There is a growing body of research showing that the methods used by an organization to manage its human resources can have a substantial impact on many organizationally relevant outcomes. Human resource management (HRM) practices have been linked with turnover (Arthur 1994), productivity (Ichniowski, Shaw, & Prennushi 1997; MacDufuy 1995), financial returns (Delery & Doty 1996), survival (Welbourne & Andrews 1996), and firm value (Huselid 1995). This growing area of research has been labeled strategic human resource management (SHRM) in that it emphasizes the strategic role of human resources management in meeting business objectives. Establishing that HRM practices are linked with firm effectiveness is an important first step in this line of research, however, there is still little understanding of the mechanisms through which HRM practices influence effectiveness. For instance, Delaney and Huselid (1996) stated that researchers still do not know "how HRM practices affect organizational outcomes, whether some practices have stronger effects than others, and whether complementarities or synergies among such practices can further enhance organizational performance" (p. 950). This line of research has been interesting in that it points to the importance of human resources, but it has provided few real insights for organizations wishing to gain a competitive advantage through human resources.

The purpose of this article is twofold. First, I provide a brief review of the major theoretical and empirical work completed to date in SHRM, focusing most closely on the fit of HRM practices into a coherent system. Second, I present several issues critical to empirical research in SHRM on which researchers must focus greater attention. To date, there is clearly no consensus as to how researchers must address these issues. It is argued throughout this article that the conceptual foundations of SHRM have been relatively weak and many of the empirical investigations have made assumptions not driven
by, or consistent with, the theoretical base. Given that this research has produced results that suggest overwhelming benefits from effectively managing human resources (Huselid 1995), I believe that we have only begun to realize how important human resource management is to competitive advantage. After all, what would an organization be without people?

THEORY IN STRATEGIC HUMAN RESOURCE MANAGEMENT

In explaining the significance of human resources to firm performance, the majority of work in SHRM has either implicitly or explicitly adopted the resource-based view of the firm proposed by Barney (1991) (Lado, & Wilson 1994; Wright, McMahan, & McWilliams 1994). The resource-based view proposes that an organization can gain a competitive advantage from the resources it possesses. This shifts the focus from the external environment and how the firm positions itself in a competitive market, to the internal resources of the firm and how the firm is able to use these resources to gain a competitive advantage. For a resource to be a source of competitive advantage, however, it must be rare, valuable, inimitable, and non-substitutable. While the HRM practices of a firm can lead to competitive advantage through developing a unique and valuable human capital pool (Barney & Wright 1997), they may also lead to competitive advantage as part of organizational capital by providing firms with both increased fit and flexibility (Wright & Snell in press).

Using the criteria of valuable, rare, and inimitable, both Wright et al. (1994), and Barney and Wright (1997) have proposed that human resources can indeed be a source of competitive advantage. Human resources, for instance, can clearly add value to a firm. This is evidenced in much of the micro HRM research and in that of utility analyses that show increases in profits with the use of a particular HRM practice. Valuable human resources are also rare to the extent that knowledge, skills, and abilities are normally distributed in the population. A firm's human capital may also be inimitable to the extent this capital has developed a unique history and culture within the organization. This focus on firm resources is important to SHRM research in that it represents a shift in focus from simply the HRM practices to the actual human resources of a firm. That is, a firm does not gain a competitive advantage from HRM practices, per se, but from the human resources that the firm attracts and retains.

While the resource based view provides a nice backdrop, explaining the importance of human resources to firm competitiveness, it does not specifically deal with how an organization can develop and support the human resources it needs for competitive advantage. I believe that most of the work in SHRM acknowledges that human resource management practices and policies are the main tools that organizations employ to develop and sustain the necessary human resources. Along these lines researchers have focused on various HRM practices organizations should use to achieve their objectives, and how these practices work together to enhance performance.
Consistent with this idea, SHRM has been based in large part on the notion that a firm must align its human resource management (HRM) practices to support business objectives. As such, researchers have focused predominantly on two forms of “fit,” vertical and horizontal (Baird, & Meshoulam 1988; Wright & McMahan 1992). Horizontal fit refers to the alignment of HR practices into a coherent system of practices that support one another. Vertical fit refers to the alignment of HR practices with the specific organizational context. Although there have now been several empirical investigations of the effectiveness of both forms of “fit,” there is relatively little empirical evidence to suggest that such alignment is necessary or beneficial. There are several possible explanations for the lack of strong empirical support for notions of fit in SHRM that will be explored in this article. First, however, it is important to understand the theoretical framework suggesting “fit.” Because the notion of vertical fit has received the most attention in the literature and there are now relatively thorough reviews of the most important issues surrounding this fit (Chadwick, & Cappelli in press), I focus here only on the horizontal or internal fit of HR practices. This area is critical for the future of research in SHRM, however, it has received substantially less attention in the literature.

**HORIZONTAL FIT OF HRM PRACTICES**

Horizontal “fit” in SHRM research deals with the internal consistency, and complementarity of HRM practices. Specifically, how HRM practices work together as a system to achieve organizational objectives. This horizontal notion of fit shifts the focus from individual HRM practices to the entire HRM system. This systems focus is discussed below.

**Human Resource Management Systems**

The following discussion explores the notion of horizontal “fit” in SHRM research. This fit deals with whether the practices used by an organization fit into a coherent system or “bundle” (Delery & Doty 1996; MacDuffie 1995) of practices that enhance and support the effectiveness of one another. The basic assumption is that the effectiveness of any practice depends on the other practices in place. If all of the practices fit into a coherent system, the effect of that system on performance should be greater than the sum of the individual effects from each practice alone (Ichniowski et al. 1997).

To test this “fit” it is important to describe in detail how practices may work together to achieve organizational objectives. Although many authors have proposed systems of internally consistent practices (Arthur 1992; Delery & Doty 1996; Miles & Snow 1984), few if any have sufficiently described how the practices support and enhance one another. Additionally, tests of horizontal fit have been limited and have shown very little support for these fit or systems effects (Delaney & Huselid 1996; Delery & Doty 1996; Huselid 1995). Complicating the study of such systems is the fact that there is no agreement as to
which practices make up a system (Becker & Gerhart 1996). This may in large part be due to the lack of a solid theoretical framework when choosing practices, and the fact that many practices may result in the same outcomes (e.g., work force characteristics). The following discussion will explore the essential components of an effective HRM system, and highlight how these components work together to achieve performance.

In an attempt to clarify issues of internal fit, Doty and Delery (1997) argued that there are countless combinations of practices that will result in the identical organizational outcomes (i.e., a high performance work force). For instance, there are several selection techniques, and combinations of techniques that should result in an equally skilled work force. To study the internal fit of practices, therefore, closer attention must be paid to whether an organization has practices in place to ensure a skilled work force, rather than focusing on whether an organization is using a particular type of selection device (e.g., a cognitive ability test). This highlights the fact that the level in the HRM system at which HRM activities are measured has important implications (Becker & Gerhart 1996). It may be much less important that an organization use one best practice, than that the organization has a system in place to ensure a skilled, motivated, and empowered work force.

If we adopt a model of SHRM where HRM practices lead to competitive advantage to the extent that they result in valuable, rare, and inimitable resources, it is essential that we identify the specific links between HRM practices and these resources. If, for instance, we agree that firms can achieve a competitive advantage by having a high performance core work force, then it is essential to identify which specific HRM practices actually develop such a work force. This process is complicated by the fact that HRM practices lead to multiple outcomes and these outcomes (e.g., work force skills) can be achieved by multiple combinations of HRM practices. As Delery and Doty (1996) noted, the configurational arguments inherent in the systems view incorporate the assumption of equifinality. It is, therefore, assumed that identical outcomes can be achieved by a number of different systems of HRM practices. In addition, some practices may only lead to certain outcomes when coupled with other practices that support it. For instance, developing valid selection devices may improve the overall skill of the work force. Implementing such a device, however, may only have the desired effect when other practices, such as high base pay that increases the applicant pool, are also in place.

Taylor and Russell (1939) long ago realized that valid selection devices only improve the work force to the extent that the firm is able to hire a small percentage of those who apply. To get the most out of a valid selection device it is essential to ensure that recruiting practices that increase the number of qualified applicants, thereby decreasing the selection ratio, are also in place (Shaw, Delery, Jenkins, & Gupta 1998). This highlights the fact that HRM practices may work together in a synergistic interdependent relationship. Other HRM practices may also enhance skills, however, in different ways. For instance, training practices may enhance skills regardless of the use of valid selection devices. From this brief discussion it is clear that HRM practices may
work together in a system in a number of different ways. In other words, individual HRM practices may show different relationships with each other in producing valuable, rare, and inimitable resources. This is a critical issue in SHRM research and it is more fully discussed below.

**Relationships among HRM Practices.** There are many different types of relationships among HRM practices that complicate the identification of systems of practices that lead to particular firm resources. It is important, however, that these relationships be acknowledged to improve our understanding of internal fit.

First, some HRM practices may show an *additive* relationship with each other in producing a desired outcome. In other words, they have independent non-overlapping effects on the outcome. This would be the case if using both practices together resulted in a greater level of the outcome (i.e., resource) than using either one alone, but not more than the sum of the individual effects of each practice. An example would be using a work sample test in conjunction with a cognitive ability test. If these two techniques measure different knowledge, skills, or abilities, together they may improve the overall skill level of the work force, but by no more than the sum of the effects due to each practice alone. While each of these practices may enhance work force skills, they may have only direct independent effects.

Second, HRM practices may have *interactive* effects in that their effectiveness depends on the level of the other practices in the system. There are several types of interactive relationships. Some practices exhibiting an interactive relationship may be *substitutes* for one another (Ichnioski, Kochan, Levine, Olson, & Strauss 1996). This occurs when both practices lead to the identical outcome. If two practices are substitutes, then the effect of adding one of them to a system will depend on whether the other practice is present. For instance, if two practices are substitutable and one is already in use, the second practice will add nothing except the expense associated with its implementation. Alternatively, if neither is present, adding one of the practices should increase the desired outcome. Using the work force skills example, it is possible that two different selection devices that both measure cognitive ability may be equally effective in selecting skilled employees. There also may be many different recruitment practices that could result in the same size and quality of applicant pool. In these cases, the firm should choose the least costly alternative. Another example of practices that may be substitutable is the case of work force empowerment. There are countless methods organizations can use to increase participation and empowerment in their work forces. Quality circles, off-line problem-solving teams, and other formal participation in decision-making programs may all lead to increased levels of empowerment in the work force. It is unlikely, however, that using all of these techniques increases empowerment beyond using only one or two techniques. To the extent these participation-enhancing techniques produce identical results, they would be considered substitutes for one another.

The more common discussion in SHRM research regarding the interactive
relationship among HRM practices has focused on a synergistic relationship between practices. HRM practices have a synergistic relationship when together they result in a substantially different effect than the sum of their individual effects would lead one to believe. Numerous researchers have argued that such synergistic effects exist (Becker et al. 1997; Delery & Doty 1996; Ichniowski et al. 1997; MacDufﬁe 1995) and that this is precisely what distinguishes the field of strategic human resource management from other research areas in human resource management. While a similar effect is described above for substitutable practices, the following relationships describe situations that lead to substantially different levels of the outcome than could be expected by either practice alone.

In describing synergistic relationships it is important to acknowledge that there are at least two possible forms. First, two practices may work together to enhance each other’s effectiveness. In this case the whole is greater than the sum of the parts. One example of this positive synergistic relationship would be when firm speciﬁc training programs are combined with highly selective staffing practices. The effectiveness of the training programs may be greatly enhanced if the organization is able to select individuals who have a greater ability to learn new material. When two practices work together in this fashion, Becker et al. (1997) have labeled their relationship a “powerful connection” signifying the positive outcomes achieved by their combined use. Second, two practices may actually work against one another. In this case, organizations that combine the two practices will have poorer performance than if they employed only one or the other. Becker et al. (1997) used the term “deadly combination” to describe such situations. An example of this would be when a ﬁrm designs the workplace to support a team production concept, but then rewards employees based only on individual performance. These rewards may work to undermine the team concept.

In sum, HRM practices may have independent or interactive effects on the human resources of a ﬁrm. This is not a new idea. Researchers have already discussed various forms of the interactive effects. Ichniowski et al. (1996) argued that some HRM practices may be substitutes, while other complements. In their description of complements, however, they only discuss what I have called positive synergistic effects. The possible negative synergistic effects have not been emphasized much in the literature, with the exception of Becker et al. (1997). Acknowledging these relationships among HRM practices has important implications for the identiﬁcation and measurement of HRM systems.

**IMPLICATIONS OF VERTICAL AND HORIZONTAL FIT FOR RESEARCH**

In the remainder of the article, I discuss what I believe are the most important issues for SHRM researchers to address in the following years. Each of these issues has already been identiﬁed and at least partially addressed in the literature. Therefore, I do not give equal weight to each issue. Rather, I focus
most closely on those issues that I feel have not received their due attention. First, I discuss issues surrounding the appropriate level of analysis for SHRM research. Some researchers have focused solely on the firm level, while others have measured and studied the effectiveness of HRM practices within only parts of the firm. Second, I discuss the issue of construct definition in SHRM research. Becker and Gerhart (1996) point out that researchers have measured HRM philosophies, policies, and practices often without clearly articulating the differences. There is clearly no consensus as to what a HRM practice really is. Third, I review and discuss the issues involving the measurement of HRM systems. I pay particular notice to the relationships among HRM practices noted above. Fourth, I discuss the often unmeasured mediating variables in SHRM research. While I do not argue that these variables must always be measured, they are clearly important in that they provide the answers as to how HRM policies and practices can influence firm performance. It was argued above that HRM systems influence firm performance through creating valuable, rare, and inimitable firm resources. In defining and measuring HRM systems, we must pay closer attention to these relationships. Finally, I discuss the choice of a dependent variable. In much of the SHRM literature, it has been assumed that certain types of HRM systems (e.g., high performance work system) maximize many performance indicators. It, however, is unlikely that a system will exhibit the same relationship with all firm level outcomes.

Level of Analysis

The level of analysis at which HRM practices are measured has important implications. Huselid (1995), for instance, measured HRM practices for exempt and non-exempt employee separately and then created a firm level measure by computing a weighted average of the practices. Arthur (1992), Delery and Doty (1996), Ichniowski et al. (1997), and MacDuffie (1995) measured HRM practices for a particular group of employees performing the same or similar job within the firm. Similarly, Tsui et al. (1997) measured HRM practices at the job level both within and across organizations. The key issue is that the constructs of interest must be measured at the appropriate level.

Doty and Delery (1997) proposed a framework of SHRM where it is more important that some employees, namely the core work force, be managed by a high performance work system. Such a framework has direct implications for the level of analysis at which HRM practices should be measured. The important construct would be the HRM system for the core work force. The other HRM systems in the organization may be less important. Measures like those obtained by Huselid (1995), where practices are measured across a corporation and a weighted average of the practices of exempt and non-exempt employees is computed, may be less appropriate. While organizations using high performance work practices for their core work force would likely receive higher scores on such measures, researchers are probably under- or overestimating the degree to which that organization is effectively managing its human resources. The main problem with this approach is that practices used to manage
the non-core work force may be less important, but may have equal or greater weight, depending on the number of employees in this category, in the overall measure.

The main issue here is that researchers must pay closer attention to the level of analysis when measuring HRM practices. Creating a measure by averaging the use of practices across an organization seems to assume that all employee groups are equally important. This clearly may not be the case. Tsui et al. (1997) provide relatively strong evidence that organizations use different HRM systems for employees in different jobs. This appears to suggest that the particular HRM system may be required to fit a particular job. Snell (1992), for instance, provides evidence that the knowledge of cause-effect relations and the presence of crystallized standards of desirable performance in jobs influences the HRM control systems selected by organizations for those jobs. It would appear that these are characteristics of jobs that vary within organizations. Focusing on HRM systems at the level of the job in organizations may, therefore, provide greater insight into how organizations can gain a competitive advantage through people. Researchers should collect information on multiple jobs in organizations and then test whether HRM systems influence performance to a greater extent among the core work force.

Construct Definition of HRM Practices

As Becker and Gerhart (1996) noted, there is little agreement as to the exact practices that make up a coherent HRM system. This is compounded by the fact that some of the most commonly mentioned practices could be used to elicit different role behaviors under different circumstances. For instance, incentive pay, or pay-for-performance, is sometimes believed to support one strategy and at other times another. It may be because incentive pay can elicit different types of role behaviors depending on the exact behaviors or outcomes the organization is rewarding.

Becker and Gerhart (1996) were among the first to point out that what researchers call HRM practices has varied across studies. They identify three levels of HRM systems at which researchers have measured HRM practice related concepts. They argued that much of the research to date has focused solely on HRM practices, but it may be equally important to focus on HRM policies and what they termed the HRM system architecture. What they draw attention to is an issue of construct definition.

The problem created by this issue is best understood through example. Researchers wishing to measure organizational-level staffing practices, for instance, must make choices in the staffing-related constructs they choose to measure. At one level one may be concerned with whether organizations are selective in hiring employees (Snell & Dean 1992). At a slightly lower level one may be interested in whether the organization uses valid selection devices, or whether it has a low selection ratio. Going even lower one could be interested in knowing whether the organization uses a particular type of selection device,
such as a situational interview. The level at which one finally decides to take measures has important implications that have only briefly been discussed in the literature. Each of these measures is assessing a different staffing-related construct. The construct ultimately measured should be chosen based on the chosen theoretical rationale. In the literature, however, these measures are often all described as measures of HRM practices.

Most SHRM research is focused on predicting organization- or facility-level outcomes. Given this focus, I believe that it is important to focus on measuring higher level systems constructs, in addition to measuring individual HRM practices. This is particularly true given the complications in using combinations of individual HRM practices to construct HRM system measures. Snell and Dean (1992) provide an example of one method of assessing a higher level HRM system construct, that of selective staffing. Instead of asking respondents whether their organizations used cognitive ability tests, or situational interviews and whether they had a low selection ratio, Snell and Dean (1992) simply asked respondents to report on issues such as “How extensive is the employee selection process for a job in this unit?” This selective staffing measure is broader in scope than measures of individual HRM practices. This may allow better measurement of the underlying construct than those obtained by combining individual staffing practices, particularly unit weighted additive combinations. Two organizations using quite different selection practices that in each case lead to selection of highly skilled employees may receive very different scores on a selective staffing composite created by combining individual practices. Using the scale developed by Snell and Dean (1992), however, they may score equally on selective staffing. I see this problem as similar to that encountered in multi-level research when a researcher attempts to combine individual-level measures to assess a group-level phenomenon.

This approach, however, is not without problems. One criticism of the approach taken by Snell and Dean (1992) is that it is too subjective a measure of staffing practices. It is clearly somewhat more subjective than the approach used to measure HRM practices by Ichniowski et al. (1997). Before abandoning the idea of measuring individual HRM practices, however, it is important that researchers investigate the construct validity of these higher level systems measures. Delery, Gupta, and Shaw (1997), for instance, presented evidence that a measure of selective staffing adapted from Snell and Dean (1992) was strongly correlated with both the number of valid selection devices employed by an organization and their selection ratio. Firms scoring higher on selective staffing used a greater number of valid selection devices and had lower selection ratios. This provides initial evidence that researchers may be able to measure higher level system characteristics directly, rather than inferring them from individual HRM practices, some of which may be substitutes for one another or act synergistically.

Although direct measurement of higher-level system characteristics may provide better measures of HRM systems, researchers are still left with the question of which level to measure. The example above is focused on staffing
practices alone. The broad selective staffing measure may be a valid measure of the staffing system, but clearly not of the entire HRM system. The only solution to this problem may be better theory development.

Measurement of HRM Systems

Incorporating the four types of relationships among practices (additive, substitutable, positive synergistic, and negative synergistic) in influencing firm resources has important implications for the measurement of HRM systems in SHRM research. The majority of empirical studies investigating the relationship between HRM systems and performance have measured a large number of individual practices and either created a scale or index “system” measure by summing scores on groups of practices identified through some form of factor or principal components analysis (Huselid 1995; Snell & Dean 1992; Tsui et al. 1997; Youndt et al. 1996), or classified firms as having a particular type of HRM system based on the results of a descriptive cluster analysis procedure (Arthur 1992, 1994; Ichniowski 1997). Both of these approaches are based on assumptions that may limit their usefulness in SHRM research.

Scales and indexes. Several researchers have attempted to measure HRM systems by combining scores on individual HRM practices into a scale score or index (Becker & Huselid 1998; Huselid 1995; Snell & Dean 1992; Tsui et al. 1997; Youndt et al. 1996). One of the more common procedures researchers have used to identify which practices should be combined into a single measure is through the use of factor or principal components analysis.

Researchers in the organizational sciences have used various factor and principal components analyses for a very long time. Often, researchers have several measures that they assume to be equivalent indicators of an underlying construct of interest and they use factor analysis to explore this possibility. Other times researchers have several measures of variables that are conceptually related and the researcher is interested in reducing the number of variables to enhance interpretability of analyses and to reduce problems associated with multicollinearity. In either case, some form of factor or principal components analysis is performed to discover the underlying factor structure of the data. As Kim and Mueller (1978, p. 8) point out, “factor analysis assumes that the observed variables are linear combinations of some underlying (hypothetical or unobservable) factors.” While principal components analysis is technically distinct from factor analysis (Bentler & Kano 1990; Dunteman 1989; Snook & Gorsuch 1989; Velicer & Jackson 1990), it is quite similar and has often been applied in a manner consistent with that of factor analysis (e.g., Huselid 1995).

In the SHRM literature, Huselid (1995) and Tsui et al. (1997) each used factor analysis procedures to group individual HRM practices into scales. Scores for each organization or unit were calculated for each factor as the average of the scores on the individual HRM practices that loaded most strongly on that factor. A similar procedure was used by Youndt et al. (1996) and
MacDuffie (1995) although the grouping of practices was derived based on theoretical rationales rather than factor analysis. What these studies have in common is that they computed measures of “systems” of HRM practices by grouping practices together and then taking the average or sum of those practices. These scales or indexes were then used to estimate the effect of the HR system on performance.

Researchers in this area have begun making distinctions between and acknowledging the differing assumptions of scales and indexes. As DeVellis (1991, p. 9) pointed out, using scales assumes the responses to items are “caused by an underlying construct.” An index, however, is made up of items that “determine the level of a construct.” The importance of this distinction is apparent when one investigates how these combinations of practices, either scales or indexes, are to be used. If one assumes that the HRM practices are caused by an underlying construct and are equivalent measures of this construct, they should have relatively high intercorrelations. In this case, Cronbach’s alpha is an appropriate indicator of the reliability of such a measure. On the other hand, if one assumes that these practices determine the level of a construct (e.g., work force skills), then they should be combined in a linear fashion, weighted by their influence on the construct and they may not be highly correlated. In a very real sense, the researcher is creating a regression equation and computing a predicted level of the outcome construct. Under these circumstances, it is not assumed or expected that the practices should have high intercorrelations and Cronbach’s alpha is not appropriate.

Huselid (1995) provides an example of how factor analysis has been used in SHRM research. In that study, Huselid (1995) conducted an exploratory factor analysis of HRM practices that resulted in the extraction of two factors. For example, one of the factors was labeled employee motivation. One could argue that this factor should be interpreted as an employer’s concern for employee motivation and it is this concern that led to the adoption of the practices loading on this factor. In fact, that the practices load together on this factor, suggests that organizations adopt these practices simultaneously. Huselid (1995, p. 647), however, stated that that these practices are “designed to recognize and reinforce desired employee behaviors.” This seems to be an argument more in line with treating this group of practices as an index rather than a scale. Youndt et al. (1996), however, explicitly combined measures of practices into indexes based on prior theory that indicated how the practices influenced firm outcomes. This is apparent in their labeling of one of the indexes, “human-capital-enhancing.” They stated that “such an additive approach to combining HR practices into an index suggests that firms can improve performance either by increasing the number of practices they employ within the system or by using the practices in an HR system in a more comprehensive and widespread approach” (p. 849). What they really argued here is that HRM practices can influence firm performance by increasing the human capital of the firm. Thus, the index they created could be interpreted as a predicted level of human capital in the firm. A measure of internal consistency, like that reported in the
study, is inappropriate. This may explain the relatively low internal consisten-
cy measures reported in that study and other similar studies (Huselid 1995; MacDuffie 1995).

It is important to note that many HRM practice measures are quite different than scale items typically used in industrial/organizational psychology to mea-
ure such attitudes as job satisfaction or organizational commitment. Items in those types of scales are written to measure the underlying construct. HRM 
practices, on the other hand, are activities an organization engages in to help it achieve goals. While there may certainly be an underlying logic or philosophy 
guiding the organization in choosing practices, I do not believe that we can assume that the practices are equivalent measures of this underlying philoso-
phy.

This becomes apparent when one acknowledges that some practices may be substitutes for one another in causing a particular outcome, or that some practices may influence the outcome in a synergistic way. In other words, two organizations both equally concerned with employee skills may adopt different staffing systems that have identical results. I believe that this distinction is a strong argument for treating the combinations of HRM practices as indexes rather than scales. Acknowledging these relationships also questions the unit 
weighting approach to combination employed by all of the studies thus far using either the scale or index approach.

The unit weighted additive approach used to combine practices into an index is appropriate only when practices are actually additive in their effects on firm resources and each practice has an equivalent effect. If any of the practices in the index are substitutes for one another, or have a synergistic relationship in producing their outcome, the level of that outcome in the firm is either overestimated or underestimated. For instance, if two practices are substitutable in their effects on work force skills, a firm would receive an index score of two if they employed both practices. Another firm that only employed one practice would get an index score of one, however, they would achieve the same level of work force skills and probably for lower costs. In this case, the level of work force skills of the firm with both practices is overestimated. If two practices act synergistically to create work force skills such that they have no influence when used independently, but achieve high levels of skill when used in combination, this could lead to an underestimate of the skills created in firms that use both practices. It is, therefore, essential to understand how the HRM practices in combination influence firm resources. Assuming an additive approach is not appropriate under all conditions.

Classification Techniques. While some researchers have opted to arrive at a HRM system measure by computing a scale or index, others have borrowed from the extensive work associated with configurational theory (Meyer, Tsui, & Hinings 1993) and have used cluster analysis to classify firms based on the components of their HRM system. In fact, Becker and Gerhart (1996) recently called for greater consideration to be given to cluster analysis as an empirical approach to identifying systems of HRM practices. This call was at least partly
because of some of the concerns with principal components analysis and factor analysis noted above. Aldenderfer and Blashfield (1984, p. 7) described cluster analysis as "a multivariate statistical procedure that starts with a data set containing information about a sample of entities and attempts to reorganize these entities into relatively homogeneous groups." There are several different empirical techniques for conducting cluster analysis, however, they have in common the attempt to group or classify cases in a data set based on several characteristics of these cases. In SHRM research cluster analysis is a very appealing technique because it enables a researcher to discover the underlying patterns or systems of HRM practices that firms have adopted. The question then becomes which HRM systems are more effective than the others? The researcher avoids the problems associated with creating scales or indexes discussed above. There are, however, other potential problems that limit the usefulness of cluster analysis in SHRM research. Many of these concerns are also being noted by configurational researchers in the area of business strategy and structure (Ketchen, & Shook 1996).

Cluster analysis is simply a descriptive empirical technique that looks for similarities across cases. While this could be quite useful in exploratory phases of research, it is clearly less useful when testing theoretical frameworks. Classifying firms into groups using cluster analysis procedures, however, is a descriptive tool not without merit. Proponents of the technique argue that it does not assume linear relationships between practices as factor analysis does (Becker & Gerhart 1996). This is an important issue. If one believes that practices can be substitutable, additive, or synergistic, and that an effective HRM system may be composed of a group of practices where all of these relationships may be present, then cluster analysis may offer several advantages. In any set of firms it can be expected that a subset will have discovered and implemented a truly high performance HRM system of practices. By conducting a cluster analysis, the researcher is able to look more closely at the systems of practices actually used by firms and make judgements as to their effectiveness. If there is a single most effective HRM system and a large enough group of firms have adopted it, it is likely that they will be identified in the cluster routine. If, however, there are multiple effective HRM systems, as suggested above, and there are small numbers of firms actually employing each of them, it may be very difficult for the cluster solution to identify them.

There is no clear rule established to decide on the number of clusters that are present in the data. Arthur (1992), for instance, arrived at six clusters from the clustering routine. He then reduced that number to two by subjectively combining similar clusters. This is at best an art. Deciding on the exact number of clusters becomes critical when either a very small number of firms have adopted an effective HRM system or the most effective system is very similar in most respects to other less effective systems. First, when there are a very small number of firms employing the most effective HRM system, these firms may easily be collapsed into a larger cluster as the researcher reduces the number of clusters. Second, this effect is compounded when the most effective system is nearly identical to a less effective or even poorly effective system.
This could occur when only one practice in the system differs between the two groups, but this practice has a synergistic effect with the others to enhance performance.

Another problem inherent in the way cluster analysis results are often used is that after deciding on the exact number of clusters in the data, a nominal level variable that represents cluster membership is created and used in future analyses. Using this approach, all firms within each cluster are treated as equivalent. The within cluster variance is usually lost at this point. Given that the number of clusters is somewhat arbitrary, and researchers are employing this technique to reduce the data, a small number of clusters (e.g., two, three, or four) is usually chosen (Arthur 1992; Ichniowski et al. 1997). This magnifies the problems created by the lost variance within clusters. In a two-cluster solution firms within a cluster are likely to be much more heterogeneous than in a four or five cluster solution generated from the same data set.

It should be apparent that while cluster analysis appears to have several advantages over other procedures, it is not without serious disadvantages. These are predominantly in how the researcher uses the technique. Deciding the number of clusters is not a trivial task. In addition, it must be acknowledged that there may be several equally effective clusters. This is a result of the different combinations of practices that may achieve the same outcomes. All of this complicates the use of cluster analysis in SHRM research.

**Alternative Procedures.** Given the inherent problems with the approaches discussed above, a few researchers have attempted to develop alternative HRM systems measures. Delery and Doty (1996), for instance, borrowed a technique used in organizational configurational research (Doty, Glick, & Huber 1993). This technique involved identifying a priori internally consistent ideal HRM systems and using profile deviation methods to classify firms into one of these ideal HRM systems. Once a firm was classified, the deviation from the ideal HRM system was then used to predict performance. While this approach has some advantages, such as the a priori construction of the ideal HRM systems based on theory, it also has several weaknesses. These weaknesses include the assumption that it is the fit of all of the practices that matters and that each practice is given equal weight. At the moment, it is not clear if this technique is truly appropriate for continued use in SHRM research.

Becker and Huselid (1998) proposed an additional alternative measure which attempts to quantify the degree to which a firm's HRM system is truly a high performance work system. They have termed it a "homogeneity index." This index is computed by summing the number of high performance work practices a firm has adopted and is using beyond the 75th percentile in the sample. In other words, a firm receives a higher score only by using a practice to a much greater extent than other firms in the sample. Although this measure may capture to some degree the extensiveness of the use of these practices and the internal consistency of the HRM system, it is clearly problematic. This is particularly the case in their data given that Becker and Huselid (1998) acknowledge that some of the elements of their high performance work system
indicators (e.g., formal grievance procedure and a policy of promotion from within) may actually be part of a bureaucratic HRM system. In addition, the 75th percentile cut-off is somewhat arbitrary. At this point it is unclear whether this measure is appropriate for further use in SHRM research. At the very least researchers should explore its assumptions prior to its use.

One final alternative is not to compute a single HRM system measure. For instance, if it is believed that there are a small number of HRM practices or policies that are the critical elements of the system, then researchers could use individual measures of these practices and their interactions in regression equations predicting performance. This approach, however, is still not without concerns. To use this approach it would be important to have a very large sample of firms. Using individual practices and their interactions uses up many degrees of freedom and reduces the power to detect such interactions. Additionally, there could be problems of multicollinearity among the HRM practices and the interactions. Finally, higher order interactions may be impossible to interpret. This technique may show promise, however, if the list of practices can be narrowed down substantially and only a handful of theoretically meaningful interactions are tested (e.g., Delery, Gupta, & Shaw 1998).

From this discussion it is clear that each of the methods used to compute a measure of an HRM system is problematic. I believe that this is a critical issue for SHRM research to explore and encourage future research that addresses this issue.

Mediating Variables

The consensus among SHRM researchers is that HRM practices and systems do not lead directly to firm performance. Rather they influence firm resources, such as the human capital of the firm, or employee behaviors, and it is these resources and behaviors that ultimately lead to performance (Wright et al. 1994). This implicit model assumes that there are mediating variables between HRM practices and firm performance, yet to date, few if any researchers have measured these mediators or adequately addressed their importance.

It is important that future research attempt to specify the mediators and attempt to collect measures of these constructs. This issue is crucial for continued theoretical development in SHRM. By putting greater attention on these variables, we may be able to better test how HRM practices influence these constructs, and ultimately firm performance. In fact, focusing on these mediators may help SHRM researchers identify systems of HRM practices that produce them. Doty and Delery (1997) argued that HRM practices influence firm performance by creating a work force that is skilled, motivated, and empowered. Using this framework, researchers would focus attention on describing how HRM practices influence these work force characteristics and how these characteristics can be measured. This should help in identifying and measuring HRM systems.

Regardless of the specific framework, researchers need to attempt to mea-
sure mediating variables or more fully address their importance if we are to understand the role that HRM practices play in competitive advantage. Given that this research must take place across a large number of firms or facilities, measuring such mediators may be extremely difficult. One approach, although not optimal, is to have key respondents rate the level of the mediating variable in the firm. Wright, Smart, and McMahan (1993), for instance, had NCAA basketball coaches rate the level of several skills that their teams possessed. Other more direct measures are preferable, but much harder to obtain.

A noteworthy attempt at measuring mediating variables was presented by Tsui et al. (1997). In that study, the employee-organization relationship, as measured by HRM practices, was found to be a significant predictor of employees' performance, citizenship behavior, and affective commitment. Although the data for that study were collected across a small number of organizations, they were collected across a number of jobs within each. This research design is appealing in that it may be difficult to collect such data across a large number of organizations. The major problem with this design, however, is that the researcher is unable to collect organizational-level performance measures. Future research should attempt to combine the approaches of Tsui et al. (1997), where jobs are studied across a small number of organizations, with those of Huselid (1995), where organization-level HRM systems are measured across a large number of organizations.

Until more attention is given to the identification and measurement of mediating variables, it is unlikely that effectiveness of HRM systems will be completely understood. It is this lack of consensus about the processes through which HRM systems influence firm-level outcomes that highlights the fact that there is casual ambiguity. This coupled with the problems associated with identifying internally consistent HRM systems greatly limits the ability of firms to replicate a competitor's HRM system or valuable resources that are created by it.

Dependent Variables

Some would argue that what sets SHRM apart from the rest of the research in HRM is the focus on organizational performance. Although much of SHRM research has measured some form of organizational performance (Delery & Doty 1996; Huselid 1995), other researchers have measured performance in only part of an organization (Arthur 1994; Ichniowski et al. 1997; MacDuffie 1995). At the beginning of this paper I stated that HRM practices have been shown to be related to turnover, productivity, financial returns, survival, and firm value. This is evidence that organizational performance has been operationalized in many different ways. If researchers are most interested in financial returns to organizations, then they must specifically address how the various performance measures lead to these returns.

Huselid (1995) is one of the only studies to date that has measured multiple outcomes. In that study, HRM practices were shown to be related to turnover, productivity, and firm value. He argued that at least part of the financial
returns from adopting a high performance work system came from reduced turnover. This implies a simple linear relationship between turnover and financial performance. Arthur (1994), however, found that the influence of turnover on productivity was a function of the HRM system in place. Likewise, Abelson and Baysinger (1984) argued that each firm has an optimal level of turnover given its unique situation. Turnover rates below or above the optimal rate should result in poorer financial performance.

The issue is whether all the dependent variables measured in SHRM research are interchangeable measures of organizational performance. At times it appears that researchers have argued that certain HRM systems lead to just about every positive outcome for organizations (Tsui et al. 1991). Clearly, HRM systems in different parts of the organization lead to more proximal outcomes than financial performance (Becker & Gerhart 1996). It is important that researchers focus more attention on measuring these outcomes and identifying specifically how they result in financial performance. Simply measuring financial performance and testing its relationship to HRM systems may be important and provide some information. This, however, gives very little insight into the specific mechanisms through which HRM systems have their effects.

CONCLUSIONS

In this paper, several issues regarding the two dominant “fit” issues in SHRM research were addressed. Although a few studies have found virtually no support for either type of fit (Delery & Doty 1996; Huselid 1995), there are a small but growing number of studies supporting some notions of fit. For instance, some research shows there is growing evidence that the internal fit HRM practices is essential. Ichniowski et al. (1997) provided relatively strong evidence of the synergistic effects found in tightly coupled systems of practices.

Despite this small but growing support for the effectiveness of the horizontal fit of HRM practices, there is still much to be learned. Several issues for future research were discussed in this paper. First, researchers need to assess HRM systems throughout the organization. Research has demonstrated that HRM systems differ within organizations (Tsui et al. 1997) yet most research has only measured either the average system used by an organization (Huselid 1995) or the HRM system for only one job or group of jobs (Arthur 1992, 1994; Delery & Doty 1996; Ichniowski et al. 1997; MacDuffie 1995). Second, the focal independent variables in SHRM research, HRM practices, have been ill-defined. Researchers must more clearly define the HRM practice and systems constructs to fully understand the underlying processes at work. Third, the assumptions underlying the techniques some researchers have used to combine individual HRM practices into systems measures have not been fully explored. Clarifying these assumptions and developing better methods to identify HRM systems may be the most critical issue for the future of SHRM research. Fourth, it is clear from SHRM research and theory that HRM practices and systems influence organizational performance through their influ-
ence on other organizational characteristics, possibly the resources of the organization. Yet, few studies have attempted to measure any of these mediating constructs. Fifth, research in SHRM has focused on organizational effectiveness and empirical studies have measured effectiveness in many different ways. While it is helpful to have multiple measures of effectiveness, researchers must more clearly articulate the relationships among these measures. An HRM system that minimizes turnover, for instance, may not maximize productivity. Much greater attention to these five issues is needed over the next decade. Below, in hopes of stimulating research on these issues, I outline several suggestions to guide future research in SHRM.

Although no one study will probably be able to incorporate all of the issues addressed above and adequately address each, there are a number of features that studies should incorporate. The ideal SHRM study would have the following characteristics. First, this study would take place within an industry across a large number of organizations. Although inter-industry studies may provide some evidence of the effectiveness of HRM practices, it is doubtful such research will identify meaningful strategic contingencies across such a wide range of competitive markets. By focusing in a single industry, the researcher is better able to gain much needed insight into the competitive market and identify strategic issues that influence the management of human resources. Second, because HRM practices vary across jobs within an organization, it is imperative that measures of HRM practices are obtained for multiple jobs or groups of jobs. Third, multiple HRM systems constructs must be measured. Not only is it important to know which specific HRM practices are in use for each job, but also to get measures of other levels in the HRM system architecture. This would help address the problems of combining practices into a systems measure. Fourth, because it is believed that HRM practices influence organizational resources, researchers must attempt to measure these resources. If it is hypothesized that HRM systems influence the human capital of the firm, it may be necessary to obtain some measure of that human capital. Fifth, multiple outcome measures that would allow better tests of the processes through which HRM systems influence organizational outcomes must be obtained. These measures would include turnover, productivity, returns, and profitability. It would then be possible to examine the tradeoffs inherent in choosing a HRM system to maximize any one of these outcomes.

While such a design may answer the most critical questions in SHRM, it would be difficult to conduct. It would be necessary to obtain the cooperation and participation of a large number of organizations and individuals. Alternatively, researchers could answer many questions by working with a large organization that has many independent operations, or a small group of organizations that have many similar operations. In any case, it is doubtful that a single study will be able to answer the wide range of questions in SHRM.

In conclusion, there is growing evidence that the management of human resources is critical to productivity and performance, however, research is a long way from explaining the underlying processes through which it has these effects. Although researchers have developed more complete frameworks that
better incorporate ideas from the areas of business strategy and finance, the empirical literature is relatively weak and underdeveloped. This is not unexpected at this early stage of SHRM research. It is quite possible that researchers have attempted to do too much in individual studies. As Dyer (1985) suggested, SHRM research encompasses a wide range of hypotheses and because of this evidence will have to be obtained incrementally. There are several issues that must be addressed if research is to provide more meaningful advice to firms. Researchers need to pay more attention to how HRM practices influence valuable, rare, and inimitable organizational resources, such as the work force. It is through closer attention to these issues that we may truly find the answer to the question "Does fit matter?"

REFERENCES


