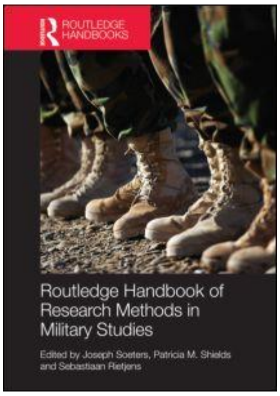

Routledge Handbook of Research Methods in Military Studies

Joseph Soeters, Patricia M. Shields, Sebastiaan Rietjens

Multilevel Analysis

Publication details
https://www.routledgehandbooks.com/doi/10.4324/9780203093801.ch18
Irina Goldenberg, Joseph Soeters
Published online on: 09 Jun 2014

18
MULTILEVEL ANALYSIS
The examination of hierarchical data in military research

Irina Goldenberg and Joseph Soeters

J.M. Schaubroeck, S.T. Hannah, B.J. Avolio, S.W.J. Kozlowski, R.G. Lord, L.K. Trevino, N. Dimotakis and A.C. Peng (2012) 'Embedding ethical leadership within and across organizational levels,' *Academy of Management Journal* 55(5): 1053–1078.

Like most organizations the military is made up of hierarchical layers. Starting at the bottom layer, the organization consists of individual soldiers, teams or squads, platoons, companies, battalions, brigades and sometimes divisions. One level is nested into another, together constituting the military organization.

Most organizational research, including military research, focuses on what is happening among samples of employees belonging to organizational units at the same level, which in this case translates into soldiers belonging to squads or platoons. Horizontal or within-unit research, however, disregards influences on individual behaviour that stem from other – 'higher' – organizational levels. To get a better understanding of what is happening in complex organizations one needs to study cross-level linkages in addition to within-level influences.

The study by Schaubroeck and associates examines the impact of ethical leadership and culture in the US Army, at the squad, platoon and company level, on the ethical behaviour of soldiers in those squads, platoons and companies. On the basis of theoretical and practical insights, it was assumed that ethical leadership and culture not only impact soldiers' behaviour within the same hierarchical layer, but also across layers, directly through trickle-down or bypassing effects, indirectly through higher-level leadership on lower-level ethical culture, or by moderating the within-level relations.

Soldier's ethical behaviour was measured via the perceived frequency of transgressions against non-combatants (e.g. mistreatment of bystanders or causing unnecessary damage) and against the Army (e.g. stealing), as well as with respect to peer exemplary behaviour and moral efficacy. Ethical leadership was measured by means of perceptions of the leader's style of discussing ethical issues

(continued)

(continued)

and setting examples, whereas ethical culture was measured by scoring on standard practices, for instance issuing penalties, in the squad, platoon or company. To assess these variables, soldiers in the respective squads rated squad ethical leadership and culture, the squad leaders rated the platoon leadership and culture, whereas platoon leaders and sergeants rated the ethical leadership and culture of their company.

The data were collected in a cross-sectional survey of 2,048 US soldiers engaged in combat operations in Iraq in 2009. The questionnaire data was aggregated and combined with multiple other ratings into statistical information on 172 squads (consisting of the scores of at least 4 members), 78 platoons (consisting of at least 2 squads) and 40 companies (consisting of on average 4 platoons). These data were used for elaborate multilevel model testing.

The results largely confirmed the hypotheses. In particular, the analyses demonstrated that ethical leadership and culture show strong horizontal, within-level effects (i.e. soldiers within the same squads have more similar scores compared to soldiers from different squads). This is the well-known immediate supervisor effect. In addition, the data revealed important direct vertical effects of leadership at one level on the adjacent lower-level leadership, which is the cascade- or trickle-down effect. There were strong indirect effects as well, indicating that ethical leadership at the company level exhibits significant effects on lower levels, including effects on ethical culture at both the platoon and squad levels. Another important finding was that ethical leadership at a lower level had a stronger positive influence on ethical culture at that level when the leader at the next higher level was reported to exhibit a high level of ethical leadership. This was a notable moderating effect of ethical leadership at higher levels. Apparently, leadership at the higher levels can facilitate or reinforce lower-level leadership, whether in a positive or in a detrimental direction.

This large-scale, complex study underlines the importance of the combined impact of ethical leadership and culture within and across hierarchical levels, i.e. throughout the whole military organization. Clearly, the 'final proof' could not been given, as this would have required a longitudinal, experimental design that would have been replicated a number of times in different contexts. Yet, the findings of this study are unequivocal. Their practical implications are clearly evident and can be instructed in courses and training programmes in military organizations all over the world.

In many research applications such as our ethical leadership illustrative study, data is hierarchically structured and consists of lower-level observations nested within higher levels. Multilevel models are developed for analysing these hierarchically structured data. Consider Figure 18.1: soldiers are *nested* in squads, squads are nested in platoons, which are nested in companies or more generally in larger military environments (Army, Navy, Air Force). In such cases there is often a significant relationship not only between individual-level variables and the outcome or construct of interest, but also between group level variables and these outcomes, as well as between the individual and group level variables themselves. As in our illustrative study, this suggests that understanding the effects of both the individual and group level or contextual variables, as well as how they work together, is important to understanding the phenomenon under study (Rousseau 1985; George and James 1994; Hox 2010).

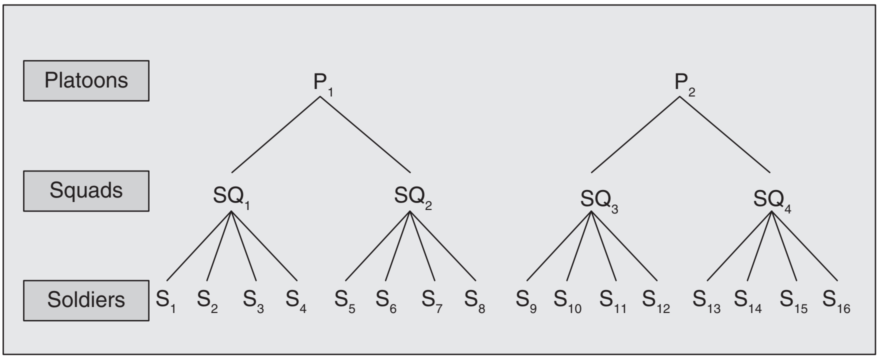

Figure 18.1 Hierarchical structuring of military organizations

Accordingly, multilevel data analysis is becoming increasingly popular in many fields and disciplines, including military research, because it allows researchers to take into account how individual-level processes or outcomes operate within different levels of analysis, such as organizations, geographical locations, time frames, or other higher-level units. For example, soldiers' organizational commitment may be predicted by perceptions of organizational support and satisfaction with leadership at the individual-level of analysis, leadership style at the unit level of analysis, and organizational culture at the service/environment (Army, Navy, Air Force) level of analysis. Further, it is then also possible to model both within-level interactions among predictors (e.g. between individuals' perceptions of organizational support and individuals' satisfaction with leadership), as well as cross-level interactions (e.g. between unit leadership style and organizational culture), as shown in the ethical leadership illustrative study.

Quantitative and qualitative applications

The term multilevel data analysis or multilevel modelling refers to a set of related approaches for analysing quantitative data measured at two or more levels of analysis. Although quantitative analyses are predominant in this field of research, qualitative multilayer studies are recognized as important. In his famous study about the accidental shootdown of US Black Hawks over Northern Iraq, Snook (2000) was not satisfied with either individual-level accounts (why did the F-15 pilots misidentify the Black Hawks?), the group level account (why did the AWACS fail to intervene?) or organization-level account (why wasn't the Army aviation detachment integrated into task force operations?). Instead of just listing the separate layer-explanations, which are interesting and important themselves, Snook developed a theory that explains the mishap in terms of cross-level mechanisms. His 'practical drift' theory emphasizes the slow, steady uncoupling of practice from written procedures, occurring throughout the whole organization and leading to this tragic event in the military that is less exceptional than one would perhaps expect.

Despite this impressive multilevel qualitative study, multilevel research is predominantly developed and applied to quantitative research. Multilevel models are called by various names such as hierarchical models, mixed models, cluster models, growth curve models, contextual models, and random coefficient models. These models extend traditional linear models to take into account situations in which individual-level data is clustered or nested into a higher-order structure. The lowest-level measurements are often said to be *micro-level*, and all higher-level measurements are said to be *macro level* (Kreft and de Leeuw 1999). In the basic two-level linear model, micro-level data or level 1 units (e.g. soldiers) are nested within macro level or level 2 units (e.g. platoons), and variables from both levels of analysis are included in the mode l (Bliese and Jex 2002).

Multilevel models in the most basic form are generally regression models that are linear in their coefficients and can be expressed in two algebraically equivalent forms. They may be expressed as an equation relating a micro-level outcome to a set of micro-level variables along with a set of equations in which the coefficients of this micro-level model are expressed as functions of macro level variables. Alternatively, these multilevel models may be expressed in a single equation where the micro-level outcome variable is expressed as a function of both micro and macro variables. This second form generally includes interactions between the micro-level and macro level variables, or cross-level interactions, to be discussed below (DiPrete and Forristal 1994).

Several terms related to variable level are important to note in the consideration of multilevel data analysis (Chan 2006). *Global variables* are those that are measured at their natural level (Hox 2010). For example, soldier years of service and gender are global variables at the individual-level, and platoon size is a global variable at the group level. It is of note that variables measured at a given level may be "moved" to a higher level through *aggregation*, which entails grouping lower-level units to form a smaller number of higher-level units. For example, soldiers' morale scores may be aggregated by computing a mean morale score to form unit morale scores. Such new variables aggregated to a higher level from a lower level are referred to as *analytical variables*. Similarly, variables can be moved from a higher level to a lower level through *disaggregation*, which entails decomposing variables at a higher level into a larger number of lower units. For example, in the disaggregation of platoon size to the soldier level, each soldier within a platoon is assigned the same value for platoon size. This creation of a new variable resulting from disaggregation provides information on the higher-level context (platoon) to the lower-level units (soldiers), and thus such variables disaggregated to a lower level from a higher level are referred to as *contextual variables* (Chan 2006). However, it is important to note that creation of such variables should be done judiciously with significant concern for construct validity. A detailed framework specifying functional relationships between constructs at different levels that may be used for the composition of such variables is discussed in Chan (1998).

Relations between the levels of analysis

Traditional methods of data analysis, such as ordinary least squares (OLS) regression models, that attempt to include both individual and contextual-level variables are not well-suited to understand these multilevel effects. In particular, the clustering of individual-level observations within these higher-level or contextual units violates the assumption of independent errors, which leads to biases in both the parameter estimates and standard errors (Bliese et al. 2002). Of course, individual-level observations within the same group are not truly independent because there is some underlying similarity resulting from the group membership that leads to dependence among the observations and the errors within these group level units (Fullerton et al. 2007).

A significant advantage of multilevel modelling then is that independence of observations is not required, and in actuality, independence is often violated at each level of analysis (Tabachnick and Fidell 2007). For example, recruits within Army combat training sessions are likely to influence each other or be influenced by their instructors, and are therefore likely to be more similar than recruits in other Army combat training sessions. Similarly, recruits within

Army training schools are likely to be more similar than recruits in different training schools. In addition, there may be interaction across levels of the hierarchy (or cross-level interactions, as in our illustrative study and as discussed below). For example, student characteristics within Army training sessions may interact with the approaches of instructors at the session level. As such, ignoring the hierarchical or nested nature of the data and analysing it as if it was on the same level can lead to both interpretational and statistical errors (Chan 2006; Dansereau et al. 2006; Snijders and Bosker 1999).

One of the main types of interpretive error is the *ecological fallacy*, also referred to as *the Robinson effect*, whereby relationships at the group level are thought to imply relationships at the individual-level (Hox 2010). For example if data on military services (e.g. Army, Navy, Air Force) is used to make inferences about individuals (e.g. soldiers, sailors, airmen or air-women). Failing to acknowledge the within-group variability that is present in the data can distort the relationships under examination (Heck and Thomas 2009). The ecological fallacy may be either positive or negative. A positive ecological fallacy occurs when a relationship at the group level (e.g. military families in communities where Military Family Resource Centres (MFRCs) provide a greater number of services are healthier) is used to make conclusions about the same relationship at the individual-level (e.g. military families that use MFRC services more frequently are healthier). A negative ecological fallacy occurs when a lack of relationship at a group level (e.g. there is no relationship between the proportion of soldiers in a unit that suffer from posttraumatic stress disorder and units' rates of attrition) translate into conclusions about a relationship at the individual-level (e.g. there is no relationship between posttraumatic stress disorder and attrition from the military). In this case the group is erroneously used as the unit of analysis, which results in a lower n (based on the number of groups instead of the number of individuals within each group). Statistically, this type of analysis leads to increased standard errors and reduced statistical power, and therefore a greater likelihood of committing Type II error, or erroneously failing to observe an effect when it actually exists (Chan 2006).

The other main type of interpretative error that may result from ignoring the hierarchical nature of data is the *atomistic fallacy*, sometimes referred to as the *individualistic fallacy*, whereby conclusions about group level variables are inferred based on data collected at an individual-level. For example, soldiers who have been deployed more frequently have a stronger warrior (versus peacekeeper) identity, therefore leadership in platoons that deploy more often place greater emphasis on the warrior identity. Statistically, Type I error (i.e. erroneously concluding that an effect exists when in fact it does not) is inflated if analyses performed at the lower level are used to make inferences at the group level because analyses are based on too many degrees of freedom that are not actually independent, and standard errors are therefore erroneously reduced resulting in an overestimation of the precision of the parameters of interest (Heck and Thomas 2009).

Intra-class correlations

As discussed, in hierarchically nested data it is often assumed that individuals or first-level units within the same group or second-level unit are more similar to each other than individuals in different second-level units. This homogeneity or similarity of individuals within groups can be measured by calculating the *intra-class correlation* (Bliese et al. 2002). Computing these intra-class correlations at the squad, platoon and company level was a major first step in the data analysis of the illustrative study on the impact of ethical leadership and culture. The range of these correlations indicated that the data at the various levels were suitable for aggregation to the adjacent, higher levels (Schaubroek et al. 2012: 1063–1064).

A high intra-class correlation indicates that individuals within groups are homogeneous or that the groups are very different from each other. In general, a low intra-class correlation indicates that the groups are only slightly different from each other. An intra-class correlation of zero means that no group differences exist on the variable of interest, and that individuals within the same group are as different from one another on this variable as individuals across groups. Thus, if the intra-class correlation is zero, clustering the data has no consequence for the relationship between the variables of interest and can be ignored in further analyses. Conversely, if there is a substantial intra-class correlation, modelling the intra-class correlation is appropriate in that it takes the nested structure of the data into account, and will lead to a better understanding of the phenomenon of interest (Kreft and de Leeuw 1999).

Since the intra-class correlation can also be thought of as a measure of the degree of dependence of individual units (Bliese 2000), the existence of an intra-class correlation indicates that the assumption of independent observations applicable in traditional data-analytic techniques is violated. The more individuals share in terms of common experiences due to closeness in space or time, the more similar or dependent they are likely to be. Further, taking into account the existence of intra-class correlations is important in that these correlations change the error variance in traditional statistical analysis techniques, as alluded to above. Under the assumption of independence of observations, error variance represents the effect of all omitted variables and measurement errors. As such, in traditional analyses omitted variables are assumed to have a random, rather than a structural effect, which, as discussed in this chapter, is often not the case if the data is hierarchical or nested in nature (Kreft and de Leeuw 1999). Of note, intra-class correlation coefficients are often used in decisions regarding aggregation of data (for greater detail refer to Bliese 2000 and Kozlowski and Klein 2000).

Cross-level interactions

Cross-level interactions or influences, as the name implies, entail interaction effects between different levels of analysis, such as interactions between individual-level predictors and group level predictors. When such interactions are present, the effects of the individual-level variable on the outcome is said to be *moderated* by the group level variable. As an extension, the relationship between two individual-level variables on the outcome may be moderated by a group level variable (i.e. the nature of the relationship between the individual-level variables may be different in different groups). For example, research on the effects of deployment on retention of military personnel has yielded mixed results. To clarify these mixed results, it has been suggested that the type of deployment (at the group level) may moderate this relationship, such that deployment may increase retention for individuals deployed in less hostile or peacekeeping operations, whereas deployment may lead to decreased retention in those deployed in more hostile operations (Fricker et al. 2003; Wisecarver et al. 2006). In the ethical leadership illustrative study one could see a larger impact of ethical leadership on ethical culture in the same unit if the higher-level leadership was ethical as well, and a more limited impact if the higher-level leadership was not very ethically oriented. These types of interaction effects can only be tested using models that take the nested structure of the data into account, and specify the estimation of these cross-level interaction effects.

Hierarchical linear modelling (HML)

The two major types of established multilevel data-analytic techniques are multilevel linear hierarchical models (HLM) and multilevel latent variable models. Understanding HLM models requires basic knowledge of multiple regression. Further, understanding multilevel latent variable models requires basic knowledge of latent variable (i.e. covariance structures) analyses (Chan 2006). A detailed discussion of multilevel data-analytic techniques is beyond the scope of this chapter. However, HLM will be presented briefly in this chapter since these represent the most common class of multilevel approaches used in multilevel research (Hox 2000). HLM was also the analytic technique used in the illustrative study by Schaubroeck et al. (2012) on the impact of ethical leadership and culture within and across different levels.

HLM is an extremely useful and popular approach to multilevel data analysis. It allows for the identification and partitioning of the different sources of variance in the outcome variable, and further, provides a means for modelling these different sources of variance using multiple predictors at different levels of analysis. Moreover, HLM provides a powerful tool for assessing both cross-level main effects and cross-level interactions.

As such, HLM is usually used to assess the influence of both individual and group level predictors on an individual-level outcome, as well as the moderating effects of group level variables on the relationships between individual-level variables (Gavin and Hoffman 2002).

More specifically, multilevel modelling deals with potential group effects, and the potential dependence among individual-level observations within groups by allowing intercepts (means) and slopes (the criterion-predictor relationships) to vary between groups. For example, the relationship between recruits' training scores (the criterion) and recruits' aptitude scores (the predictor) is allowed to vary between training schools. This variability is modelled by treating the group intercepts and slopes as criterion variables in the next level of analysis. For example, in the next level of analysis, differences in means and slopes within training schools may be predicted by instructors' leadership between the training schools.

Conceptually, HLM can be thought of as a two-level approach where the Level 1 analysis consists of regressing the outcome (e.g. training scores) onto the criterion (e.g. aptitude scores) separately for each group or training school. The regression equations estimated for each group generate intercept and slope terms summarizing the relationship between training scores and aptitude for each group or training school, and it is not assumed that this relationship is consistent across training schools. The Level 2 analysis in HLM would then assess the degree to which these intercepts and slopes can be predicted by training school. Thus, in this HLM model the main effect of aptitude would be estimated in the Level 1 part of the model, the main effect of training school would be estimated in the Level 2 part of the model through regressing the intercept terms onto training schools, and the cross-level interaction between aptitude and training school would be assessed in the Level 2 model by regressing the slope terms onto training school (Castro 2002; Gavin and Hofmann 2002). Significant cross-level interactions indicate that group level variables moderate the relationship between two individual-level variables (e.g. the aptitude scores-training scores relationship is moderated by training school) because the value of the Level 1 relationship (i.e. the value of the Level 1 regression slope) differs or depends on the value of the Level 2 or group level variable. As with traditional single-level Ordinary Least Squares regression analysis, a main effect should be interpreted with caution when there is the potential for moderation of the relationship (Gavin and Hofmann 2002).

Although the earliest applications of multilevel data analysis focused on two levels in relation to a continuous outcome, this basic model has been extended in a number of ways to include three or more levels of analysis and a variety of different types of outcomes. The researcher should consider the features of the particular data and the overall goals of the research in the selection of the specific multilevel data-analytic approach (Heck and Thomas 2004).

Longitudinal studies: Growth curve analysis

Cross-sectional multilevel models examine individual data nested within higher-level units. In longitudinal analyses, multilevel models examine patterns of repeated measures nested within individuals. Although the lowest level of data in most analyses is usually an individual, in longitudinal designs the lowest level of data is repeated measurements of individuals, with these measurements said to be nested within individuals. Thus, in longitudinal analyses, instead of analysing inter-individual differences in the context of different groups, the primary focus is usually on analysing intra-individual changes over time (Bliese et al. 2007; Han and Andres 2014, Chapter 17 this volume). In such cases, the intra-class correlation measures the degree to which behaviour of the same individual is more similar to his/her previous behaviour in comparison to the behaviour of other people (Kreft and de Leeuw 1999).

Because there are separate analyses of each case over time, individual differences in *growth curves* may be evaluated (Heck and Thomas 2009). For example, do military children differ in their adaptation over the course of their military parent's deployment (e.g. through the pre-deployment, deployment, and post-deployment phases)? If so, are there variables, such as support from other family members or military family services, which predict these differences?

Considerations in multilevel modelling

The role of theoretical guidance and the choice of predictors

Given that a main purpose of multilevel data analysis is to consider predictors at different levels of analysis, the correlations among predictors at all levels of analysis are considered together and adjusted for each other. As a result, it becomes more likely that none of the regression coefficients associated with these predictors will be statistically significant. Thus, it is of key importance to select the right number and combination of predictors to maximize the utility of the analysis in explaining the phenomenon of interest. In particular, it is suggested that only a very small number of predictors be selected, and further, that these predictors are relatively uncorrelated with each other (Tabachnick and Fidell 2007). Of course, strong theoretical and conceptual rationale should be used in the selection of predictors, and will be helpful in selecting a limited number that optimizes explanation of the phenomenon of interest.

Only the most theoretically relevant predictors are included in the model to begin with. Following this, predictors may be added in order of importance, and those that do not improve prediction of the phenomenon of interest are dropped from the analysis (unless they add meaningfully to cross-level interactions) (Raudenbush and Bryk 2001). If there are a large number of potential predictors, they may first be screened using simple modelling techniques, such as linear regression, to eliminate those that do not contribute to explanation of the phenomenon of interest from the start. Further, if the sample is large enough, it is suggested that these exploratory type analyses be cross-validated by using half of the sample to build the model, and the other half of the sample for cross-validation. This will attenuate the degree to which these exploratory model-building techniques are influenced by chance (Tabachnick and Fidell 2007).

Statistical assumptions and sample size

The statistical assumptions and limitations that pertain to traditional data analyses techniques generally apply to multilevel data-analytic techniques as well (Castro 2002). As such, preliminary data cleaning and analyses needs to assess conformity with distributional assumptions as well as outliers.

Ideally, it is recommended that screening of lower-level predictors should be conducted within each higher-level unit. However, this may be impractical, especially when the number of higher-level units is large. In such cases, the lower-level units may be combined over the higher-level units. Likewise, second-level predictors should be examined within third-level predictors if possible, or they may be aggregated across the third-level predictors (Tabachnick and Fidell 2007).

By their nature, multilevel models are generally more complex than traditional models, because they entail calculating a greater number of equations at various levels of analysis. As discussed, in addition to the effects of interest within each level, effects of parameters, including the intercepts and slopes, are also of interest at each level. As such, these types of models generally necessitate larger sample sizes at each level to counteract the instability inherent in such models (Chan 2006). The illustrative study was based on a sample of 2,048 soldiers which was necessary in order to build multilevel models that will then consist of much fewer cases at each subsequent level, such as squads, then platoons and finally companies.

As in most analyses, power increases with increased sample sizes, larger effect sizes, and smaller standard errors. However, there are some more complex issues related to power in these types of analyses that pertain to having effects at different levels of analysis. For example, it has been shown that power grows with the size of the intra-class correlation (i.e. the difference between groups relative the differences within groups, as discussed above), particularly for tests of higher-level effects and cross-level interactions. In general, it has been demonstrated that power is greater with a greater number of groups (or second-level units) and fewer cases per group (first-level units) than the other way around, but that power increases with greater sample size at both levels (Tabachnick and Fidell 2007).

Limitations and cautionary notes

Although multilevel data analysis provides sophisticated methods for analysing complex phenomenon and considering the influence of predictors at different levels of analysis, it is important to recognize the trade-offs between using these types of models as compared to more traditional approaches. One particularly important point for consideration is that although multilevel models may yield more realistic explanations of real-life phenomena, these types of models are generally more complex, and thus may not always be the best approach. Of note, multilevel data analysis usually results in more complex statistical models that are generally more difficult to interpret than simpler models.

Further, the results of complex models are usually more difficult to replicate across samples and across different studies. This is because complex models are more sensitive to changes in what is a more complex system, which may entail a greater number of explanatory variables, measured at multiple levels of the hierarchy, and may include multiple cross-level interactions among variables of different levels, which leads to instability in parameter estimates across models that differ in minor ways (Kreft and de Leeuw 1999). At the least, these types of models are not generally recommended for exploratory data analysis or extensive modifications to increase model-fit (Kreft and de Leeuw 1999).

In addition, it is important to pay attention to the number of higher-level units and not to confuse the number of higher-level units with the total study sample size or the total number of lower-level observations. Although the statistical power of the tests of significance of lower-level estimates is dependent on the total lower-level sample size, the statistical power of estimates for the higher-level estimates and cross-level interactions is based on the number of higher-level units. Because HLM and other multilevel techniques assume large sample sizes it is important to ensure a sufficient number of groups in order to properly test and interpret the effects of multilevel analyses (Castro 2002; Chan 2006).

Finally, although multilevel models are very flexible and allow for the testing of a variety of hypotheses concerning variables and relationships at multiple levels, as well as various cross-level relationships, this flexibility also makes such models more vulnerable to misuse resulting in misleading or erroneous inferences (Chan 2006), such as the ecological fallacy and atomistic fallacy discussed above.

Conclusion

Military research is replete with phenomena that are multilevel, or hierarchical, in nature. As such, military researchers are interested in understanding individuals (or other micro-level units) within their social or organizational contexts. Individuals within higher-level groups or contexts often share common properties or characteristics or are subject to common experiences. Similarly, properties of groups or contexts may also be influenced by the individuals within them. It is clear that multilevel analysis is not only a useful, but in fact is often an essential approach for military research, enabling researchers to understand the phenomena under investigation more accurately and more completely.

References

Bliese, P.D. (2000) ‘Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis’, in K.J. Klein and S.W. Kozlowski (eds), *Multilevel Theory, Research, and Methods in Organizations* (pp. 349–381), San Francisco, CA: Jossey-Bass.

Bliese, P.D., Chan, D. and Ployhart, R.E. (2007) ‘Multilevel methods: Future directions in measurement, longitudinal analyses, and nonnormal outcomes’, *Organizational Research Methods*, 10(4): 551–563.

Bliese, P.D., Halverson, R.R. and Schriesheim, C.A. (2002) ‘Benchmarking multilevel methods in leadership: The articles, the model, and the data set’, *The Leadership Quarterly*, 13(1): 3–14.

Bliese, P.D. and Jex, S.M. (2002) ‘Incorporating a multilevel perspective into occupational stress research: Theoretical, methodological, and practical implications’, *Journal of Occupational Health Psychology*, 7(3): 265–276.

Castro, S.L. (2002) ‘Data analytic methods for the analysis of multilevel questions: A comparison of intraclass correlation coefficients, $r_{wg(j)}$, hierarchical linear modeling, within- and between-analysis, and random group sampling’, *The Leadership Quarterly*, 13(1): 69–93.

Chan, D. (1998) ‘Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models’, *Journal of Applied Psychology*, 83(2): 234–246.

Chan, D. (2006) ‘Multilevel research’, in F.T.L. Leong and J.T. Austin (eds), *The Psychology Research Handbook* (2nd edn) (pp. 401–418), Thousand Oaks, CA: Sage.

Dansereau, F., Cho, J. and Yammarino, F.J. (2006) ‘Avoiding the “Fallacy of the Wrong Level”: A within and between analysis (WABA) approach’, *Group and Organization Management*, 31(5): 536–577.

DiPrete, T.A. and Forristal, J.D. (1994) ‘Multilevel models: Methods and substance’, *Annual Review of Sociology*, 20, 331–357.

Fullerton, A.S., Wallace, M. and Stern, M.J. (2010) ‘Multilevel models’, in K.T. Leicht and J.C. Jenkins (eds), *Handbook of Politics: State and Society in Global Perspective*, Handbooks of Sociology and Social Research (pp. 589–604), Springer Science + Business Media.

Fricker, R.D., Hosek, J. and Totten, M.E. (2003) 'How does deployment affect retention of military personnel?' RAND Corporation (research brief).

Gavin, M.B. and Hofmann, D.A. (2002) 'Using hierarchical linear modeling to investigate the moderating influence of leadership climate', *The Leadership Quarterly*, 13(1): 15–33.

George, J.M. and James, L.R. (1994) 'Levels issues in theory development', *Academy of Management Review*, 19(4): 636–640.

Han, J. and Anders, M. (2014) 'Studying Longitudinal Influences on the adjustment of new recruits in the Chinese military', in J.M.M.L. Soeters, P.M. Shields and S.J.H. Rietjens (eds) *Routledge Handbook of Research Methods in Military Studies*, London: Routledge.

Heck, R.H. and Thomas, S.L. (2009) *An Introduction to Multilevel Modeling Techniques* (2nd edn), Routledge: New York.

Hox, J.J. (2000) 'Multilevel analyses of grouped and longitudinal data', in T.D. Little, K.U. Schnaubel, and J. Baumert (eds), *Modeling Longitudinal and Multilevel Data: Practical Issues, Applied Approaches, and Specific Examples* (pp. 15–32), London: Laurence Erlbaum.

Hox, J.J. (2010) *Multilevel Analysis: Techniques and Applications*, Routledge: New York.

Kozlowski, S.W.J., and Klein, K.J. (2000) 'A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes', in K.J. Klein and S.W.J. Kozlowski (eds), *Multilevel Theory, Research and Methods in Organizations: Foundations, Extensions, and New Directions* (pp. 3–90), San Francisco: Jossey-Bass.

Kreft, I. and de Leeuw, J. (1998) *Introducing Multilevel Modeling*, London: Sage.

Raudenbush. S. and Bryk, A.S. (2001) *Hierarchical* (2nd edn), Newbury Park, CA: Sage.

Rousseau, D.M. (1985) 'Issues of level in organizational research: Multi-level and cross-level perspectives', *Research in Organizational Behavior*, 7: 1–37.

Schaubroeck, J.M., Hannah, S.T., Avolio, B.J., Kozlowski, S.W.J., Lord, R.G., Trevino, L.K., Dimotakis, N. and Peng, A.C. (2012) 'Embedding ethical leadership within and across organizational levels', *Academy of Management Journal*, 55(5): 1053–1078.

Snijders, T.A.B. and Bosker, R.J. (2011) *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling* (2nd edn), London: Sage.

Snook, S.A. (2000) *Friendly Fire: The Accidental Shootdown of U.S. Black Hawks over Northern Iraq*. Princeton, NJ: Princeton University Press.

Tabachnick, B.G. and Fidell, L.S. (2007) *Using Multivariate Statistics* (5th edn), Toronto, ON: Pearson.

Wisecarver, M.M., Cracraft, M. and Heffner, T.S. (2006) *Deployment Consequences: A Review of the Literature and Integration of the Findings in a Model of Retention*. U.S. Army Research Institute for the Behavioural and Social Sciences.