Relationship between Cognitive Ability and Training Effectiveness by Feedback Nested within Criteria removed from the Influence of Training Content

				90% CI					90% CV	
Criteria/feedback	*k*	*N*	*r*	L	U	ρ	*SD* ρ	% Var. explained	L	U
Overall	54	7416	0.30	0.29	0.32	0.39	0.13	40.10	0.18	0.61
Feedback										
High	5	1286	0.19	0.14	0.23	0.24	0.08	49.81	0.11	0.37
Medium	23	3875	0.33	0.31	0.35	0.40	0.09	48.45	0.25	0.55
Low	26	2255	0.32	0.28	0.35	0.40	0.20	29.03	0.07	0.73
Cognitive										
High feedback	4	1152	0.29	0.24	0.33	0.37	0.08	41.99	0.24	0.50
Medium feedback	11	1808	0.38	0.31	0.35	0.41	0.11	35.57	0.27	0.63
Low feedback	16	1468	0.29	0.25	0.33	0.37	0.20	33.03	0.04	0.70
Declarative knowledge										
High feedback	4	1152	0.29	0.24	0.33	0.37	0.08	41.99	0.24	0.50
Medium feedback	10	1724	0.39	0.36	0.42	0.47	0.07	53.75	0.35	0.59
Low feedback	13	1307	0.30	0.26	0.34	0.41	0.21	31.11	0.06	0.76

Table 14 continued
Relationship between Cognitive Ability and Training Effectiveness by Feedback Nested within Criteria removed from the Influence of Training Content

				90% CI					90% CV	
Criteria/feedback	*k*	*N*	*r*	L	U	ρ	*SD* ρ	% Var. explained	L	U
Other										
Medium feedback	3	976	0.27	0.22	0.32	0.30	0.27	4.26	-0.14	0.75
Low feedback	5	446	0.13	0.05	0.21	0.14	0.01	99.80	0.13	0.15
Skill										
High feedback	5	1256	0.15	0.10	0.19	0.20	0.09	44.89	0.05	0.35
Medium feedback	19	3423	0.34	0.34	0.34	0.41	0.00	100.00	0.41	0.41
Low feedback	16	1213	0.39	0.35	0.43	0.49	0.15	43.53	0.25	0.73
Proceduralized										
High feedback	4	1197	0.13	0.09	0.18	0.19	0.05	72.04	0.11	0.27
Medium feedback	18	2916	0.33	0.33	0.33	0.39	0.00	100.00	0.39	0.39
Low feedback	9	723	0.38	0.33	0.44	0.45	0.13	42.84	0.23	0.66
Adaptive										
Medium feedback	5	1068	0.39	0.39	0.39	0.44	0.00	100.00	0.44	0.44
Low feedback	7	490	0.40	0.40	0.40	0.48	0.00	100.00	0.48	0.48
Attitude										
Medium feedback	3	620	0.16	0.09	0.22	0.17	0.04	79.69	0.11	0.23
Low feedback	9	564	0.08	0.01	0.15	0.10	0.00	100.00	0.10	0.10

Note. *k* = number of observations (i.e., studies). *N* = number of participants. *r* = sample-weighted mean correlation. CI = confidence interval (L = lower, U = upper). ρ = estimated true validity after correcting for criterion and predictor unreliability. *SD* ρ = standard deviation of the estimated true validity. % Var. explained = % variance explained by sampling error and criterion unreliability. CV = credibility value (L = lower, U = upper).

Similar to the results with the structure variable, removing the influence of training content from the analysis had no effect on the results with regard to the cognitive evaluation criteria. However, in contrast to the results with the structure variable, the removal of training content from the analysis did not affect the results with regard to the skill evaluation criteria—an ATI is still evident in Table 14.

Main Effect Analysis. Table 15 shows the main effect results for the feedback variable. Table 15 shows that the effect of feedback on training effectiveness is small but positive (*swmr* = 0.20, ρ = 0.26), suggesting that increased feedback benefits trainees in general but because some of the credibility values include negative numbers, increased feedback, as structure, is not always beneficial to trainees (Kluger & DeNisi, 1996).

Table 15
Main Effect Results for Feedback, broken down by Cognitive and Skill Criteria

				90% CI					90% CV	
Feedback	*k*	*N*	*r*	L	U	ρ	*SD* ρ	% Var. explained	L	U
Overall	7	828	0.20	0.14	0.25	0.26	0.28	15.26	-0.20	0.72
High vs. Low	2	162	0.18	0.00	0.44	0.19	0.16	66.71	0.19	0.19
Medium vs. Low	5	666	0.20	0.14	0.26	0.27	0.27	15.26	-0.17	0.72
Cognitive	4	353	0.22	-0.18	0.63	0.25	0.25	43.09	0.25	0.25
High vs. Low	2	162	0.06	0.06	0.06	0.06[1]	0.00	100.00	0.06	0.06
Medium vs. Low	2	191	0.37	-0.09	0.82	0.37[1]	0.28	35.89	0.37	0.37
Skill	4	535	0.12	0.05	0.19	0.17	0.00	100.00	0.17	0.17
Medium vs. Low	4	535	0.12	0.05	0.19	0.17	0.00	100.00	0.17	0.17

Note. *k* = number of observations (i.e., studies). *N* = number of participants. *r* = sample-weighted mean correlation. CI = confidence interval (L = lower, U = upper). ρ = estimated true validity after correcting for criterion and predictor unreliability. *SD* ρ = standard deviation of the estimated true validity. % Var. explained = % variance explained by sampling error and criterion unreliability. CV = credibility value (L = lower, U = upper). [1] = ρ and *r* are the same because the reliability estimates for both the criterion and predictor were not reported.

Similarly as for the main effect of structure, Table 15 also shows that the different types of feedback (i.e., high vs. low and medium vs. low structure) that were examined as moderators are consistent with the idea that increased feedback benefits trainees (*swmr* = 0.18 for the high vs. low comparison and *swmr* = 0.20 for the medium

v. low structure comparison). Because the percent variance explained suggested the operation of moderators the type of criterion was examined as a further moderator. Table 13 also shows similar results to those of structure with regard to cognitive and skill criteria: the *swmr*s for the various amounts of feedback support the benefit of increased feedback in training programs.

To determine the nature of the ATI found for the skill-based evaluation criteria a figure was created such that the main effect of feedback was plotted with the different slopes for low versus high feedback. Figure 7 shows that high feedback generally benefits trainees when their training effectiveness is measured with skill criteria.

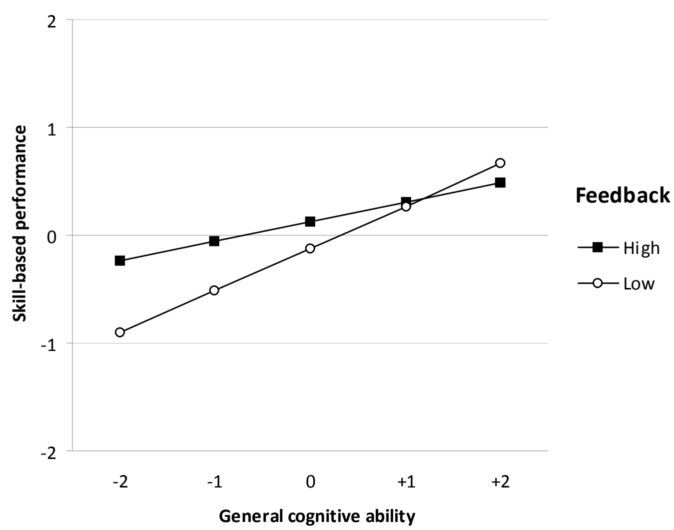

Figure 7. ATI (Feedback) for Skill Criteria. The values on the axes reflect standard deviations.

Figure 7 also shows, similarly to the ATI based on structure, that the effects of feedback are stronger for trainees who are at the lower ends of the general cognitive ability distribution. At the extreme high end of the general cognitive ability continuum, trainees appear to perform better in training programs with low feedback—this crossover effect was not obtained with the ATI based on structure.

In summary, the results for ATIs based on feedback are loosely consistent with those found based on structure. Although the same trends with regard to cognitive and skill evaluation criteria are observed, the effect with regard to cognitive evaluation criteria is much less strong, while the effect with regard to skill evaluation criteria is slightly stronger. These results lend further support to the existence of ATIs, even with different treatment variables.