Classroom context and aggression:

Understanding the effects of and change in classroom social structure

Ana M. Velásquez

A Thesis

in the Department

of

Psychology

Presented in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy at

Concordia University

Montreal, Quebec, Canada

October 2010

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Ana M. Velásquez

Entitled: Classroom context and aggression: Understanding the effects of and
             change in classroom social structure

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

complies with the regulations of the University and meets the accepted standards with
respect to originality and quality.

Signed by the final examining committee:

Jon Capobianco	Chair
Wendy Troop-Gordon	External Examiner
Hariclia Petrakos	External to Program
Anna-Beth Doyle	Examiner
Norman Segalowitz	Examiner
William M. Bukowski	Thesis Supervisor

Approved by

Andrew Chapman
Chair of Department or Graduate Program Director

Brian Lewis
Dean of Faculty

Abstract

Classroom context and aggression:

Understanding the effects of and change in classroom social structure

Ana M. Velásquez, Ph.D.
Degree of Doctor of Philosophy
Concordia University, 2010

Peers relations have been regarded as a powerful developmental context that can influence multiple aspects of children's functioning and well being (Rubin, Bukowski, & Parker, 2006). Given that to a great extent these peer interactions take place within classrooms, this particular social context is an essential domain for the study of children's socialization. The present dissertation sought to further our understanding of the multiple effects that the classroom may have on children's social development, particularly for antisocial behaviour. Through a series of three interrelated studies, the classroom was examined as a fundamental context of experience and measurement, as a predictor of individual development, and as dynamic system that changes over time. These issues were addressed in a sample of 1594 Colombian children, who were assessed at four time points during a school academic year. The first study (Chapter 1) describes a statistical technique that capitalizes on the valuable information that peer-nomination-based assessments provide about individual members of the classroom, while ensuring the reliability and validity of the measures. Particularly, a regression-based technique that corrects scores obtained from sociometric and peer assessments to prevent the risk of bias due to differences in classroom size (i.e., number of nominators) is proposed and examined. The advantages of the regression-based technique over the widely-used

standardization technique, in terms of practicality, efficiency, and interpretability are presented. The second study (Chapter 2) examined the effect of classroom cohesion and aggression-related norms on changes in individuals' aggressive behaviours across the school year. Results from this study revealed that classroom norm salience predicted higher and increasing levels of students' aggression. In addition, it was found that classroom cohesion prevented direct forms of aggression, while it enhanced conformity to aggression norms in girls. Finally, the third study (Chapter 3) explored how classroom norm salience changes over the course of the school year and to what extent teachers' cognitions and behaviours explain such changes. Evidence indicates that teachers' support prevented classroom norm salience from increasing over time. Results from these studies are discussed in light of their practical implications for interventions directed at preventing aggression and promoting social competence in the school context.

Acknowledgements

I am truly thankful to Bill, my supervisor, who not only enlightened my professional career with his wise guidance and everlasting encouragement, but also inspired me to be a better human being with his support, sympathy and generosity. Thank you for sharing your precious expertise with me and for believing that I was worth inheriting a piece of it.

Thanks to Anna-Beth for her interest in my dissertation and for her thorough feedback. I greatly appreciate Norman's willingness to be on my committee and the opportunity I was given to participate in discussions we had during my program. Special thanks to Enrique, who motivated me to pursue this goal in my life and who is always willing to make the effort of putting what we have learned from research into practice.

I would also like to thank the Interpersonal Relationships and Development Laboratory, the Centre for Research in Human Development, Concordia University, the Social Sciences and Humanities Research Council of Canada and the Fonds Québecois de Recherche sur la Societé et la Culture, for providing the financial support to pursue my studies.

I want to acknowledge the participation of each one of my lab mates on this achievement, Felicia, Holly, Jonathan, Caroline, Kate, Marie-Eve, and Melissa, who offered me invaluable companionship, support and feedback during this process. I want to express my gratitude to Caro, Primo and all my empathic web friends who followed every progress I made, encouraged me to keep going and celebrated my accomplishments. I also wish to thank Robin and Angie, my dancing buddies, who filled my spirit with joy and energy to keep walking during the most arduous phase of this pathway.

Yiyo, you were my husband, partner, and friend during this journey. Thank you for being by my side, for your confidence in me, for your patience and for all your support. Thank you for our son, Samuel, who is the inspiration of all the good things I want to do to change the world.

Finally, there are no words to express how grateful I am with my best friend, my partner in crime, Lina. Thank God who put us on each other's way to share our thoughts, feelings, goals, obstacles, achievements, failures and dreams. This thesis would not have come true if it were not for the buffering effects of having a friend like you. So let peers be the bridge towards well-being…

Table of contents

List of figures

Chapter 1

Chapter 2

List of tables

Contribution of authors

Chapter 1. Adjusting for group size effects in peer nomination data.

Authors: Ana M. Velásquez, William M. Bukowski, Lina M. Saldarriaga

Initial concept of the chapter:	William M. Bukowski
Ownership of the data:	William M. Bukowski,
	Ana M. Velásquez,
	Lina M. Saldarriaga
Analyses:	Ana M. Velásquez,
	Lina M. Saldarriaga (equally)
Introduction section:	Ana M. Velásquez,

William M. Bukowski (equally)

Methods section:	Ana M. Velásquez
Results section:	Ana M. Velásquez
Discussion section:	Ana M. Velásquez

Chapter 2. The effects of classroom peer norms and cohesion on aggression trajectories:

A short-term longitudinal study.

Authors: Ana M. Velásquez, William M. Bukowski, Lina M. Saldarriaga

Initial concept of the chapter:	Ana M. Velásquez,
	William M. Bukowski
Ownership of the data:	William M. Bukowski,
	Ana M. Velásquez,
	Lina M. Saldarriaga
Analyses:	Ana M. Velásquez
Introduction section:	Ana M. Velásquez
Methods section:	Ana M. Velásquez
Results section:	Ana M. Velásquez
Discussion section:	Ana M. Velásquez

Chapter 2. Predicting changes in classroom aggression-related norms. The role of teachers' normative beliefs and support.

Authors: Ana M. Velásquez, William M. Bukowski, Lina M. Saldarriaga

Initial concept of the chapter:	Ana M. Velásquez,
	William M. Bukowski
Ownership of the data:	William M. Bukowski,
	Ana M. Velásquez,
	Lina M. Saldarriaga
Analyses:	Ana M. Velásquez
Introduction section:	Ana M. Velásquez
Methods section:	Ana M. Velásquez
Results section:	Ana M. Velásquez
Discussion section:	Ana M. Velásquez

General Introduction

It is known already that children's development is affected by the ecologies that surround it. A specific for of influence comes from interactions that occur within and between the different proximal and distal social systems in which the individual is embedded (Bronfenbrenner, 1976, 1992, 2006). One context of particular importance for the child is the classroom, the principal social setting in which children initiate and develop peer relationships. It is well known that peers have a significant impact on child development; they serve various socialization purposes (Bukowski, Brendgen, & Vitaro, 2007), and satisfy individuals' needs (Bass, 1960; Schutz, 1966). Evidence shows that peers are of significant value for the individual (Shaver, 1977) and that they represent the values, beliefs and attitudes from which the individual derives his/her identity (Cotterell, 1996). It is known also that peers are used for social comparison (Levine, 1983; Merton, 1968; Pettigrew, 1967), and they supply social integration and reassurance of one's self-worth through social validation (Tesser, 1988). Given that experiences with peers typically occur with the classroom context this social environment is a fundamental domain for the study of socialization processes.

In studying the classroom context as related to individual development, at least three issues should be considered. First, it must to be acknowledged that, to a great extent, most of child and adolescent development occurs in school contexts. Therefore, the classroom setting is one that developmental researchers have consistently used to assess individual outcomes. In particular, peer relations researchers have capitalized on the rich knowledge that peers have of their classmates to evaluate a variety of individual characteristics. For this reason, procedures such as sociometric questionnaires or peer

assessments, which are mainly based on nomination procedures, have been very popular. In spite of their usefulness, nomination-based procedures have faced the long-standing challenge of controlling for differences in the size of the classroom (i.e., number of nominators) to obtain valid and reliable measures. Therefore, when studying the classroom context an important methodological issue is ensuring that the valuable information that peers can provide about their classmates is free of any possible bias generated by classroom size differences.

The second issue concerns the effects that the classroom context has on individual development. To date, there is a wealth of research showing the impact of the classroom environment on the development of a variety of outcomes in children, including social competence (e.g., Gazelle, 2006; Hoglund & Leadbeater, 2004; Howes, 2000; Rimm-Kaufman, La Paro, Downer & Pianta, 2005). However, there is little agreement as to which characteristics of classrooms should be studied. For instance, educational research usually looks at properties of the learning, social and organizational climate of the classroom (e.g., Fraser, 1994) whereas developmental research often looks at aggregates of individual attributes (e.g., aggression) to characterize classrooms (e.g., Kellam, Ling, Mersica, Brown, & Ialongo, 1998). The present study proposes that classrooms, as social groups, can be defined on the basis of the pattern of interpersonal interactions that take place within classrooms, namely, the social structure (Shaw, 1981). Based on this and on theories of peer relations (Rubin, Bukowski, & Parker, 2006), we propose that the structural characteristics relevant to the study of the social organization of classrooms are behavioral norms, cohesion, and status. This study examines the effect of these features on the development of children's aggression.

The third issue relates to the dynamic nature of the classroom as a social system. In addition to looking at how classroom social structure influences its members' behavior, it is necessary to understand how the structure itself changes over time. If it is true that the classroom social environment influences students' adjustment, then it is necessary to understand how such social structure develops. This is important to elucidate for at least two reasons. First, research suggests that classrooms, as groups, are dynamic systems: they change. This is evident in studies showing that interactions among classroom peers differ across time (e.g., Ramsey, 1995) or in studies that have consistently found that the classroom social climate tends to worsen from the beginning to the end of the school year (e.g., Greenberg, 2007). The second reason is that to generate positive transformations in a context such as a classroom, one must understand how it develops naturally and what factors play a role in such development. This knowledge is an essential prerequisite to proposing solid research-based intervention strategies designed to change classrooms and, as a consequence, to create environments that appropriately foster children's adjustment. In spite of the importance of this social environment for child development, little is known about how classroom structure changes over time. Most research efforts have studied these contexts as predictors but not as outcomes; in consequence, change of these settings themselves has been overlooked. This raises questions surrounding which factors may generate changes in the social system of these contexts.

The present thesis will address these concerns through three separate but inter-related studies that have complementary goals. The first study, methodological in nature, proposes and tests the efficacy of a statistical procedure that adjusts peer-nomination-based scores obtained from sociometric and peer assessment procedures administered in

classrooms to avoid the potential for bias due to variations in class size. In this way, we seek to capitalize on the value of classroom-based assessment techniques, by assuring the reliability and validity of these measures. The second study is concerned with classroom effects on children's development. Specifically, we are interested in the way in which norms, cohesion and status, as structural aspects of the classroom social organization, can be measured, and in the predictive value of these context characteristics on the development of children's aggression. The third study follows from the results of the second study, by identifying and analyzing changes in those classroom norms that seem to have a substantial impact on students' development. Specifically, we look at how these norms change over the course of a school year and at the variations across classrooms in those changes that can be predicted by other factors (i.e., teachers' characteristics). This approach brings a new view to the study of social development, one in which it is understood that, groups, just as individuals, develop over time.

The classroom as a social context

Classroom characteristics have been assessed in a variety of ways and have been based on a diverse array of approaches. Although this diversity has facilitated the availability of a large set of useful classroom social features to be observed, it has also created uncertainty as to the most useful framework for studying the organization of the classroom as a social setting.

Allport (1924a,b) argued that groups can only be understood based on the characteristics of their individual members. Several investigations seem to follow this premise, as they have assessed classrooms based on aggregates of their individuals' characteristics. For example, many studies evaluate classroom climate by calculating the

average of students' aggression within the classroom (e.g., Barth, Dunlap, Dane, Lochman, & Wells, 2004; Thomas, Bierman, & The Conduct Problems Prevention Research Group, 2006), or of their normative beliefs (e.g., Henry, et al., 2000). This approach suggests that the classroom group is defined by its members' attributes and not by its own emergent characteristics. In contrast, other social perspectives propose that groups can and should be understood as entities per se and not just as collections of individual characteristics (Durkheim, 1895). In this regard, Moos stated that "environments, like people, have unique personalities" (1976, p. 320). The present study follows this second perspective, suggesting that classrooms, as groups, should be considered as a whole that is more than just the sum of its parts.

The holistic organization of the social systems that take place in classrooms can be understood according to the perspectives of three theories. One is the ecological view outlined by Kelly and colleagues (2000). According to this perspective, a social system consists of social structures and social processes. Structures are defined as the elements of the system that create the framework for the interaction among participants in the system. Structures "indicate what makes up the characteristic elements of the system" (p. 135). Processes, on the other hand, are the actions that permit the creation, maintenance or transformation of the structures of the system. Processes "indicate how interactions take place within a social system" (p.135).

A second framework is that of theories of small groups. These theories consider groups as complex, adaptive, and dynamic systems (McGrath, Arrow & Berdhal, 2000) and also emphasize the importance of analyzing group' structures and processes as the main dimensions of the interactions that occur within a group. From this perspective, the

group structure consists of the patterns of relationships among its members and includes aspects such as norms, roles, status, leadership style and cohesion (Cartwright & Zander, 1953; Levine & Moreland, 1990; Shaw, 1981). Group processes, on the other hand, pertain to the dynamics that emerge and change in the interactions among members of a group, and include such things as role differentiation, conformity, and social influence, among others (Levine & Moreland, 1990).

A third frame of reference that enriches our understanding of the classrooms as social systems comes from the systems framework proposed by Tseng and Seidman (2007). They state that social settings are dynamic systems that consist of social processes, resources and the organization of resources. Social processes refer to the transactions that occur among the interacting members of the setting; resources include any physical, economic, human or temporal elements available for individuals in a social setting; and the organization of the resources refers to the particular way that such elements are arranged or managed. According to this view, social processes are the main engine for youth development, and can be captured by studying the norms, relationships and participation of the members of the social system.

These frameworks serve as our points of departure to organize our theoretical approach to the study of the classroom social processes. Fundamental to our perspective is the view that the classroom is a social context with features that are defined by the interactions between classmates. Our approach is especially concerned with the social structure and processes that comprise the classroom environment. Consistent with the ideas of Tseng and Seidman (2007) we define social processes as the transactions and

interactions between classmates that manifest themselves in the structure of the classroom social organization.

The focus of this thesis is on processes of conformity within the classroom. For this reason, the structural aspects we are interested in examining are norms, cohesion, and status. Norms define the acceptable and expected behavior of the group members; cohesion is the extent to which the students are attracted to each other; and status is related to the importance, prestige and value that each individual has within the group (Shaw, 1981). These characteristics of the classroom are important for the study of conformity processes for several reasons. First, it is important to assess classroom norms because individuals will conform to what they believe are the behavioral expectations in their group. Second, high-status is granted by the group to those individuals that conform to group norms (Coie & Kupersmidt, 1983) and thus, these individuals become role models to whom others look up (Prinstein & Cillessen, 2003). Third, because membership to classrooms is not children's choice, the motivation to conform to the peer group norms may depend on the level of importance that students ascribe to their classroom peer group. To become a reference for behavior, the classroom has to be meaningful for the students (Richer, 1976). One way in which this can be evaluated is by looking at the level of connectedness of the students with their peer group, that is, their level of cohesion.

In sum, the present study seeks to examine properties of the classroom that transcend the characteristics of its members. Therefore, we will focus on the structure and processes that take place within these social systems. Given our interest in looking at conformity processes, we will center our attention on the structural characteristics of

classroom groups that may be relevant for this, namely, norms, cohesion, and status. To understand these social processes we will study the classroom context in three different ways: as a context of measurement, as a predictor of individual development, and as a dynamic system that changes over time.

Study 1: The classroom as a context of measurement

A basic premise of our approach is that peers, collectively and individually, possess a large amount of valid and reliable information about constructs related to the social processes of their classrooms. More so than teachers and parents, peers have extensive access to each other's behavior. On a daily basis and in a wide range of contexts they can see what their classmates do and hear what they say. By observing how their peers spend their time, what they like to do or to avoid, and how they react to experiences, children can form a powerful impression of what the peers in their classroom are like. Like their teachers, children are able to observe each other in the classroom; unlike teachers, children can also observe each other in other important school-related contexts such as in the lunchroom, the playground, and in the hallways of the school. We believe that this rich data base is an essential source of information about the characteristics of the classroom and its members.

Although observations from peers have tremendous potential for measuring psychological constructs, a basic component of the type of techniques used to collect this information might nevertheless hinder their validity. That is, peer nominations of sociometric status or of behavioral assessments have been widely used. However, the fact that classrooms vary in their size and therefore differ in the number of nominators challenges the reliability and/or validity of these measures. In spite of the development of

techniques to control for differences in classrooms size, limitations still exist that might bias results from research that uses these types of measures.

Based on this, the first study in the present thesis proposes a new mathematical procedure to maintain the value of the classroom as a context of measurement while controlling for the differences in classroom size. It is the first in a series of three studies related to classroom context and provides a basis for the measurement of the social structure of the classroom and of the students' behaviors used in two subsequent studies. We expected to demonstrate that this procedure would yield more reliable and valid results when studying classroom social processes.

Study 2: The classroom context as a predictor of individual development

The bulk of the literature on classroom contexts has regarded them as a source of variation in student's development. Indeed, research findings suggest that the social and learning environment of the classroom has an important impact on children's social and cognitive development (e.g., Fraser, 1994; Ghaith & Bouzeineddine, 2003; Hoglund & Leadbeater, 2004; Weinstein, 1991; Willms & Somers, 2001). Following our conceptualization of the classroom social organization, the focus of this particular study was on the effects that the classroom social norm and cohesion might have on children's aggression.

The contribution of this study to our understanding of classroom processes is three-fold. First, we looked at developmental changes in students' behavior over the life of the classroom, that is, the school year. In this way, we were able to better understand how children develop as a function of their context environment. Second, we propose new ways to measure classroom norms that are based on a systemic view of the classroom

functioning. Instead of looking at classroom aggregates of students' behaviors or cognitions, we assessed the acceptability of aggression by measuring the positive feedback of the social system following exhibition of this behavior. Third, we tested two opposite roles of classroom cohesion: one as a main predictor of lower levels of aggression, and the other as an enhancer of conformity to aggression norms. By looking at the latter role we acknowledged that the context might vary in the meaning it has for an individual and hence, we evaluated whether normative processes vary as a function of the ascribed significance of the context.

Based on this, we addressed three main research questions. First, we tested whether classroom aggression-related norms at the beginning of the school year would predict high initial or increasing levels of aggression across the school year. Second, we examined whether higher levels of classroom cohesion at the beginning of the year would prevent children from showing aggression. Third, we evaluated whether classroom cohesion would increase the level of conformity of students to aggression-related classroom norms. By answering these questions, we expected to have a better understanding of the effect of the classroom social structure on children's social development.

Study 3. The classroom context as dynamic system

One of the purposes of the present study was to better understand how the social structure of the classroom develops. Our theoretical point of departure was Arrow's systemic view (2004) that describes the forces that drive social groups towards change and stability. In particular, we focused on teachers as internal agents of classroom change.

Based on the results from Study 2, we chose the indicators of classroom aggression-related norms that had the most predictive value in understanding the development of students' aggression. We collected data on these classroom indicators across four waves during the school year, which allowed us to explore changes in the extent to which aggression is tolerated in classrooms. We did not have a previous hypothesis about the direction of these changes, but we expected to find variations across classrooms in the way they changed. Given the role of the teacher as the main leader of the classroom activities, we examined the influence of their cognitions and behaviors on the social environment of the classroom. Specifically, we tested whether teacher's normative beliefs regarding aggression and/or their supportive behaviors toward students explained increases or decreases in the classroom aggression-related norms over the course of the school year. Results from this study are discussed in terms of their usefulness for informing interventions aimed at promoting positive transformations in the social structure of these settings.

Summary

In sum, the goal of this thesis was to study the classroom context in a comprehensive way that includes the classroom as a context of measurement, as a predictor of individuals' development and as an outcome. Based on this, three sets of research questions were explored through three inter-related studies. The first study acknowledged the value of the peer-based measures collected in classrooms and proposed methodological solutions to control for the effect that variations in classroom size might have on statistical results conducted with this type of measures. The second study used these peer-based measures to capture the classroom social structure (i.e. norms, cohesion

and status) and individuals' behavior (i.e. aggression), and examined the classroom context as a predictor of individual development. Based on these results, the third study examined how the classroom features that predict individual change may themselves change over the course of the classroom life (i.e., a school year), and how teachers may have an impact on those transformations. As such, the overarching goal of this study was to further our understanding of the role and significance that the classroom context may play in child development.

Chapter 1. Adjusting for group size effects in peer nomination data.

Sociometric and peer assessment techniques continue to be standard measurement procedures in research on peer relations. The value of these widely used techniques derives from their capacity to provide efficient and direct indices of individual differences of functioning within the peer group. Each of these techniques takes advantage of the richness of the knowledge that exists in the peer network. Sociometric techniques use data from individual children concerning which of their peers they like and which they dislike to compute indices of acceptance, rejection, and friendship for each child within the group. Peer assessment techniques, such as the Revised Class Play (Masten, Morrison, & Pelligrini, 1985), in which each child indicates whether or not each of his/her peers exhibits particular characteristics (e.g., "someone who helps others"), capitalize on the extensive and diverse observations that each child has of her or his peers to create measures for each child on basic dimensions of social functioning, instrumental competence, and personal well-being. These techniques typically use peer nominations[1] and produce frequency measures indicating how often a child was chosen as liked or disliked and how often each child was seen as fitting a particular characteristic.

Given their ubiquitous use and critical value to the study of peer relations, it is important that these techniques are used in ways that will maximize validity and reliability and minimize bias. The quality of peer research that uses these procedures is, in part, dependent on the ability of these techniques to produce accurate scores whose meanings are invariant across contexts. In order for nomination-based measurement

---

[1] Techniques that use either peer ratings or paired comparisons are used much less frequently than peer nomination measures.

procedures to produce meaningful scores they need to create a common frame of reference across groups. One challenge to the creation of a common frame of reference derives from variability in the characteristics of the peer group contexts in which peer assessments and sociometric methods are used, typically classrooms in schools. These contexts often vary in size, sometimes by substantial amounts. This variation in size is potentially problematic as it may affect the frequency scores observed with peer assessment and sociometric procedures. The purpose of this paper is to present and evaluate solutions to this potential source of bias in the frequency measures derived from these procedures. The proposed solution is a regression-based technique that adjusts peer assessment and sociometric scores for variations in the size of the peer groups in which these procedures are used.

The introduction to this paper is divided into three sections. First, we discuss why and how group size variations are a potential problem. In the second section we discuss the techniques that have been used already to account for variations in group size, and limitations of each technique. In the third section we describe our proposed regression-based adjustment procedure.

Adverse effects of variability in group size for peer nomination methods.

Throughout the eight-decade history of the use of nomination-based peer assessments and sociometric techniques, it has been believed that the scores they produce may be positively correlated with the size of the nominating pool (Cillissen, 2009). That is, scores on nomination-based measures from larger peer groups are believed to be inflated relative to scores from smaller groups. This presumed inflation may be due to the larger number of opportunities to be chosen, that is, as the number of persons giving

nominations increases, the likelihood that a particular child would be chosen by her or his peers will increase also. Yet, in spite of the general belief that nomination-based scores will be positively associated with class size, there is also reason to expect that this association may be negative. It is conceivable that as the number of potential nominees increases, the likelihood of a particular child being chosen will decrease. Considering chance only, if a nominator can choose from a larger pool of peers, the likelihood that she or he will chose any particular peer will be smaller than when there are fewer peers in the pool.

If these two mechanisms (i.e., an increase due to a larger number of nominators and a decrease due to a larger number of nominees) had the same strength, the effects of one would compensate for the effects of the other. Presumably, the effects of these mechanisms are more likely to be equal when limited choice procedures[2] are used than when unlimited choice[3] procedures are adopted. Because unlimited choice procedures are currently used more widely than are limited choice procedures and also because limited choice procedures might be affected more strongly by the first mechanism under certain circumstances (described next), the presumed positive relation of scores to the size of the group needs to be considered.

A second consideration related to group size is just as important as the factors discussed above; however, it may affect the observed scores in a different way. Whereas the first set of considerations would affect the overall group mean of the observed scores,

---

[2] Such as when only one peer can be chosen for each item in a peer assessment technique and when one can chose only 1, 2, or 3 peers as friends or as disliked peers in a sociometric questionnaire

[3] When one can choose as many peers as one wants to for the roles in a peer assessment technique and as friends and disliked peers in a sociometric measure.

a second issue affects the variations in the range of the distributions across groups, in addition to group means. The central point of this issue is that the maximum possible value on a score might be higher in larger groups than in smaller groups. Consider this example. Suppose that someone is so unlikeable that nearly everybody in the group dislikes her or him. If there are ten people in her or his group, then as many as nine people may identify her or him as a disliked peer. His/her rejection score can be as high as nine but no higher. If he or she is in a group of 20, then the rejection score could be as high as 19, more than twice as high and in the group of ten. The same unlikeable child would have two very different scores merely as a function of the size of the group. Hence, this outlier may have different effects on the means and distributions of the two classrooms. .

In summary, it is reasonable to expect that variations in group size will add a source of variance to measures based on peer nominations. This variance constitutes a form of error in the sense that it is unrelated to the true score variance of the construct that the item is intended to measure. This error is problematic because it may detract from a measure's accuracy and validity. Moreover, group size can also affect the size of the correlations observed between measures based on peer nominations. In so far as a portion of the shared variance between two nomination-based measures comes from their mutual association with the size of the group, the correlations between these measures will include an artificial positive inflation. Accordingly, the observed correlation between two measures whose true scores are positively correlated will be even more positive (i.e., inflated), whereas the observed correlation between two measures whose true scores are negatively associated will be less negative (i.e., more positive, or attenuated). When

these scores are adjusted to account for the effects of group size, positive correlations will become less positive and negative correlations will become more negative. So even under circumstances where correcting for this source of error may not seem to matter (e.g., when only within-group comparisons are being used), these problems need to be eliminated. In the next section we discuss forms of correction that have been used in previous research.

Prior procedures for controlling for group size differences by adjusting individual scores

Already several procedures have been adopted in an effort to correct for the bias in observed scores that result from differences in group size. Use of these correction procedures has been motivated by two goals. One is to estimate the amount of variability in nomination-based scores that is due to variations in group size *per se*. This estimate gives researchers a chance to correct their scores so as to create scores that are unaffected by group size. A second goal has been to make individual scores comparable across groups, while maintaining their interpretability. These techniques are aimed at creating a common interpretive frame of reference. In this section we describe the techniques that have been used to accomplish these goals, pointing to their strengths and limitations.

Proportion scores. One procedure, whose main advantage is its apparent simplicity, is to create proportion scores by dividing each observed score by the number of nominators[4] in the group (see Chang, 2004). Here is how the proportion approach works: Suppose that two students from different groups, one with 11 members and the other

---

[4] In bears mentioning that when we refer to the number of nominators, we exclude the nominated person from the count. The reason for this is that, in most nomination procedures the nominee is asked to not nominate him or herself. In other words, the number of nominators is equal to the number of participants filling out the sociometric and peer assessment questionnaires minus one (the nominee).

with 21, each received 5 nominations. By dividing each child's score by the size of the group, the score of the first student would be .50 and the score of the second would be .25. These corrected scores are expressed as proportions of the theoretical maximum. That is, the score is expressed as a ratio of the observed value relative to the potential. This technique is attractive as it provides an easy way of accounting for group size and it provides scores that are easy to compute and use. These advantages, however, are seen as specious when the limitations of this technique are acknowledged. The first limitation has to do with the metric, which is a recurrent limitation across the different techniques available to correct for group size. Although changing the metric of the scores seems convenient, researchers may lose their capability to interpret their results in relation to the metric of the actual measurement system. In its original form, the unit of increase is one nomination. That is, a one unit increase in the score actually means that one more assessor has recognized that the target child possesses a particular behaviour or that one more person likes or dislikes the target child. With proportion scores, the interpretation of a one unit increase is removed from this original metric making its meaning somewhat unclear.

A second limitation is that this procedure assumes a strong one-to-one linear effect of the group size on individual scores. That is, if a person in a group of 10 nominators receives 5 nominations, a person in a 20% larger group (i.e., 12) will receive a 20% larger score (i.e., 6). As we will see later, although a linear effect is present, the relation is not one-to-one.

Finally, a third limitation, related to the previous one, is that this procedure overlooks the size of the pool of nominees. It is evident that as the number of nominators increase,

the size of the potential nominees also increases. As we outlined earlier, just based on probabilities, the chances of being chosen decreases as the potential choices increase. In addition to this, nominators might feel discouraged from choosing many peers when faced with long lists of names. Then, a *plateau effect* might take place where the positive linear effect of the size of nominators might progressively decrease as the number of nominees increases. Given that proportion scores do not account for this effect, comparisons of individuals who come from groups that differ significantly in size might be misleading.

Proportion of maximum score. A possible solution to the problem of using the theoretical maximum score as a divisor to create proportions would be to use the observed maximum score. Using the same example described above, if in the group of 10 nominators the maximum number of nominations received by any of the students in the group was 8, and in the group of 20 nominators the maximum was 12, then the student with 5 nominations in the first group would receive a corrected score of .62 whereas the student with 5 nominations in the second group would have a corrected score of .42. In parallel to proportion scores, this technique also allows for between-subjects and between-group comparisons. In addition, using the maximum score as the denominator in the proportion eliminates the assumption of a strong linear effect between group size and the size of the observed scores. However two other problems arise. One problem is the same as with the proportion score. Specifically, the computed proportions will again have a very different metric than the observed scores, making the interpretation difficult. A second problem concerns the bias that would occur when an outlier is present in a particular group. Consider the following example. Two highly likeable children are in

two different groups (Group A and Group B). Each group includes 13 children. Each of the two likeable children is nominated by 8 of their peers as a friend (i.e. their raw acceptance scores will be 8). In Group A, however, there is another child who is also very well liked. This well liked child is chosen as a friend by each the peers in this group (i.e., this child has an acceptance score of 12). If the child in group B has the maximum score in the group's distribution, his/her proportion score would be 1.0 (i.e., the observed score (8) divided by the maximum for his/her group (8)). The child in group A, however, that also includes the extremely likeable peer, would have a proportion score of .66 (i.e., 8/12). As we see, two children who are liked to exactly the same degree have very different scores. The problem with this technique is that the observed metric is too dependent on the presence (or absence) of an outlier in the distribution. This issue can cause particular problems with longitudinal data. Suppose that at Time 2 the extremely likeable child in Group A has left the study and our example child has the same score as at Time 1 but now has the highest score in the group. The example child's score will increase from .66 to 1.0 even though his/her acceptance is exactly the same. The apparent change was the result of the absence of the outlier. For these reasons proportion scores do not seem to provide a solution to the problem of bias that can result from differences in group size.

Standardization. The most long-standing solution to the potential problem of group size variations has been standardization. The conversion of raw scores into within-group standardized scores has been used for several decades as a means of eliminating size differences between peer groups. This procedure transforms individual scores into standardized values, taking as the frame of reference variations within each group. In this

case, the scores of each person are subtracted from the mean number of nominations received in a group and then divided by the standard deviation of the scores within the group. In our example, if the student with 5 nominations is in a classroom where, on average, students received 4 nominations and the standard deviation is 3, her or his standardized score would be .33 ((5 − 4) / 3). If the other student, who also receives 5 nominations, is in a class where the mean is 3 and the standard deviation is 4, his/her score would be .50 ((5-3)/4). Corrected scores are expressed as deviations from the group mean, which are mathematically comparable across groups. Similarly to the other techniques described above, this procedure offers a common metric, which in this case is a score relative to the group functioning. When it comes to sociometric assessment, where status is the object of measurement, relativity to the group is very reasonable. In fact, Coie and colleagues (Coie, Dodge & Coppoteli, 1982) based their well-known status classification system on this procedure. This evident advantage for the assessment of status and the fact that standardized scores are very common in statistical analyses is likely what makes this procedure the preferred one by peer relations researchers.

In spite of the wide spread use of standardized scores, this type of transformation can cause several problems, some of which have been pointed out already (Newcomb & Bukowski, 1983; Rubin et al., 2006). Although the use of standardized scores appears to promote easy interpretation, it can actually have the opposite effect. For instance, two children who are from groups of exactly the same size and who are chosen exactly the same number of times for a particular peer assessment item will have different standardized scores if the two groups have different means and variances. If both children are chosen 5 times for an item but one is from a group that has a mean of 4.5

and a standard deviation of 2 while the other is from a group that has an outlier and, as a result, has a mean of 5.5 and a standard deviation of 2.5, then the first child's Z score will be .12 while the other child's score will be -.2. They would appear to differ by 1/3 of a standard deviation even though their observed scores were identical. In consequence, although this strategy is used as a way to make comparisons between individuals, it may nevertheless result in creating even more interpretive challenges.

This procedure also distorts differences between groups. Standardization creates a common frame of reference by giving each group the same mean (i.e., 0) and standard deviation (i.e., 1). Initially researchers who used nomination procedures applauded the apparent consistency it created by eliminating observed differences between groups. It was a handy way of dealing with issues such as differences in group size that were difficult to deal with in other ways. The downside to this approach is that by creating similarity where it does not exist one eliminates important between-group differences (e.g., in mean and standard deviation). In this way, standardization can create distorted scores at both the individual and group levels.

Probability. Problems associated with standardization procedures were recognized several decades ago at a time when sociometric research typically relied on classification techniques (Bronfenbrenner, 1943, 1944). To avoid these problems, an alternative procedure based on probabilities was proposed. This method calculated the likelihood of particular scores given the group size and the number of choices made by each nominator. These values were used to create cut-off scores for assignment to sociometric groups (Bronfenbrenner, 1943, 1944; Newcomb & Bukowski, 1983). Although a probability method has some clear advantages for making assignments to

sociometric groups, its value as a means of correcting for biases that derive from group size differences in continuous measures is less clear. Accordingly, in spite of their intellectual elegance and their utility for making classifications in sociometric categories, approaches based on probability considerations are not likely to be useful for correcting dimensional scores based on peer nominations.

Residualized scores. Another approach that can be taken to eliminate the effect of the group size on raw nomination scores is to calculate unstandardized residual scores through a regression procedure. First, regression analysis is used to estimate the real effect that the number of nominators has on the participants' scores. Given the previously described plateau effect, it is reasonable to include both linear and curvilinear effects of the group size. In that case, it might be found that as the size of the group increases, the effect would progressively decrease. The next step would be calculating predicted scores based on the size of the group. Residual scores, that is, the difference between the real and the predicted scores, would then represent the score of each individual after the effect of the group size has been removed. This is a procedure that can be easily conducted with almost any statistical package. It is very advantageous in that it takes into consideration the actual effect of the group size, accounting for the number nominators and any plateau effect that might take place with the increase in the size of the nominees pool. Despite these strengths, two limitations need to be noted. On the one hand, the estimation of the residual scores would be based on an analysis conducted with the individual as the unit of analysis. Given that individuals within each group share the same number of nominators, the assumption of independence of observations would be violated and the coefficients might be distorted. The second limitation is related to the distribution of the

data. Given that residual scores have a mean of zero, between-group and repeated measures differences become impossible to discern.

Multilevel modeling. The limitations of the previous techniques call for a procedure that avoids transforming the metric of the original scores. One way in which peer-based variables can be examined is by estimating parameters of interest while controlling statistically for the effect of the group size. Given the nested nature of the data, group size should be included as a group-level covariate of any individual-level estimate, in a multilevel analysis. For example, in a study looking at the association between aggression and peer rejection (as measured with peer assessment techniques), such an association should be modeled at the individual level, while group size is included as a group-level predictor of both the individual-level intercept and slope. In this way, any effect due to the number of nominators would be removed for the estimated parameters. This technique is well-suited for studies with relatively simple models where the main focus of study lies on individual level associations. However, for complex models with many different types of predictors (e.g., moderators, mediators), or with group-level predictors derived from individual peer-based variables, controlling for the effect of group size may complicate model estimation. In multilevel structural equation modeling, for example, the inclusion of group size as a covariate adds unwanted residuals with other variables that might worsen the fit of the models, with covariances that lack any substantial meaning. Longitudinal analyses would be even more complex as group size at each time of measurement would need to be included and even more unwanted residuals might add noise to the estimation of model fit. As a consequence, for complex models, a

technique that eliminates the effect of group size from the individual scores prior to any statistical analysis is preferable.

A proposed regression-based procedure

The review of the procedures available to control for group size effects on peer nomination assessments leave us with the need to develop new techniques that overcome several challenges: (a) to maintain a metric that can be easily interpreted, (b) to account for the real effect that group size has on the number of nominations received, (c) to take into consideration the possible plateau effect that might take place when nominators are faced with long lists of potential nominees, (d) to allow for the possibility of estimating differences between individuals, groups, and across time of measurement, and (e) to keep statistical models as free as possible of any noise that might complicate their estimation. In the present study, we propose a technique that may overcome the limitations of the current procedures and that is very accessible to researchers as it is based on basic regression procedures.

Similarly to the residual scores and multilevel procedures, this technique makes use of regression analysis to estimate empirically the effect of group size on peer nomination scores. This procedure is not intended to derive a global effect that would be applicable to any study. Rather, it should be conducted for the sample of each particular study. In addition, our interest is not to generate any inferences about population estimates, but to eliminate the group size effect from the scores of the study in question. In consequence, this procedure does not deal with hypothesis testing. Even when very small effects are found, it is worth removing those effects from the variation in the nomination scores.

To estimate the effect of the group size on the nomination-based scores, the first step is to determine whether the dependent variables in the regression analyses will be the items, the variables or the questionnaires. For illustration purposes, we will focus our attention on two types of questionnaires. One is a sociometric questionnaire in which students rate, from one to five, the extent to which they like each one of their participating peers. Acceptance scores are derived by adding the number of times a person is chosen as a five, and rejection from the number of times he or she is chosen as a one. Given that only extreme ratings are considered, these scores can be compared to "most like" and "least like" nominations (Bukowski, Sippola, Hoza, and Newcomb, 2000). The other is a peer assessment questionnaire, where students nominate all of those peers who exhibit certain characteristics. In this case scores are calculated from the number of nominations received for each of the items. Scores within each one of these two questionnaires are comparable. For example, one might want to compare acceptance with rejection levels, or the prevalence of different types of aggression. Consequently, any technique used to adjust scores for group size differences should maintain the comparability of the items within the questionnaires.

Additionally, it is important to recognize that there may be substantive reasons to expect differences in the means of scores taken from groups that vary in size. For example, it is conceivable that larger groups, relative to smaller groups, may have a less cohesive atmosphere and that this lower cohesion might lead to either higher levels of aggression or lower levels of helpfulness. Accordingly, any correction should be the least content specific possible. Based on this, we propose that a mean number of nominations

received across all items within the same questionnaire (e.g., the peer nomination questionnaire) should be calculated.

The second step is to conduct a multiple regression analysis with the average of received nominations in the questionnaire as the dependent variable, and the group size (i.e., total number of nominators minus one) as the predictor. To account for any plateau effect, both linear and quadratic effects need to be estimated. To do this, group size should be centered at the mean and its quadratic value calculated, so this can be included a second predictor. The main interest of these analyses is estimating the unstandardized B coefficients, which will be used to correct the scores. The regression formula would be:

$$y = A + B_1 \, (groupsize) + B_2 \, (groupsize^2) + e$$

where y is the mean number of nominations received in a questionnaire, A is the intercept, B_1 is the linear effect of group size, B_2 is the curvilinear effect and e is the residual error. Given the nested nature of the data, one of two procedures should be conducted to estimate the B coefficients of interest. One is to use multilevel modelling and add the linear and curvilinear effect of group size as group-level predictors of the individual-level variations in the intercept of the questionnaire's average nominations. The other procedure would be to aggregate the questionnaire's average nominations at the group level and conduct multiple regression analyses with the group as the unit of analysis. Both procedures would yield almost identical coefficients.

Once linear and curvilinear coefficients have been estimated, one will know, on average, how many units of increase in the nomination scores of individuals will result from a unit increase in the group size. The next step would be to choose a reference group size to equalize the scores from individuals in the rest of the groups to that

particular size. This could be derived either from the sample group size mean or from a group size that the researcher finds meaningful for her or his interpretation. In classroom research, for example, one could think of 20 as a standard size that could be used as a reference. Creating a variable representing a deviation in group size from that reference size would help with the transformation of the scores. This way, for each unit increase in group size, scores would be decreased based on the coefficients. Similarly, for each unit decrease in group size, scores would be increased. The correction formula would look like this:

$$y' = y - [(\textit{size difference} * B_1) + (\textit{size difference}^2 * B_2)]$$

where y' is the adjusted score, y is the raw score, *size difference* is the deviation from the reference group size, B_1 is the linear coefficient, and B_2 is the curvilinear coefficient.

An important caution is that scores representing zero nominations should not be transformed. First, transforming zeros into negative scores would again make scores difficult to interpret. Second, in groups smaller than the reference group this procedure would force scores of zero to increase, which would be unrealistic for most of the individuals who were not nominated in any particular item. For this reason, it is important to keep zeros as untransformed.

This technique presents several advantages over the ones that have been traditionally used in nomination-based assessments. First, it maintains the original metric (i.e. number of nominations received) of the scores so that any statistical estimates as well as effect sizes can be easily interpreted. Second, it accounts for the actual effect of the group size, without any assumptions that the relation between the scores and number of nominations is strictly linear. Third, behavioural assessments become comparable across groups

because the estimation of the scores is not relative to the group (as it is the case in the standardization procedure). Fourth, this technique takes into consideration that the effect of the group size might be smaller for larger groups (long list of nominees decrease the probabilities of being chosen) by estimating curvilinear effects. Fifth, within-group means and variances can be estimated which allows for between-group comparisons. Sixth, any variations across times of measurement due to changes in the group size are controlled, so the adjusted scores are a better representation of true scores of individuals in longitudinal studies. Finally, regression analysis are widely known and used, which makes this procedure very accessible to any researcher.

Summary and goals

Variations in group size are a potential source of bias in measures derived from nomination procedures. Scores for individuals in large groups may be inflated or attenuated simply due to a higher number of nominators. We have described different procedures for eliminating this source of bias and have as outlined their potential limitations. Based on this, we propose a regression-based technique that we believe will overcome many of the limitations of previously used techniques. In the following sections we present an empirical evaluation of this technique. The specific goals of this evaluation are: (a) to illustrate with real data the regression-based procedure to correct nomination-based scores for differences in group size, (b) to compare the efficacy of standardization and regression-based corrections in estimating the stability of the variables over time, between-group comparisons and mean changes in the variables over time, and (c) to analyze the extent to which associations among nomination-based raw

scores are inflated due to a confound between true scores and shared variance related to the common group size.

Method

Participants

The sample for this evaluation study consisted of 1595 boys and girls (mean age = 10.2 years; 53% male; 47% female) from fourth, fifth and sixth grade in nine schools in Bogotá, Colombia. Participants were enrolled in mixed-sex schools drawn from neighborhoods that varied in socio-economic status (SES). SES was assessed via parental reports about their household stratification, which is defined in Colombia by the local government and ranges from 1 (low) to 6 (high). School SES was then specified according to stratification level of the majority of the students. A total of 63 classrooms took part in the study and their size ranged from 11 to 32 participants (mean size = 21.98).

Procedure

The purposes and procedures of the study were described to the principals or coordinators at the participating schools. After administrative approval was received, the students in the schools were informed of the objectives and procedures of the study in their classrooms, and were given letters of information as well as parental consent forms to be brought to their parents. Only the participants whose parents returned a signed consent form were included in the study. Using this recruitment procedure, the overall participation rate of the study per classroom was, on average, 79%. At the end of the study, participants received school supplies as an expression of our gratitude to them for being in the study.

Using a classroom administration format, participants completed a multi-section questionnaire during a one-hour in-class session. Data were collected at four time points, at roughly eleven-week intervals during a single school year

Given drop-out, inclusion of new students or incomplete responses, a range of 2% to 15% of the data was missing at some points during the study. Results from Little's MCAR test indicated that data was not missing completely at random ($\chi^2_{(2256)}$= 3280.30, p < .05). Under the assumption that the data was missing at random, multiple imputation of 20 data sets was conducted using the software known as Amelia II (Honaker, King, & Blackwell, 2009). Although only sociometric and peer assessment data were used in the present analyses, all the variables assessed in the broader study were included in the imputation model. To account for the hierarchical nature of the data set, classroom identification was "dummy coded" and included in the imputation procedure. Based on Rubin's (1987) rules for estimating within- and between-imputation variance, the fraction missing for our variables of interest ranged from 9 to 10%. This rate of missing information, combined with the 20 imputations conducted, represented a level of efficiency of 99% (McKnight, McKnight, Sidani, & Figueredo, 2007) in the estimation of the variable means.

Instruments

Sociometric assessment. A liking questionnaire was administered to assess how much each student liked his/her participating peers. Each student was given a list of their participating classmates and was asked to rate the extent to which he/she liked each one. These ratings were made with a five point scale in which 1 was labeled "do not like this person" and 5 was labeled "like this person very much". These rating were used to create

measures of acceptance and rejection for each participant. The number of times a child received a rating of 1 was used as the unadjusted measure of rejection; the number of times the child received a rating of 5 was used as the unadjusted measure of acceptance. For the purposes of the present study, each child's acceptance and rejection scores were added together and then averaged (i.e. divided by 2) at each of the four measurement times. This score represents the average number of nominations that a child received in the sociometric questionnaire. The number of participants within each classroom minus 1 (to account for the fact that a child did not rate him/herself) was used as the index of group size for these sociometric measures. Group sizes ranged from 9 to 31 at Time 1, from 11 to 29 at Time 2, from 11 to 30 at Time 3, and from 9 to 30 at Time 4. The group size of 20 was used as the median value.

Peer assessment. A peer assessment procedure (for a description, see Rubin, et al., 2006) was used to assess four broad dimensions of behaviour. Six items were used to measure aggression (e.g., someone who hits or pushes other people), two items for helping (e.g., someone who helps others when they need it), three items for sociability (e.g., someone who gets along well with others), and one item for perceived popularity (e.g., someone who is popular). For each item students were asked to nominate any number of their participating classroom peers who they thought matched the item. For each item, a raw score was calculated based on the number of nominations received. To calculate the mean number of nomination received in the peer assessment questionnaire, the 12 item scores were averaged for each student at each time point. Similarly to the sociometric assessment, group size was calculated as the number of students who gave nominations minus the nominator. Under rare circumstances, such as when a child failed

to fill out the sociometric questionnaire, group size could vary slightly from one questionnaire to the other. Group sizes for the peer assessment ranged from 10 to 31 at Time 1, from 11 to 30 at Time 2, from 11 to 32 at Time 3, and from 10 to 30 at Time 4. Again, a group size of 20 was used as a median value.

<div align="center">*Results*</div>

Illustration of the regression-based correction procedure

The regression-based procedure designed to control for group size differences consisted of three steps. First, students' scores (i.e. average number of nominations received in each questionnaire) were aggregated at the classroom level, to account for the embedded nature of the data. This aggregated value was a direct index of the total number of times that the children in a particular group had been nominated in each type of questionnaire (i.e., sociometric or peer assessment). Second, a series of multiple regressions was conducted with group as the unit of analysis. Average nominations for each questionnaire (i.e., sociometric and peer assessment) were the dependent variables and the linear and quadratic effect of classroom size (i.e., number of raters/nominators minus one) were the predictor variables. To facilitate the interpretation of the intercept and to reduce collinearity between the predictors, classroom size was centered at the median value of 20 and then its quadratic effect was calculated. This calculation was done separately for each of the four times of measurement.

The results of these analyses are shown in Table 1.1. This table shows the intercepts, and unstandardized *B* coefficients found for each scale at each time point for the two types of questionnaires. Given that we were not interested in making inferences about the population as a whole, no significance testing results are reported.

As shown in Table 1.1, results from the regression analyses showed that group size explains an important proportion of variance of the nomination-based scores, particularly for sociometric variables (e.g., effect size on the sociometric measures were .591, .650, .530 and .604 at T1 through T4 respectively). Also, we found that the linear and curvilinear effects tended to be very similar across time for both types of measurement. As predicted, the negative curvilinear coefficients indicate that the linear effect of the number of nominators on the nominations received was weaker for larger classes than for smaller classes. This pattern of results provides evidence for a plateau effect where, in face of large lists of peers, students tend to nominate a smaller proportion of available peers, compared to nominators in groups were lists of potential nominees are shorter.

The third step was to correct the scores according to the estimated coefficients. This procedure was carried out with the students' raw scores for each of the items on the sociometric and peer assessment questionnaires. It is important to note that corrections were performed only for scores that differed from zero. In other words, students who received zero nominations in any of the items were left with the same score, regardless of their classroom size. Centered classroom size, which represented the deviation from a reference classroom size of 20 students, was used as the basis for the scores' adjustments. The following formula was used to conduct the corrections at each time of measurement

$$y' = y - ((ccs * B_1) + (ccs^2 * B_2))$$

where y' is the adjusted score, y is the raw score, ccs is the centered classroom size (i.e., the classroom's deviation from the median score of 20), B_1 is the linear coefficient of classroom size and B_2 is the quadratic coefficient of classroom size. Based on this, students' scores in classrooms with a size larger than 20 were reduced, while students'

scores in classrooms smaller than 20 were increased. Following the application of this formula the adjusted scores were an estimate of the number of nominations a child would have had if he/she had been in a classroom with 20 nominators. As an example, a student with a raw acceptance score of 6 at Time 1 in a classroom with 22 nominators would have had her or his score adjusted to 5.724 based on this calculation: $6 - ((2 * .148) + (2^2 * -.005))$.

Comparison of regression-based procedure with standardization

To examine the performance of the scores adjusted for classroom size with the regression-based relative to the standardization technique, a series of comparative analyses were conducted. First, we standardized the scores for each item within each classroom. Next, we calculated peer-assessed variables by averaging the corrected items of the aggression, helping and sociability subscales. Acceptance, rejection and perceived popularity scores were measured with only one item, so no additional calculations were conducted for these variables. Then, we conducted three types of analyses to compare parameter estimations for the two procedures. First, we tested the stability of the variables across the four times of measurement. Second, we compared the variables at Time 4, across groups defined by socio-economic status. Third, we examined changes in peer-assessed aggression and rejection over time for the whole sample, as well as separately for girls and boys.

Stability over time. Measures of stability are reported in Table 1.2. The vast majority of the variables show high levels of stability from one time of measurement to the next, particularly for peer-assessed behavioural variables. More importantly, we found evidence that the regression-corrected scores were more stable than the standardized

scores, especially for sociometric variables. Specifically, for 16 of the 18 comparisons, regression-corrected scores were more stable than standardized scores. Only for two comparisons (i.e. aggression at time 2 to 3 and time 3 to 4), were both procedures equally stable.

Between-group comparisons. A second set of analyses was conducted to examine differences between high and low SES children in the variables evaluated at Time 4. Students from only high SES (i.e., stratification 5 or 6) and low SES (i.e., stratification 1 or 2) schools were selected for these particular analyses. A series of one-way ANOVAs were run with each of the sociometric and peer assessed measures as dependent variables, and SES as an independent dichotomous variable. No between-group differences were found with the standardization corrected scores, given that means of the groups are forced to be zero. In contrast, regression-based corrections allowed this variability to be captured, and revealed a number of substantive differences between low and high SES schools. For example, mean differences presented in Table 1.3 show that high SES children exhibit higher levels of acceptance, perceived popularity, helping and sociability. Low SES children, on the other hand, show higher levels of rejection and aggression.

Changes over time. Finally, we examined changes over time in the scores corrected by each technique. For illustrative purposes we ran these analyses with rejection and with a subset of items measuring physical aggression. We examined overall changes in the whole sample and also examined whether those changes differed between boys and girls. Repeated measures ANOVAs were run to achieve these goals. Results of the rejection scores corrected with the regression technique showed a significant main effect of time

(Wilks lambda = .981, $F_{(3,1590)}$ = 10.202, p < .05) and a non-significant effect of the interaction between time and sex (Wilks lambda = .998, $F_{(3,1590)}$ = .985, p > .05). On the other hand, both effects were non-significant when the analyses were run with standardized rejection scores (Wilks lambda = 1.00, $F_{(3,1590)}$ = .003, p > .05, for the main effect; Wilks lambda = .998, $F_{(3,1590)}$ =.883, p > .05, for the interaction). This pattern was expected given that means are pushed towards zero with the standardization procedure. Figure 1.1 and 1.2 depict the mean changes over time in rejection scores corrected with both procedures. These figures show how, while regression corrected scores represent changes in the behaviours of individuals over time, the standardized scores represent changes in boys' and girls' deviation from their classrooms' mean. Therefore, analyses conducted with each technique should be interpreted in completely different ways.

The same analyses were conducted for changes in physical aggression. Results from the repeated measures ANOVAs revealed significant changes over time in regression corrected aggression scores for the whole sample (Wilks lambda = .980, $F_{(3,1590)}$ = 10.797, p < .05), as well as a significant time by sex interaction (Wilks lambda = .990, $F_{(3,1590)}$ = 5.368, p < .05). In contrast, none of these effects were significant when standardized scores were used (Wilks lambda = 1.00, $F_{(3,1590)}$ = .003, p > .05, for the main effect; Wilks lambda = .998, $F_{(3,1590)}$ =.879, p > .05, for the interaction). An examination of the mean scores at each time point calculated after both techniques were conducted (see Figures 1.3 and 1.4) indicates that while the standardization suggest a similar pattern of no change for boys and girls, regression corrected mean scores show that, while girls do not seem to change over time, means scores for boys fluctuate from one time to the other. Simple-effects pairwise comparisons revealed that, while no

significant changes occurred for girls, boys showed a significant increase in aggression from Time 1 to Time 4, as well as from Time 3 to Time 4, and a significant decrease from Time 2 to Time 3. These findings confirm that, with the standardization procedure, changes over time are obscured.

Inflation of associations between nomination-based raw scores

Given that part of the variance in the nomination-based variables is due to the group size, we expected any association among these variables to be misleadingly more positive. To test this, we estimated the level of association among all of the variables at time 1, for the raw scores and for the corrected scores.

Table 1.4 presents the correlations among all the sociometric and peer-assessed variables at time 1, before and after being adjusted with both procedures. As can be seen, all correlation coefficients estimated with raw scores are more positive that those estimated with the regression-based corrected scores. For example, the negative association between helping and aggression, is more positive when it is estimated with raw scores ($r = -.371$), compared to regression-based corrected scores ($r = -.425$). This confirms our expectation that the use of raw scores might result in more positive associations (i.e., attenuation of negative associations and inflation of positive associations), due to the positive correlation between the group size variance in both nomination-based variables.

. In addition, these results show that the association between the scores corrected with either the regression or the standardization techniques are high (all coefficients are above .80). Also, correlations among the variables do not seem to differ considerably from one procedure to the other. However, it is worth noting that most of the associations

are slightly stronger when the scores are adjusted with the regression-based procedure, compared to the standardization approach.

Discussion

The purpose of this study was to propose and evaluate a solution to the potential bias in nomination-based sociometric and peer assessment measures that may derive from variations in the number of the nominators and nominees across peer groups. In spite of the well-known advantages of peer-based measures, researchers have faced the long-standing challenge of controlling for group size differences. Previously proposed solutions to this problem are unlikely to be effective as they present new practical or interpretive problems. Accordingly, we introduced a new regression-based correction technique. We proposed that this procedure has important advantages over some of the most commonly used techniques such as standardization. We described the procedure and compared its performance with that of the standardization technique in estimating different parameters.

The proposed regression-based procedure allows researchers to empirically estimate the effect of group size on nomination-based scores obtained with sociometric and peer assessments. Under the assumption that a higher number of nominators produces higher scores, the procedure estimates an index of the linear effect of group size on the scores. A second assumption of this model is that the effect of group size is nonlinear, specifically, that at some point there is a plateau in the effect of group size. That is, we expected that at some point along the distribution of class size the linear effect would decrease or even level off. We account for this plateau effect in our procedure by including estimates of quadratic effects. The estimation of linear and quadratic effects is conducted with the

group (e.g., classroom) as the unit of analysis, by means of multiple regression analyses. The linear and quadratic coefficients are then used to correct the scores by adjusting for the expected inflation or deflation in scores that would result from variations in size from a reference group. Scores are then corrected so as to reflect the score any individual would have received, had she or he been in a classroom that was the size of the reference.

Results from the multiple regression analyses revealed that group size accounted for an important proportion of variance of the nomination-based scores. The effect size, as estimated by the R^2 values, ranged from 20 to 65%. Moreover, results that contrast the correlations among nomination-based variables, both for raw scores and for corrected scores, provide evidence that these associations are slightly inflated when the effect of group size is not eliminated. As noted by previous researchers (for a review see Cillessen, 2009), this pattern confirms the need to control for such effects to avoid bias in any estimation based on this type of measure. The linear and curvilinear coefficients found in these analyses provide strong evidence of the limitations described for the proportion scores technique. First, if the relation between group size and nominations received was strictly linear and if this effect were strong and invariant across groups, a linear coefficient close to 1.0 would be expected. As we see, the observed coefficients fluctuated between .105 and .188. Second, although the observed quadratic effects were small, they nevertheless indicate that as the size of the group becomes larger, the linear effect becomes weaker. This plateau effect, which becomes increasingly important when the range in group size is large, is not accounted for when proportion scores are used.

The results from the comparative analyses between the scores derived from the proposed regression-based procedure and the widely used standardization technique are

striking. They show clearly that between-group differences become almost impossible to detect, due to the artificial transformation of the means to zeros. This difficulty was evident in the results examining SES differences between schools. Even when group comparisons are discernable, as was the case for sex differences in rejection and physical aggression, interpretation of the findings is very unclear, given that the scores are expressed as deviations from the group mean. Thus, in order to interpret these differences in a meaningful way, one would have to assume that all groups function similarly and that deviations from the group mean have the same meaning across groups. Researchers have typically avoided this limitation by focusing their research questions on the associations among variables, where classroom averages do not seem to matter. However, we saw that even the associations among standardized variables can be slightly distorted compared to the regression-based correction, both among concurrent and longitudinal variables. Moreover, given that the variance of the variables subjected to standardization procedures is drastically transformed (to become 1.0), it must be questioned whether longitudinal studies looking at changes overtime through lagged correlations can really capture developmental processes in the individuals. In sum, the problem with the standardization technique is that it serves its function too well. By equating parameters across groups it eliminates group differences in central tendency and in variability that derive from any group-level processes, not just from variations due to differences in the number of nominators. Thus, we would argue that this overly thorough correction may create more interpretive problems than it resolves.

Limitations

A potential limitation of the regression-based technique described in this study may rest in its own goal of eliminating differences in observed scores that are associated with group size. Eliminating the effect of the group size is predicated on the idea that such differences are a methodological artifact. In some circumstances, however, it may be reasonable to expect that a group size difference may be substantive rather than artifactual. For instance, as outlined earlier, larger groups may lack cohesion, which in turn could promote less caring behaviours among their members. As another example, where school performance is measured via peer assessments, such performance may be enhanced in smaller groups, given the greater support that teachers can provide to their students. In these cases, entirely erasing differences related to group size would eliminate important true score variance. Our proposed procedure minimizes this problem by basing the estimate of the group size effect on a very broad index in which all available measures are taken into consideration. That is, the possibility of eliminating true score variance is reduced by using the average number of nominations received on all the different items included in a particular questionnaire as the dependent measure in these computations. By making the dependent variable as content-general as possible, any substantive variance in specific variables related to the size of the group would be retained. However, this could represent a problem in situations where only the dependent variables of interest are evaluated or, even more so, in cases where only one measure is used to assess a particular construct. The latter is the case when peer acceptance is measured by friendship nomination procedures, where the effect of group size would be completely eliminated. In such cases, the recommended action would be to assess the characteristics of the context that might be mechanisms through which group size has an

effect on the outcomes. In our examples, this would mean having a measure of group cohesion or a measure of teacher support. In this way, the richness of nomination-based procedures could still be taken advantage of, even in cases where group size is regarded as a substantive predictor.

Future directions

In our view, the main advantages of our regression-based procedure to correct for group size differences in nomination-based measures are practicality, efficiency, and interpretability. Since the procedure is based on a widely used statistical technique (i.e., multiple regression), any researcher should be able to implement the procedure in her or his sociometric and peer-assessment data. In this sense, it shares the same advantage that made the standardization technique so popular. As such, we would expect potential applications of this technique across several types of studies, which will contribute in the construction of peer relations knowledge in two ways. First, it will provide more information regarding the overall effect of group size. It would be interesting to examine the extent to which the linear and curvilinear effects are stable across studies, and evaluate if it is possible to establish a common ground for the correction of nomination-based measures.

Second, we believe that the use of this regression-based procedure provides a stronger basis to conduct meta-analyses, compared to the findings already available with standardization corrections. Z-scores are expressed as deviations from the group mean, which makes them relative to the group functioning. As a consequence, results based on these scores would only be comparable under the assumption that all groups within and between studies function in very similar ways. It is well known that social contexts differ

in their organization, and therefore this assumption is highly questionable. With regression-based scores, the interpretation of scores across groups and samples is the same. As a result, meta-analysis would be more feasible and valid with studies using this type of correction.

In sum, it is our contention that correction of group size differences by means of a regression-based procedure is a very powerful and efficient tool to improve research in peer relationships and in other areas that can make use of peer evaluations. It has several advantages as it maintains the original metric of the nomination-based variables, accounts directly for the real effect of group size, considers linear and curvilinear effects, and especially, it facilitates the accurate estimation of individual, group and time of measurement effects. Hence, the regression-based technique represents a very promising tool for enhancing the validity of peer relationships research.

Table 1.1

Regression unstandardized B coefficients for the effect of classroom size on average number of nominations received in each questionnaire

Time	Intercept	*B1* Linear effect	*B2* Quadratic effect	R^2
Sociometric assessment				
Time 1	3.596	.148	-.005	.591
Time 2	3.647	.188	-.001	.650
Time 3	3.537	.175	-.001	.530
Time 4	3.732	.187	-.001	.604
Peer assessment				
Time 1	4.985	.150	-.006	.392
Time 2	4.838	.108	-.002	.237
Time 3	4.701	.116	-.004	.203
Time 4	5.188	.145	-.003	.281

Table 1.2

Stability of the variables over time, compared for the two procedures

Variable	Time 1 and 2	Time 2 and 3	Time 3 and 4
Acceptance			
RB[1]	.746	.717	.720
SD[2]	.682	.663	.676
Rejection			
RB	.724	.694	.711
SD	.640	.615	.615
Perceived popularity			
RB	.840	.826	.830
SD	.795	.805	.802
Aggression			
RB	.838	.866	.866
SD	.849	.866	.866
Helping			
RB	.839	.838	.830
SD	.792	.788	.777
Sociability			
RB	.813	.839	.813
SD	.792	.815	.775

[1] Correlation coefficients for the scores corrected with the regression-based technique.
[2] Correlation coefficients for the scores corrected with the standardization technique.

Table 1.3

SES differences in the nomination-based variables, as captured by each correction technique

Variable	Low SES M	High SES M	p
Acceptance			
RB[1]	3.726	5.787	0.000
SD[2]	0.000	0.000	1.000
Rejection			
RB	3.992	1.147	0.000
SD	0.000	0.000	1.000
Perceived popularity			
RB	4.938	6.121	0.000
SD	0.000	0.000	1.000
Aggression			
RB	4.564	2.774	0.000
SD	0.000	0.000	1.000
Helping			
RB	4.737	8.338	0.000
SD	0.000	0.000	1.000
Sociability			
RB	5.953	9.602	0.000
SD	0.000	0.000	1.000

Note. Low SES $n = 422$, High SES $n = 772$.
[1] Scores corrected with the regression-based technique.
[2] Scores corrected with the standardization technique.

Table 1.4

Correlations among variables at time 1 corrected with the standardization and regression-based techniques

Variable	1	2	3	4	5	6
Acceptance	**.893**					
RW[1]	-					
RB[2]	-					
SD[3]	-					
Rejection		**.809**				
RW	-.403	-				
RB	-.480	-				
SD	-.460	-				
Perceived popularity			**.894**			
RW	.621	-.269	-			
RB	.600	-.312	-			
SD	.579	-275	-			
Aggression				**.915**		
RW	-.209	.408	.084	-		
RB	-.271	.373	.059	-		
SD	-.241	.302	.069	-		
Helping					**.874**	
RW	.644	-.410	.545	-.371	-	
RB	.619	-.477	.522	-.425	-	
SD	.578	-.375	.458	-.434	-	
Sociability						**.793**
RW	.703	-.486	.692	-.214	.827	-
RB	.667	-.588	.677	-.277	.816	-
SD	.662	-.471	.641	-.283	.711	-

Note: Bold coefficients represent the correlation between the scores corrected with each technique (regression-based and standardization) for the same variable.
[1] Correlation coefficients for the raw scores.
[2] Correlation coefficients for the scores corrected with the regression-based technique.
[3] Correlation coefficients for the scores corrected with the standardization technique.
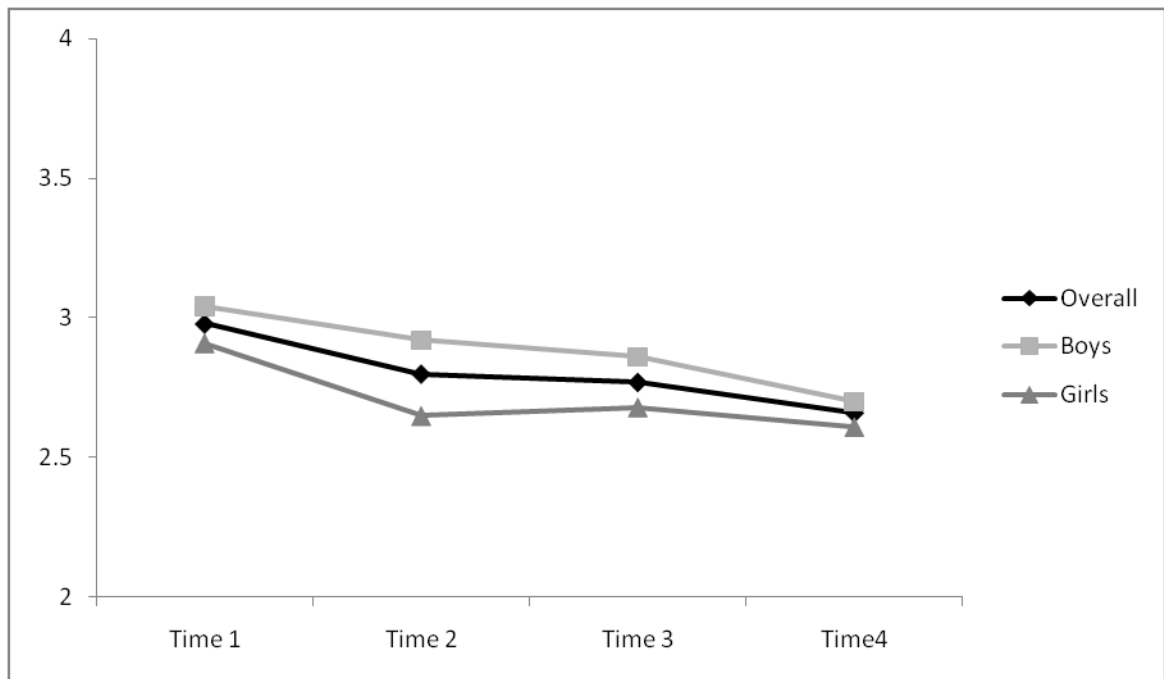
Figure 1.1. Changes over time in rejection for boys and girls, based on regression-corrected scores
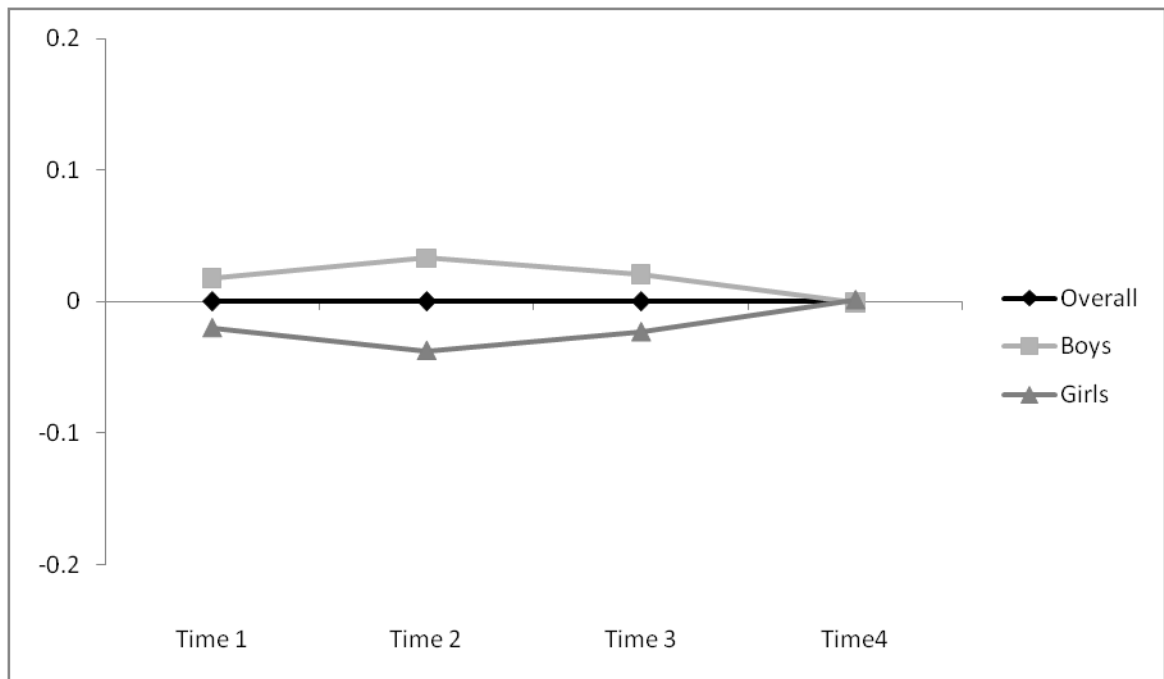
Figure 1.2. Changes over time in rejection for boys and girls, based on standardized-corrected scores
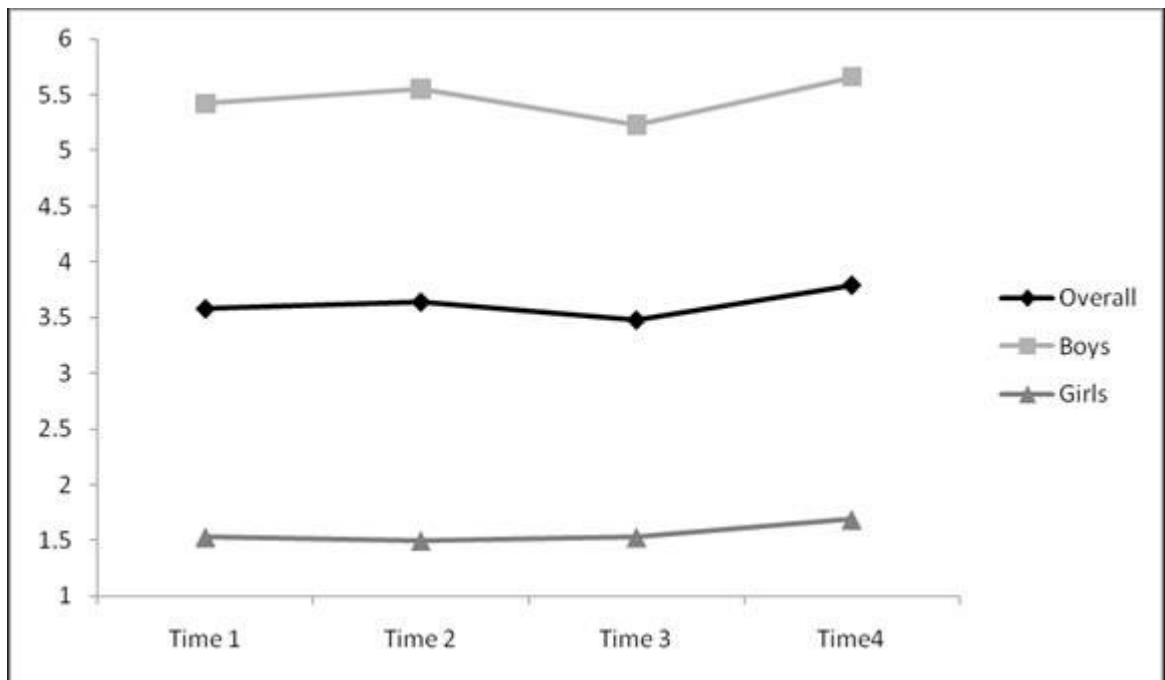
Figure 1.3. Changes over time in physical aggression for boys and girls, based on regression-corrected scores
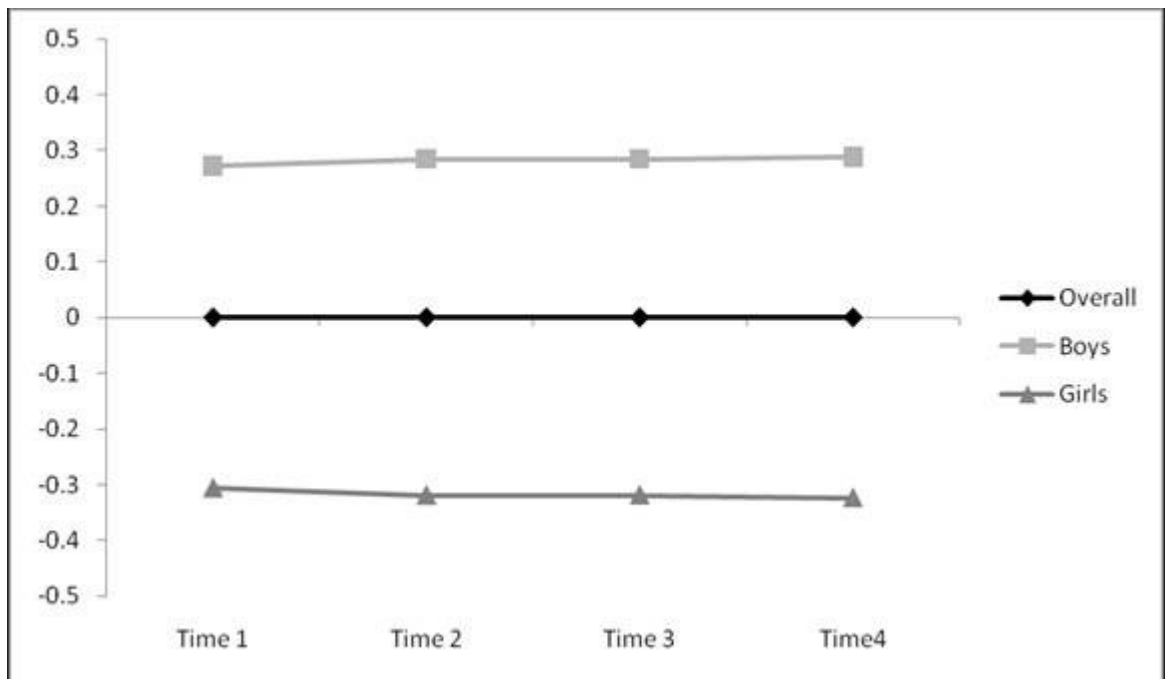
Figure 1.4. Changes over time in physical aggression for boys and girls, based on standardized-corrected scores

Chapter 2. The effects of classroom peer norms and cohesion on aggression

trajectories: A short-term longitudinal study.

Peers relations have been regarded as a powerful context of influence for multiple domains of children's social development such as antisocial behaviour. There is already a wealth of evidence showing that associating with deviant peers may increase the likelihood that a child will exhibit externalizing problems (Vitaro, Boivin & Bukowski, 2009). In spite of the size of the literature on this topic, many questions about how characteristics of the peer group itself (e.g., classroom context), rather than specific experiences within it, influences children's behaviours remain unanswered. Already important progress has been made regarding the characterization of peer groups in terms of their structure. For instance, there is some agreement that the properties of groups that transcend aggregations of individuals' characteristics include norms, cohesion and hierarchical organization (Rubin, Bukowski & Parker, 2006). However, more attention needs to be paid to the role of the peer group in children's social experiences (Bukowski & Sippola, 2001). The particular concern of the present study is with the effect of classroom norms on aggressive behaviour. The specific strategy of this study is to assess aggression-related classroom norms at the beginning of the school year and examine their influence on changes in children's aggression over the course of the school year. Beyond recognizing the effects of classroom norms per se, there is reason to believe that children who are especially attracted to the group may be especially likely to be influenced by peer norms. Therefore, the second purpose of this study is to test the moderating effect of classroom cohesion on the association between norms and aggression development.

It is broadly acknowledged that in order to understand the individual, it is necessary to also understand the ecologies in which he/she is embedded. Lewin (1936) claimed many years ago that the behaviour of the individual needs to be understood as a function of his/her personal characteristics, of the characteristics of the environment and of the interaction between the two. Similarly, ecological theories (Bronfenbrenner, 1976, 1992, 2006) stress the importance of the interaction between the individual and the shared and embedded social systems in which he/she develops. Based on this, one important context that seems to deserves special attention is the classroom, as it represents the main social setting where children initiate, develop and experience relationships with their peers.

Classroom peer groups are a form of obligatory rather than voluntary or selective membership. Children are assigned to them; children are not members of a particular group due to their own free choice. As social groups, classrooms consist of collections of individuals among whom there can be a variety of interactions and relationships. In the process of group formation and development, patterns of interactions among the individuals emerge, generating the features and structure of the group which include norms, status and cohesion. The goal of the present study is to understand how these group phenomena that occur within the classroom, such as normative influences, hierarchical organization and interpersonal attractions and repulsions, can explain individual differences in developmental trajectories of aggression.

Classroom norms

Norms have been defined as the acceptable and/or expected behaviour of the members of a social group (Shaw, 1981). Theory and research from social psychology shows that individuals are concerned with these norms and that they will use them to

guide their own behaviour (Miller & Prentice, 1994). Conformity to norms has been a well documented process which seems to operate even when individuals do not agree with the normative behaviour (e.g., Asch, 1987). The motivational mechanisms through which norms have an effect on individual behaviour have been explained in different ways. Social learning theories (Bandura & Walters, 1963) argue that children guide their own behaviour by observing the positive or negative consequences on others for exhibiting specific behaviours. From Bukowski and Sippola (2001) it can be inferred that conformity to norms might result as general tendency of the group to maintain homogeneity, that is, a sense of agreement among the group members about the group's goals, activities, and acceptable behaviours. Prinstein and Dodge (2008) explain that children's desire to fit the expectations of an admired group will drive them to conform to the group's norms. Crosbie (1975) proposed rewards, sanctions and persuasion as mechanisms that bring deviant members back to the limits approved by the group.

The ecological framework proposed by Baker (1968) represents a process-oriented integration of these ideas. He proposed that the functioning of a social setting depends upon feedback circuits that maintain the group's dynamic homeostasis. Applied to normative influences, this idea implies that individuals receive feedback from their environment to determine what behaviours are appropriate. According to Baker, these feedback circuits develop as group self-regulating mechanisms to maintain the setting's activities and stability. Two feedback circuits described by Baker seem particularly relevant for normative processes. One consists of *deviation-countering circuits* which are forms of reinforcement or punishment from the group to maintain individuals within the limits of acceptable behaviour. The *modus operandi* of these circuits is consistent with

the idea that the group tends to reward persons who promote group functioning with power, attention and status (Bukowski & Sippola, 2001). Another feedback system consists of *veto circuits* in which individuals who do not respond to deviation-countering feedback tend to excluded from the group. Victimization, for example, can be seen a peer pressure mechanism to prevent individuals from exhibiting non-normative behaviours (Bukowski & Sippola, 2001; Velásquez, Santo, Saldarriaga, López & Bukowski, 2010). In the present study we focus on the deviation-countering circuits that take place in classrooms to regulate students' aggressive behaviours according to the peer group norms.

Feedback from these circuits can be fed by different forms of norms. Cialdini and colleagues (Cialdini, Reno, & Kallgren, 1990) distinguished between the *is* (descriptive) and the *ought* (injunctive) meaning of social norms. According to them, *injunctive norms* refer to the expected standards of what is approved or disapproved in a social context, while *descriptive norms* refer to the typical observed behaviours in the group, from which individuals may infer what is acceptable. In addition to this, Henry and colleagues (Henry et al., 2000) proposed *norm salience* as a third feedback mechanism through which individuals can deduce the expected behaviours within the group. These authors explain that injunctive norms can be made salient within the group by rewarding approved behaviours or punishing disapproved behaviours.

These different forms of norms have been empirically assessed in prior research related to aggression. Descriptive norms have usually been measured as the prevalence of aggressive behaviours within groups and have been related to aggression-related phenomena. For example, some studies have found that children in classrooms with

many aggressive peers tend to persist in conduct problems over time (Barth et al., 2004; Kellam et al., 1998). Other studies showed that aggressive behaviour was negatively associated with peer acceptance only in groups with low mean levels of aggression (Boivin, Dodge, & Coie, 1995; Chang, 2004; Stormshak et al., 1999; Wright, Giammarino and Parad, 1986). According to the person-group similarity model proposed by Tversky (1977), prevalence of aggression within a group creates an environment that normalizes the behaviour making it more socially acceptable.

In spite of its apparent cogency, this approach to the use of peer norms as predictors of behaviour poses some challenges. One challenge derives from the distinction between descriptive and injunctive norms. The prevalence of a certain behaviour, as manifested in a descriptive norm, may not reflect whether the behaviour is expected, as manifested in an injunctive norm. That is, a group may not actually expect or sanction a behaviour even though it is prevalent. Accordingly, in a study of group effects one should assess descriptive and injunctive norms. However, there are also statistical challenges. When studying the effect of group processes on individual outcomes, one needs to account for the hierarchical structure of the social context. In the case of groups and individuals there is a need to recognize the lack of independence between these two levels of social complexity. Proper analysis of this kind of data requires the use of hierarchical linear models which convert the estimation of individual outcomes into group means, thus making the outcome being the same as the predictor. As a consequence, estimating group features from individual aggregates needs special caution.

This concern about recognizing the distinction between descriptive and injunctive norms has been discussed before. Opp (1982), in his evolutionary theory of norm

formation, states that beyond the frequency of a particular behaviour, it needs to be accepted by the group so as to become a norm. In this sense, injunctive norms would be a more accurate approximation to normative processes. According to Henry (2008), this type of norm can be operationalized in two ways: as *actual* or as *perceived* injunctive norms. Actual norms correspond to individuals' normative beliefs, that is, personal cognitions regarding the acceptability of a particular behaviour (Huesmann & Guerra, 1997). Perceived norms, on the other hand, are individual perceptions of what others think is acceptable (Ozer, Weinstein, Maslach, & Siegel, 1997). The first concept is based on social cognitive theories that underscore the role of beliefs in the regulation of behaviour (Bandura, 1986), whereas the second is based on the theory of reasoned action (Fishbein & Ajzen, 1975), which states that individuals guide their behaviour based on what they think others expect.

There are at least two reasons to expect that using perceived injunctive norms might be a better approach to understanding individual change than is using actual injunctive norms. First, the same argument outlined for descriptive norms with regards to the need to use hierarchical lineal models as the preferred analysis technique for embedded data applies to actual injunctive norms. Therefore, it would be not be possible to establish whether the outcome (i.e., group mean levels of normative beliefs) is a characteristic of the individuals or the group. Second, previous research has shown that individuals tend to act in accordance with the collective norm rather than with the private norm (Miller & Prentice, 1994). For instance, in a well known set of studies, university students were observed to adjust their consumption of alcohol, their drug use and their sexual behaviour so as to approximate what they believed to be the social norms of their campus, even

when the perceived norm differed from the actual norm (e.g., Martens et al., 2006). Another example comes from research in bullying. Findings in this area suggest that bystanders might passively approve the bully's aggression by not challenging their aggressions. Then, aggression might become increasingly tolerated by the group, although it might not reflect individuals' normative beliefs (Juvonen & Galván, 2008). Based on this, one of the purposes of this study is to assess aggression-related injunctive norms, based on the perceptions that students have about their classmates' approval of aggression.

As a third feedback mechanism, Henry and colleagues (Henry et al. 2000; Henry 2008) have looked at ways in which injunctive norms are made salient (i.e., norm salience). Particularly, they focused on the social rewards and punishments that children receive from their peers for exhibiting aggressive behaviours. They proposed social status, in terms of acceptance and rejection, to be fundamental reinforcement mechanisms that are integrated in the deviation-countering feedback circuits within classrooms. In a study that looked at changes over time in aggression, they operationalized norm salience as the within-classroom association between aggression and acceptance, and between aggression and rejection. They also assessed the extent to which the teachers made norms against aggression salient by using reprimands. Their results showed that peers alone did not have an impact on changes in aggression. Decreases in aggression were found only when both peers and teachers made the norms salient. This finding can be linked to findings from studies looking at variations in the association between acceptance and levels of aggression in peer groups (Boivin et al., 1995; Chang, 2004; Stormshak et al., 1999; Wright et al., 1986). Although these studies

regarded group mean levels of aggression (descriptive norms) as the predictor (moderator) of the association between aggression and acceptance, the concurrent nature of their data allows making inferences in the other direction. That is, in groups where aggression is strongly related to acceptance, group levels of aggression are also high. These results call for a more thorough examination of the interrelation between norms and status and of the direction of these associations.

Classroom hierarchical organization

An emergent structural feature of most social organizations is the hierarchical organization of individuals' positions. As a group phenomenon, this process has been studied in peer relations research as the popularity hierarchy, that is, the group's view of an individual in terms of liking and disliking (Bukowski & Hoza, 1989; Parker, Saxon, Asher & Kovacs, 1999). In the context of normative influences, two important issues seem to be particularly important to note about the emergence and function of groups' hierarchical organization. On the one hand, it seems that status is granted to individuals in the group who exhibit acceptable behaviours (Coie & Kupersmidt, 1983). On the other hand, these patterns of interaction define who is in the position to influence who. As such, high status individuals become group leaders who are looked up to by others in the group (Prinstein & Cillessen, 2003). As a consequence, high status individuals become models for acceptable (normative) behaviours within the group.

Juvonen and Galván (2008) illustrated how group norms and hierarchical organization are interrelated in a peer social phenomenon such as bullying. They show that bullies tend to be perceived as popular by their peers during early adolescence (Gest, Graham-Bermann, & Hartup, 2001; La Fontana & Cillessen, 1998; Parhurst &

Hopmeyer, 1998). As outlined earlier, bystanders may passively approve bullying behaviours by not challenging the bully and by conferring him or her higher status. As a result, high-status bullies become the reference for aggression normativeness. In fact, evidence from a study showed that students who considered bullies as popular became themselves more antisocial during the school year (Juvonen & Ho, in press). Juvonen and Galván explain that others might try to emulate the behaviour from high-status peers for two reasons: to maintain or increase their own status, or to protect themselves. This set of ideas, in conjunction with Henry and colleagues' (2000) results regarding norm salience, indicates that the behaviours of high-status children are especially important for the functioning of normative processes in these peer groups. It may be that the weak findings from Henry and colleagues' study related to norm salience could be due to the fact that within-group correlations between aggression and popularity were based on all children in the class rather than on the high status children. This extremely general indicator may have failed to fully capture the extent to which aggression is approved and made salient in the group by means of high status. In the present study this problem was avoided by the use of normative aggression indicators based on the data from only the high-status children within each classroom instead of on data taken from all the children.

An important issue that needs to be considered when examining popularity hierarchies within peer groups is that two different forms of high status can be identified: *actual* and *perceived* (i.e., acceptance and popularity) (Rubin et al., 2006). Whereas acceptance refers to how a child is actually liked by her/his peers, popularity is a perceptual indicator of the standing or status in the group. There are at least two reasons to hypothesize that perceived popularity would be better related to group normative

processes. First, it has been stated that "whereas being liked or accepted occurs at the dyadic level (i.e., one person has affection for someone else), the perception of someone as being popular in a classroom or school reflects a group level of analysis (i.e., the person is perceived according to her/his position in the group)" (Rubin et al., 2006, p. 598). Therefore, perceived popularity might be a better indicator of who has more power to influence others at the group level. Second, Prinstein and Cillessen (2003) noted that the group leader may have to be forceful and strong at times, which may lead them to include some levels of aggression in their behaviour. Based on this, in this study we compared the power of normative influence that accepted and perceived popular children have, in terms of aggression-related norms.

Classroom cohesion

Apart from the norms and the hierarchical organization of the group, another structural aspect that has been studied in peer groups is cohesion. Cohesion has been broadly defined as the level of attraction within the group (Cartwright, 1968). Based on this definition, interpersonal attraction has been seen as the basis for attraction to the group (Newcomb, 1960; Shaw, 1974). In peer relations research, it is conceptualized as the "structural integration within the group that connects and links individuals to each other" (Bukowski & Sippola, 2001, p. 359). As such, group cohesion can be manifested in the mutual positive attitudes among classmates (Lott, 1961). From this point of view, we operationalize classroom cohesion as the density of mutual attractions among students that can be derived from sociometric measures.

There are two reasons for studying classroom cohesion as part of the processes underlying the development of aggression. First, classroom cohesion is likely to be

related to how meaningful a group is for the individuals in it. To the extent that a group is meaningful for an individual, he or she will use the group's norms as a referent point to guide their own behaviour. According to reference group theory (Richer, 1976), social settings have an effect on the individuals' covert or overt behaviour as they become a reference for their behaviour. Implicit in this argument is the assumption that for the reference process to operate, groups need to be meaningful for the individuals within them. Compared with other types of peer groups (e.g., cliques, crowds), membership in classrooms is obligatory. Given this, preference does not play a role in the selection process of individuals into classrooms. Accordingly, the extent to which the classroom is significant for the students may vary. For some students, their bond with the group and their desire to maintain their belongingness to it may be very strong, while for others it may be very weak. There is reason to believe that individuals who have a strong sense of belongingness to the classroom and an interest in remaining as member of it will probably be more concerned about maintaining its status quo by following the group norms. Accordingly, Hare (1976) proposes that members of groups "…who are highly attracted to the group either for its prestige, its productivity, or the friendship of its members will conform more to the standards of the group than will members who place low value on this criteria"(p. 30). In fact, studies conducted with work groups have shown evidence of this interaction (e.g., O'Reilly & Caldwell, 1985). Therefore it is our contention that studying classroom norms without taking into account how strongly students are connected will provide an incomplete picture of the influential role of the classroom norms on children.

The second reason has to do with the direct effects that classroom cohesion may have on students' aggression. Cohesion can be taken as an index of the quality of the relationships among the students. Thus, if students are highly bonded, they might care more about their classmates' well-being. Presumably, in more caring communities the level of harm among their members should be low, which would translate into low levels of aggression. In light of these two oppositional forces of cohesion, one that might make students be sensitive to the negative influence of aggressive norms, and another that might prevent students from harming each other, this study examined the direct and moderating effects of classroom cohesion on changes in students' aggression across the school year.

Trajectories of aggressive behaviour

Two approaches have been used to examine short-term and long-term changes in aggressive behaviour. Traditional approaches to evaluate change in aggression have examined the mean linear growth trajectory of particular samples and the deviations from the average trajectory. In this approach, individuals are assumed to belong to a single population with a common pattern of change (Martino, Ellickson, Klein, McCaffrey, & Edelen, 2008). Examples of this type of analysis are studies conducted with linear growth curve models or with multilevel analyses where repeated measures are nested within individuals (e.g., Farrel, Sullivan, Esposito, Meyer, & Valois, 2005; Mercer, McMillen, & DeRosier, 2009; Murray-Close, Ostrov, & Crick, 2007; Ostrov, Ries, Stauffacher, Godleski, & Mullins, 2008; Williams, Conger, & Blozis, 2007).

More recent approaches have looked for heterogeneity in patterns of change within broadly defined samples. It assumes the possibility of existence of distinct developmental

trajectories (Muthén, 2004; Nagin & Tremblay, 2005). As an illustration, long-term longitudinal studies looking at changes in physical aggression have given support to the existence of different classes of pathways of aggression from childhood to adolescence (Brame, Nagin, & Tremblay, 2001; Nagin & Tremblay, 1999; Schaeffer, Petras, Ialongo, Poduska, & Kellam, 2003; Shaw, Gilliom, Ingoldsby, & Nagin, 2003; Martino, et al., 2008). Results from these studies suggest that the most common trajectories are: a) the normative, low or no-aggression trajectory, b) the persistent-high aggression trajectory, c) the desisting trajectory, and d) the increasing trajectory. These trajectories are defined on the basis of an intercept (i.e., starting point) and a slope (i.e., the rate and shape of change). Based on this evidence, this study will examine whether similar pathways can be identified in short-term changes of students' aggression across a school year.

An important consideration to be taken into account in looking at changes in aggression is gender differences. According to Little, Jones, Henrich and Hawley (2003), aggression can be classified based on the form of "delivery" of the behaviour, in which case it can be either direct or indirect. Direct aggression refers to any physical or verbal act that directly hurts another person, while indirect aggression corresponds to harm caused to the other's relationships and is usually inflicted in such a way that the victim is not able to identify who hurt him/her. Aside from the apparent overall difference in the amount of aggression shown by boys compared with girls, research findings suggest that boys are more prone to exhibit direct (e.g., physical) rather than indirect forms of aggression, whereas girls are more prone use indirect (e.g., relational) rather that direct aggression (Crick et al., 1999; Leaper, 2000). Although most studies have focused on more "male-typical" forms of aggression, namely physical, some research has started to

look at different developmental trajectories of relational aggression development for boys and girls during middle childhood (Underwood, Beron, & Rosen, 2009; Vaillancourt, Miller, Fagbemi, Cote, & Tremblay, 2007). These studies suggest that pathways of physical and relational aggression might differ. This led us to test the existence of differential forms of development of direct and indirect aggression over the course of the school year. Moreover, we examined whether aggression development could be distinguished in boys and girls, both in terms of their overall levels of aggression, as well as of differential trajectories in both types of aggression.

In sum, the purpose of this study was to answer four questions. First, we wanted to explore whether aggression development during one school year can be represented as sub-populations of children whose patterns of change can be distinguished in distinct trajectories. Specifically, we looked for different developmental pathways for boys and girls in their use of direct and indirect aggression. Second, we examined whether change over time in aggression could be predicted by aggression-related classroom norms at the beginning of the year. To accomplish this, we tested if membership in different developmental trajectories was predicted by perceived injunctive norms and norm salience. Third, we asked whether classroom cohesion would prevent children from showing high or increasing levels of aggression. Thus, we tested if membership in different developmental trajectories would be predicted by levels of cohesion in the classroom at the beginning of the year. Finally, we wanted to know if conformity to negative norms would be enhanced or prevented in highly cohesive classrooms. Therefore, we examined if classroom cohesion acted as a moderator of the effect of classroom norms on aggression development. Altogether, our goal was to further our

understanding of the effect of the classroom peer group structure on aggression development during the school year. By this we expected to contribute in the building of solid research-based knowledge to inform school-based interventions aimed at preventing aggression.

Method

Participants

Boys and girls (N = 1595, mean age = 10.2 years; 53% male; 47% female) from 63 fourth-, fifth- and sixth-grade classrooms in nine mixed-sex schools in Bogotá, Colombia took part in the study. Participants were enrolled in schools drawn from neighborhoods that varied in socio-economic status (SES). Individual SES was assessed via parental reports about their household stratification, which is defined in Colombia by the local government and ranges from 1 (low) to 6 (high). Classroom size ranged from 11 to 32 participants (mean size = 21.98).

Procedures and instruments

After receiving permissions from the appropriate school personnel, the research teams visited the classrooms of the potential participants to tell the students about the purposes and procedures and to give them a letter and consent form for their parents. Children who returned a consent form signed by their parents were allowed to take part in the study. The overall participation rate of the study per classroom was, on average, 79%. At the end of the study, participants received school supplies as a gesture of our gratitude for their willingness to help us.

At each of four time points separated by roughly eleven-week intervals (Time 1 = March; Time 2 = May; Time 3 = August; Time 4 = October) the participants completed a

multi-section questionnaire during a one-hour in-class session. Four set of variables were measured. They were aggression, classroom injunctive norms, classroom norm salience and classroom cohesion. The techniques used to measure these variables are as follows.

Aggression. An unlimited choice peer assessment procedure was used to assess direct and indirect aggression (see Rubin et al., 2006 for a description). Each student was given a list of all the participating children in their class and a list of several characteristics and behaviours. They were asked to indicate which of their participating classmates fit each characteristic or behaviour in the list. Four items were used to assess direct aggression (e.g, someone who hits or pushes other people; Cronbach's alpha at Time 1= 0.94), and two items were used to assess indirect aggression (e.g., someone who talks bad about others behind their backs to hurt them; Cronbach's alpha at Time 1 = 0.87). Each child was given a score on each item indicating how often she/he had been nominated for it by her/his participating classroom peers. The average of the items corresponding to each type of aggression was calculated to obtain a score for each variable at each of the four time points.

Classroom injunctive norms. Based on Henry's (2008) description of perceived injunctive norms, we created a self-report questionnaire assessing the perceptions of students about their classmates' attitudes towards aggression. Items were taken from the Huesmann and Guerra's (1997) *Normative Beliefs About Aggression Scale* and then modified for the present study. Using a 5-point Likert scale, students rated their level of agreement (1 = total disagreement, 5 = total agreement) with three items (e.g. "My classmates think it is ok to be aggressive"; Cronbach's alpha at Time 1= 0.83). An average of the items at Time 1 was calculated for each student and these scores were then

aggregated at the classroom level. Two intraclass correlation (ICC) indices were used to test the reliability of students' ratings of this classroom feature. ICC(1), which corresponds to the ratio of between-classrooms variability and total variability (i.e., between- plus within-classroom variability), is an index of the reliability of individual student's ratings. ICC(2), on the other hand, is a function of ICC(1) and the number of students who provide ratings for each class. As such, it is an index of the accuracy of the class-mean (see Lüdtke, Trautwein, Kunter, & Baumert, 2006, for a review). Our analyses showed that the ICC(1) for this variable at Time 1 was .09 and the ICC(2) was .71. Based on the .70 cutoff value for ICC(2), our results indicate that the reliability of the student's aggregated scores of perceived injunctive norms was sufficient.

Classroom norm salience. Based on Henry et al. (2000) we estimated two types of indices of aggression-related norm salience. One corresponded to the within-classroom correlation between popularity and aggression, and the other was the mean level of aggression of the most popular children within each class. For this, we assessed the association between an overall index of aggression, and the two measures of popularity. The index of aggression was estimated by averaging the above-described indices of direct and indirect aggression. As for popularity, both actual popularity (i.e., acceptance) and perceived popularity were considered. Acceptance was assessed by means of a sociometric questionnaire (Bukowski et al., 2000; Newcomb & Bukowski, 1983). Students received a list with the name of each of the participants in their class. Same-sex peers and other-sex peers were grouped is separate lists. Children were asked to rank order their friends by indicating their first best friend with a "1", their second best friend with a "2", their third best friend with a "3", and any other friends with a "4" (they were

free to chose as many "4" as they wanted). They completed this procedure first for same-sex peers, and then for other-sex peers. The acceptance score was the number of times a child was chosen as a friend by both same-sex and other-sex peers regardless of the order in which they were chosen. This is consistent with measures of acceptance that assess the level of likeability of the children within the peer group via nominations of who are the most liked children in the group (Bukowski et al., 2000). Popularity was measured with the same peer assessment procedure used to measure aggression. Children nominated those children in their class who fit the description "someone who is popular". The score on this variable was the total number of times he/she was chosen for it by his/her peers.

To calculate the first type of norm salience, two indices were calculated with Time 1 data. They were (a) the correlation between acceptance and aggression, and (b) the correlation between perceived popularity and aggression. For the second type of norm salience index, students whose popularity scores were one standard deviation above their classroom mean were identified and their overall aggression scores were averaged for each class. Therefore, we generated two of these measures of norm salience for each classroom at Time 1: (a) the mean level of aggression of the most accepted children, and (b) the mean level of aggression of the children perceived as most popular.

Classroom cohesion. Based on Cartwright's (1968) definition of cohesion, we operationalized this variable as the number of mutual attractions among the members of a group. Accordingly, we used sociometric positive mutual nominations at Time 1 as our index of classroom cohesion. Using the sociometric questionnaire described above, we counted all positive nominations given by each child's, which consisted of all same-sex and other-sex friend choices, regardless of ranking. Then, we calculated how many of

those choices were reciprocated with positive nominations by the nominated friends. Finally, we aggregated the number of mutual positive nominations at the classroom level.

Adjustment for classroom size differences in the nomination measures. Considering the difference in the number of nominators among classrooms (range = 9 to 31) scores in aggression, perceived popularity and acceptance were corrected for classroom size (i.e. number of nominators). To maintain the original scale of the items (i.e. number of nominations received), a regression-based procedure was used, as opposed to standardization of the scores (for a review of the procedure see Velásquez, Bukowski, & Saldarriaga, 2010a). First, we examined the extent to which changes in the classroom size had an effect on the number of nominations received in each variable. To do this, the linear and quadratic effects of the number of nominators (classroom size minus one, that is, the person receiving the nominations) were used as predictors of the variables' scores at each time point for aggression and at Time 1 for acceptance and perceived popularity (because classroom norms were only examined at the beginning of the school year). These scores were aggregated at the classroom level and analyzed with the classroom as the unit of analysis, to account for the nested nature of the data.

Results for acceptance showed that the variance explained by the classroom size was 8% at Time 1. Unstandardized *B* coefficients for the linear effect indicated that for each unit increase in the group size, students' acceptance scores at Time 1 increased by .141. In addition, *B* coefficients for the quadratic effect showed that for each squared unit increase in class size, scores in acceptance decreased by .004. Based on this, a group of 20 students (which represents the median) was selected as the reference group size. Scores in classrooms that were larger or smaller than 20 were adjusted according to the

extent of this difference. The amount of the adjustment was computed using the weights observed for the linear (e.g., .141 for Time 1 scores) and curvilinear (e.g., .004^2 for Time 1 scores) effects in the regression analysis. For example, Time 1 scores of participants from a classroom of 21 students was reduced by .140, whereas the scores of participants from a classroom of 22 students were reduced by .281, and so on. In these adjustments, scores of zero were not changed.

A similar procedure was conducted for the aggression and perceived popularity scores. However, to make variables in the peer assessment questionnaire comparable to each other, in these adjustment procedures an average was initially computed based on all 12 items of the peer assessment questionnaire. This average score was used as the dependent variable in the regression analysis. The analysis showed that at each time point (i.e., 1 though 4) classroom size explained 39%, 24%, 20%, and 28% of the variance in the peer assessment scores. The observed unstandardized B coefficients for the linear effect at Times 1 through 4 were .159, .125, .127, and 147, respectively; the observed coefficients for the quadratic effects were -.006, -.003, -.004, and -.001, respectively. Perceived popularity scores at Time 1 and aggression scores at Time 1 through 4 were then adjusted according to the steps described for acceptance scores.

Finally, the same procedure was applied to the mutual friendship nominations scores, calculated for the cohesion index. Given that for the purposes of this study we were interested in initial levels of classroom cohesion, only results from Time 1 are reported. Results from the regression procedure showed that class size explained 3% of the variance. The linear B coefficient was -.038 and the quadratic coefficient was -.002. Although the linear results seem counter-intuitive in terms of the direction of the effect,

more nominators imply larger classes, something that adds a counter-balancing force to the positive effect of classroom size. In such crowded contexts, it is reasonable to believe that cohesion might decrease. In our sample, this negative force seemed to be somewhat stronger compared to the positive force of a higher number of nominators. Therefore, the scores were adjusted accordingly.

Missing data

Given drop-out, inclusion of new students or incomplete responses, a range of 2 to 15% of the data was missing for the variables used in the current study. Results from Little's MCAR test conducted with the items indicated that data was not missing completely at random ($\chi^2_{(2256)}= 3280.30$, p < .05). Under the assumption that the data was missing at random, multiple imputation of 20 data sets was conducted using the software known as *Amelia II* (Honaker, King, & Blackwell, 2009). Although for the purposes of this study only a set of variables were considered, all the items assessed in the questionnaire were used in the imputation model. To account for the hierarchical nature of the data set, classroom identification was dummy coded and included in the imputation procedure. Based on Rubin's (1987) rules for estimating within- and between-imputation variance, the fraction missing for our items of interest ranged from 9 to 10%. This rate of missing information, combined with the 20 imputations conducted, represented a level of efficiency of 99% in the estimation of the items' means. All the data analyses, were conducted with the software Mplus, version 5.21 (Muthén & Muthén, 1998-2008), using the imputation option which allows including all the imputed data sets as input and estimating accurate standard errors, following Rubin's rules.

Data analysis

Growth Mixture Modeling (GMM) (Muthén, 2004) conducted with the Mplus (version 5.21) statistical software package (Muthén & Muthén, 2007) was used (a) to identify differential trajectories of aggression development over the course of the school year, and (b) to predict membership in these trajectories on classroom level covariates. Given the presumed sex differences in the use of direct and indirect aggression, these models were estimated separately for boys and girls and for each type of aggression.

First, Latent Growth Curve (LGC) models (Bollen & Curran, 2006) were estimated for each type of aggression and sex. Via structural equation modeling, two latent factors, representing the intercepts and slopes of aggression, respectively, with shared multiple indicators were estimated. In this case, the multiple indicators correspond to univariate repeated measures of the aggression outcome at the four time points. The parameters estimated for these growth factors represent the mean initial status (i.e., intercept) and mean rate of change across the four time points (i.e., slope) for the sample, as well as the variability of individuals around those mean growth factors. Both linear and quadratic forms of change were considered and compared in the estimation of these models, to account for possible curvilinear changes.

Based on the assumption that all participants did not belong to a single population, GMM was applied next. This technique is based on a categorical latent framework where distinct patterns of growth (i.e., classes) can be identified and each individual's most likely class can be estimated based on posterior probabilities. Three criteria were used to decide on the number of classes. First, we compared the Bayesian information criterion (BIC) between models with k and k-1 classes (where k equals the number of classes). A model with k classes and a smaller BIC than a model with k-1 classes was considered

better. Second, we looked at the average latent class probabilities (ALCP). Solutions that showed probabilities above 80% for most likely latent class membership were considered appropriated models. Third, and more importantly, models in which trajectory shapes were meaningful and distinct from one another were preferred.

Once differential trajectories of direct and indirect aggression were identified for boys and girls, demographic covariates were included as predictors of the growth indices and of class (i.e. type of trajectory) membership. Given that trajectories are treated as categorical variables, the effects of these continuous covariates were estimated by means of multinomial logistic regressions. Finally, multilevel analyses were used to include classroom covariates as predictors of class membership. These analyses estimate whether there is variability across classrooms in the proportion of students in each trajectory. Specifically, they test whether the log-odds of being in a given trajectory versus being in a reference trajectory varies across classrooms. Then, classroom covariates are used as predictors of such variability, by means of regression analyses. A graphic representation of these analyses is depicted in figure 2.1.

Results

Descriptive statistics

Observed means, standard deviations, standard errors, and intraclass correlations for females and males on the measures of direct and indirect aggression are presented in Table 2.1. Overall, boys showed higher scores than shown by girls on the measure of direct aggression. No differences are seen on the measure of indirect aggression. For girls, the mean for indirect aggression was higher than the mean for direct aggression; the opposite pattern was found for boys. Both boys and girls showed an increment in the

intraclass correlations over time for both types of aggression. This pattern could be an indication that as the school year continues children become more similar to their classroom peers in the aggressive behaviours. Finally, it should be noted that girls' ICC in direct aggression was higher compared to boys' ICC.

Correlation coefficients among the individual-level variables and among the classroom-level variables are shown in Tables 2.2 and 2.3. At the level of the individual, correlations for the aggression variables over time were very strong indicating that these variables are very stable. Consistent with the findings reported already for the mean differences between boys and girls, these results show that girls had lower scores on the measure of direct aggression whereas no effect was found for sex on indirect aggression. In terms of SES and age, these demographic variables were positively correlated with aggression, but the strength of these associations was relatively weak.

On the other hand, at the level of the classroom, moderate to strong correlations were found among the norm salience measures, indicating some level of communality among these forms of measurement. The associations between perceived norms and the measures of norm salience were moderate, except for the case of norm salience measured by the correlation between aggression and acceptance. Perceived norms and norm salience, as measured by the level of aggression of accepted and popular children in the classroom, were found to be positively associated with levels of direct and indirect aggression across time. As for cohesion, we found that it was negatively related to both forms of aggression, but this relation was only statistically significant for direct aggression.

Structural equation modeling was used to estimate changes in each type of aggression across the four time points, through latent growth curves. Three nested models were compared by means of chi-square difference tests to define whether the general pattern of change of the data was linear or curvilinear. The first model was "unconstrained." In this model the first and last growth factor loadings fixed at 0 and 1, respectively; Time 2 and Time 3 loadings were free to be estimated. The second model was a linear model in which the growth factor loadings were fixed at 0, 1, 2 and 3, respectively for each time point. The third model added a quadratic growth factor in which loadings of 0, 1, 4 and 9 were used for Times 1 through 4, respectively. All models included an index factor representing the intercept to which the four factor loadings were fixed at 1. These models were estimated adjusting fit indices and standard errors to the clustered nature of the data (i.e., students within classrooms). As shown in Table 2.4 comparisons between the models revealed that the linear model did not worsen the fit of the data compared to the unconstrained model, and that the curvilinear model did not improve the fit of the data, compared to the linear model. For these reasons, only the linear growth factor was retained for further analyses.

Trajectories of indirect and direct aggression for boys and girls

Based on the empirical findings reported by studies looking at changes in aggression over the course of childhood and adolescence (e.g., Nagin & Tremblay, 1999; Martino et al., 2008; Underwood et al., 2009; Vaillancourt et al., 2007), we were expecting to find three to four distinct trajectory classes. Accordingly, we tested one-, two-, three-, and four-class models using the Growth Mixture Modeling procedure. For the estimation of the latent classes, three constraints were imposed: (a) the residual variances of the

aggression outcomes were constrained to be equal across time and across classes; (b) covariances between the growth factors (i.e., intercept and slope) were fixed at zero, for all classes; and (c) the variances of the growth factors were constrained to be equal across classes. (The means of the growth factors were allowed to be estimated separately for each class. Given that students were nested within classrooms, clustering was accounted for by using the *TYPE = COMPLEX* option of the analyses conducted in Mplus.

For direct aggression in boys, a three-class solution fit all our criteria. The BIC for three classes was smaller than the two-class solution ($BIC_1 = 15412$; $BIC_2 = 15271$; $BIC_3 = 15194$), and the ALCP ranged from .81 to .95. Because a four-class solution did not converge it was assumed that the three class model presented an optimum fit to the data. Class 1 represented a low-stable group (77%), Class 2 was a high-increasing group (15%) and Class 3 was high-desisting group (8%).

A three-class solution seemed also appropriate for indirect aggression in boys. The BIC was better than for models with less classes ($BIC_1 = 14368$; $BIC_2 = 14244$; $BIC_3 = 14155$), and the ALCP ranged from .81 to .96. The four-class solution was rejected given that two classes showed a very similar shape, with only a slight change in the intercept. Similarly to direct aggression, Class 1 corresponded to a normative low-stable group (83%), Class 2 represented a high-increasing group (8%) and Class 3 was high-desisting group (9%).

Three class solutions were also observed with the girls' data for both direct and indirect aggression. The three-class models showed smaller BICs than the two-class models ($BIC_1 = 10595$; $BIC_2 = 10403$; $BIC_3 = 10269$, for direct aggression; and $BIC_1 = 12631$; $BIC_2 = 12511$; $BIC_3 = 12489$, for indirect aggression). As for the ALCP, the

three-class solutions yielded adequate posterior probabilities, ranging from .91 to .99 for direct aggression, and from .80 to .97 for indirect aggression. Four-class solutions were not considered because the models failed to converge. Both types of aggression showed similar patterns of change. Class 1 was a normative low-stable group (92% for direct aggression, 87% for indirect aggression), Class 2 represented a low-increasing group (3% for direct aggression and 9% for indirect aggression) and Class 3 was a high-stable group (5% for direct aggression and 4% for indirect aggression). Figure 2.2 depicts the trajectories observed.

Predictors of trajectory membership

To determine whether membership in a trajectory could be predicted by classroom cohesion and aggression-related norms, data were analyzed with Multilevel Mixture Modeling. We tested several models in which classroom cohesion, classroom norms and the interaction between them were included as predictors of individual-level latent class intercept.

To test these models, the effects of demographic variables (i.e., age, SES), centered at their grand mean, were accounted for at the within-classroom level. These effects were tested in a set of multinomial logistic regressions in which the effects of the predictors on the conditional probability of class (i.e. trajectory) membership is compared to a reference group. The results of these analyses are expressed as odds ratios, where positive and significant coefficients indicate that as the predictor increases, there is a higher likelihood of following the trajectory of interest, compared to the reference trajectory. The converse would be interpreted in the case of negative odds ratios.

Separate analyses were run for boys and girls and for type of aggression. Given that the five measures of peer norms (i.e., perceived norms, aggression of accepted children, aggression of popular children, association of aggression and acceptance, and association of aggression and popularity) were in most cases strongly associated, to avoid multicollinearity we also conducted separate analyses per type of norm. For each model, variations in the classroom proportion of students in each trajectory as a function of classroom norm, cohesion, and the interaction between these two variables were examined. Results from these analyses are summarized in Table 2.5. Several findings should be noted. First, from the five measures of classroom norms, the one with the highest predictive value (i.e., major capacity to differentiate between trajectories) was the index of the level of aggression of the children perceived as being the most popular in the classroom (see Model 3 in Table 2.5). It can be seen that in groups where popular children were more aggressive there is a higher proportion of boys who increase in direct and indirect aggression and a higher proportion of girls who increase in indirect aggression, compared with the proportion of children who remain low in their levels of aggression over the course of the school year. Second, measures of norm salience based on levels of aggression of high-status children appeared to have a higher predictive value than the measures based on correlations. Third, differences in the proportion of boys who decreased in their level of aggression (i.e., desisters), compared with boys who increased, were only predicted by acceptance-based norm salience measures. Fourth, high scores on the measure of classroom cohesion was observed to be related to lower classroom proportions of children in high-risk trajectories (i.e., high-stable, increasing), particularly for direct forms of aggression. Finally, although only a few interactions between

classroom norm and cohesion were found to be significant, it should be noted that these interactions were positive for girls and negative for boys. In other words, classroom cohesion tended to strengthen the negative effect of aggression-related norms for girls, whereas it tended to protect boys from these negative effects.

Discussion

The goal of the present study was to examine the effect of the classroom social structure on children's aggression development over the course of the school academic year. To achieve this goal, our first step was to model changes in aggression over time. Following the premise that there might be different subpopulations of children who follow distinct developmental trajectories (Muthén, 2004; Nagin & Tremblay, 2005), we looked for latent growth trajectories in our sample. Taking into consideration the long-standing discussion about gender differences in the use of different forms of aggression, namely direct and indirect, we explored these growth trajectories separately by sex and type of aggression. Our descriptive statistics gave an initial support for this decision. Between-sex comparisons showed that boys exhibited higher levels of direct aggression compared to girls, while levels of indirect aggression were similar for boys and girls. Within-sex comparisons, on the other hand, showed that while boys tend to use more direct than indirect forms of aggression, the converse was found for girls. Looking at patterns of change over time, our results were consistent with previous studies that have looked at changes in aggression across childhood and adolescence (e.g., Schaeffer et al, 2003; Martino et al, 2008). Overall, we found the four most common trajectories described in those studies: a normative low-stable trajectory, a high-stable trajectory, an increasing trajectory, and a desisting trajectory. However, these results differed by

gender and type of aggression. For girls, only the low, high and increasing trajectories were found, while boys showed the low, increasing and desisting trajectories. Within gender, differences in the types of aggression were found in the mean levels of each type of aggression, but not for the shapes of change. To our knowledge, this is the first study that looks at patterns of change of children over the course of a school year. It was noteworthy that only a relatively small proportion of children showed mean changes in their behaviour over time. This pattern suggests that mixing these children in the analyses with the normative population that maintains a low-stable expression of the behaviour might reduce an important proportion of variability to be explained. Therefore, these results are a powerful demonstration of the remarkable utility of growth mixture modeling as a means of capturing variability in individual development.

Regarding the effects of classroom norms on aggression trajectories, this study supports the idea that the behaviour of individuals is affected by group norms (Miller and Prentice, 1994). On the one hand, we observed that the intraclass correlations of aggressive behaviours tended to increase from the beginning to the end of the school year. As an index of the similarity of individuals within groups, these changes in the ICC indicate that, over time, students within a classroom might ascribe to the norm of the peer group as they tend to become more similar to one another over time. On the other hand, results from the multilevel mixture models showed that norm salience had an important effect, particularly in distinguishing those children who tended to increase their levels of aggression, from those who remained low in aggression. This finding is consistent with Baker's (1968) ecological framework that explains that social settings make use of the feedback circuits to regulate the behaviours of individuals who belong to them. In

particular, the effect of norm salience functions as a deviation-countering circuit where individuals who ascribe to the norm are rewarded with status by the social system. As a consequence, others will try to emulate the behaviours of high status peers in order to increase their own status or to avoid being rejected by the group (Juvonen & Galván, 2008).

Two indices of functioning within the group, specifically perceived popularity and acceptance were used as measures of peer status to measure norm salience. As we expected, measures of norm salience based on perceived popularity had a somewhat better predictive value in normative peer influence processes compared with measures based on acceptance. Consistent with Rubin et al. (2006) this set of results provides evidence that this measure of peer status is a better index of group phenomena and captures more precisely who in the group has the power to influence others. It should, however, be noted that only norm salience measures based on acceptance made a difference in terms of the distinguishing boys who increased versus boys who decreased in the aggressive behaviours. Given that acceptance refers to experiences of affect rather than to experiences of status or power it is possible that the potential for interpersonal influence, especially positive forms of influence (e.g., , reducing maladaptive behaviours) may be heighten by personal affect-based experiences rather than by perception of whom is powerful at the group level.

The direct and moderating effects of classroom cohesion were also explored in this study. Using correlational indices and on results from multilevel mixture models, it was shown that classroom cohesion was negatively associated with aggression, particularly with direct forms of it. Specifically, in highly cohesive classrooms, the proportion of

children following high-risk trajectories of direct aggression tended to be smaller compared to low cohesive classrooms. These findings provide evidence of the protective role of cohesion in the development of aggression. On the other hand, results from the moderating effects of classroom cohesion yielded only a few significant interactions. From these interactions, it could be concluded that in classrooms where students feel mutually attracted to one-another, girls are more prone to be negatively influenced by the aggression-related norms, whereas boys tend to be protected from it. However, given the inconsistency of the interactions across the different types of norms and aggression, these results are inconclusive and need further exploration.

A particular strength of this study was that it used multiple measures of group norms. Typically, group norms have been captured in research as aggregates of individual behaviours (i.e., descriptive norms) or personal beliefs (i.e., actual injunctive norms). Although these indices can approximate what is tolerated in the group, they have two important limitations. First, descriptive norms represent prevalent behaviour, yet they do not assess whether the behaviour is expected or approved by others in the group. Second, as it has been described for the bullying phenomena, it is possible that personal injunctive norms differ from the group norms (Juvonen & Galván, 2008). Hence, we tried to tap more directly into the approval aspect of the construct, focusing on the group as the reference of measurement. This led us to center our attention on perceived injunctive norms and norm salience. We found that perceived norms explained intercept differences but not so much changes over time. Specifically, in classrooms where students perceive that their classmates approve aggression, higher proportions of children

start and remain high in levels of aggression, compared to those who start and remain low.

Norm salience, on the other hand, was a better predictor of change. Building upon the work by Henry and colleagues (2000), we looked at forms in which aggression is rewarded in the group, thus making the norm salient. Two improvements were added to the measures proposed by Henry et al. First, instead of looking at within-classroom correlations, we focused on the level of aggression of those children who are rewarded by the group with peer status. Since correlations include all children in the classroom, they could obscure the power that high status peers, in particular, might have on influencing others behaviour. Our second improvement was to measure perceived popularity in addition to acceptance. It has been argued that perceived popularity, relative to acceptance, is a better index of group phenomena (Rubin et al., 2006), so we wanted to compare the predictive value of both types of peer status. In fact, we found that norm salience, based on the level of aggression of children in the classroom perceived as the most popular, distinguished more developmental trajectories, compared to other measures of norm salience. Altogether, the comparison between different measures of group norms suggests that these might be better captured by looking at the processes by which the group regulates members' behaviour than by the perceptions.

An additional strength of this study lies in the implications that these results have for intervention. First, it seems clear that although the majority of children tend to remain stable over time in their levels of aggression, there are small groups of children who show positive or negative changes across the school-year. Understanding what makes these children different is crucial to unveil the mechanisms through which changes may

occur and to design interventions aimed at transforming those mechanisms. In this case, classroom norms and cohesion seemed to be important features to be taken into consideration to prevent aggression. Second, our results noted that popular children in the classroom have an important role in modeling group norms and influencing others to follow them. Hence, educational practitioners should pay attention to high-status children and either reinforce their positive behaviours or discourage negative behaviours that might be emulated by peers who look up to them. Finally, it appears that classrooms with prevalent problems of disruptive behaviours characterized by physical or verbal aggression may benefit from interventions aimed at improving cohesion among students.

Some potential limitations of this study should also be discussed. On the one hand, from a statistical point of view, there could be some criticism of the fact that trajectories that included a rather small proportion of the sample were considered in the analyses. Although these trajectories seem trivial in the number of children included, they had a relevant theoretical meaning, as they represented high-risk children who were presenting either persistent high levels of aggression or an increase in this behaviour over time. On the other hand, the analyses may have had a limited power to detect classroom level effects due to sample size of the classrooms. Based on the multilevel analyses rule that at least 30 to 40 units in the higher level are required (Bryk & Raudenbush, 1992-2009), we recruited 63 classrooms for the study. However, it is possible that more classrooms would have provided more variability and more statistical power to detect significant effects.

Future directions

The results of this study bring up some interesting questions that may enrich knowledge about classroom effects on children's social development. First, the differential effects of perceived norms and norm salience led us to believe that another refinement to the measurement of norms is to evaluate the perceptions that children have regarding the consequences they might experience in their group as a result of exhibiting certain behaviours. This has been previously explored in studies focused on bullying behaviours (Salmivalli & Voeten, 2004). Expanding this approach to direct and indirect aggression and comparing it with norm salience as related with both types of aggression would be interesting to examine the extent to which children are aware of the salience of their group norms. Also related to norms, it should be explored whether different expectations can be identified for girls and boys in terms of the types of aggression that are tolerated within each gender. Previous studies have already suggested that this might be the case (Velásquez et al., 2010), but this study did not explore this question given our main interest was on the climate of the classroom as a whole.

A second issue of interest raised by this study has to do with the differentiation of two particular patterns of change found in boys: the increasing versus the desisting trajectories. We were able to predict that children in classrooms where the most accepted children are aggressive will tend to follow an increasing pathway of aggression instead of a desisting one. However, it is necessary to better understand what makes some children desist in aggression. This would certainly provide invaluable information for the design of aggression prevention programs.

A third direction in which this study leads us is related to classroom structure development. Given the apparent importance of the classroom features examined here, a

next step would be to understand how that structure is formed, how it changes over time and what factors predict such changes. This would require a closer examination of the transformations that take place during the first months of the school year and of the role both teachers and students play in such a process. In addition, it should be acknowledged that classrooms, as dynamic systems, might also change over the course of the school year. For instance, some of the questions that might be interesting to explore are: How is popularity acquired during the first weeks of school? What makes some classrooms more cohesive than others? To what extent are these group features rooted in the structure of the group in previous years? An examination of when and how those changes occur would also enlighten intervention initiatives that might want to generate positive transformations in children by means of changing the social setting they develop in.

Finally, the study of classroom effects on children's social development may also be enriched by the classroom climate literature. Other features of the classroom environment, not directly related to the peer group, may also deserve some attention in the study of aggression. For example, previously studied dimensions of the classroom such as teacher-student relationships, teachers' leadership style, classroom rules, cooperative vs. competitive academic interactions, and discipline management, among others, might be relevant in the prevention or promotion of aggressive behaviours within the classroom. Hence, further studies should examine the effect of these features on children's changes over the course of the academic year.

In sum, this study contributed to our understanding of the development of aggression in the school context in four different ways. First, this is the first study that looks at heterogeneity in the patterns of change in aggression that children have across the school

years. Second, by comparing different measurements procedures for classroom norms, we were able to identify that one of the most powerful influences on children comes from high status peers who model what is expected in the group. Third, this study showed some evidence that the level of cohesion of the peer group in the classroom context may have positive and negative effects in the development of aggression, in that it may prevent children from harming each other, but it also may enhance conformity to norms that promote aggression. Finally, we provided useful information to widen the alternatives that practitioners may have to design prevention programs aimed at improving the social settings where children and youth develop.

Table 2.1

Descriptives at each time point for type of aggression by sex.

	Boys				Girls			
Variable	*M*	*SD*	*SE*	*ICC*	*M*	*SD*	*SE*	*ICC*
Direct aggression								
Time 1	4.89	4.00	.14	.09	1.82	1.82	.07	.22
Time 2	4.90	4.00	.14	.07	1.68	1.75	.06	.23
Time 3	4.64	3.98	.14	.14	1.72	1.97	.07	.28
Time 4	5.10	4.08	.14	.15	1.93	2.10	.08	.30
Indirect aggression								
Time 1	3.31	2.70	.09	.11	3.29	2.54	.09	.10
Time 2	3.36	2.84	.10	.11	3.41	2.65	.10	.16
Time 3	3.24	2.87	.10	.21	3.08	2.69	.10	.23
Time 4	3.82	3.08	.10	.18	3.60	2.81	.10	.26

Note. N = 842 and 752, for males and females, respectively.

Table 2.2

Correlation coefficients individual level variables

Variable	1	2	3	4	5	6	7	8	9	10
1. Sex	-									
2. SES	-.01	-								
3. Age	-.03	.17*	-							
Direct aggression										
4. Time 1	-.44*	-.09*	.08*	-						
5. Time 2	-.45*	-.09*	.09*	.86*	-					
6. Time 3	-.42*	-.11*	.08*	.83*	.88*	-				
7. Time 4	-.43*	-.10*	.09*	.80*	.83*	.87*	-			
Indirect aggression										
8. Time 1	-.01	-.06*	.08*	.70*	.57*	.56*	.53*	-		
9. Time 2	.00	-.10*	.10*	.57*	.67*	.61*	.57*	.72*	-	
10. Time 3	-.03	-.12*	.08*	.54*	.60*	.72*	.62*	.64*	.76*	-
11. Time 4	-.04	-.10*	.10*	.55*	.59*	.65*	.72*	.65*	.73*	.78*

Note. N = 1594; for the measure of sex, Boys = 0 and Girls = 1.
* p < .05

Table 2.3

Correlation coefficients for classroom level variables

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Cohesion	-												
2. Perceived norm	-.15	-											
Norm salience													
3. Aggression of accepted children	-.18	.38*	-										
4. Aggression of popular children	.15	.46*	.58*	-									
5. Aggression and acceptance *r*	.03	.17	.71*	.48*	-								
6. Aggression and popularity *r*	.39*	.33*	.33*	.76*	.55*	-							
Direct aggression													
7. Time 1	-.25*	.46*	.53*	.40*	.14	-.07	-						
8. Time 2	-.29*	.42*	.30*	.14	.02	-.12	.66*	-					
9. Time 3	-.27*	.44*	.39*	.24	.16	.01	.58*	.89*	-				
10. Time 4	-.30*	.48*	.42*	.31*	.17	.04	.60*	.82*	.85*	-			
Indirect aggression													
11. Time 1	-.18	.59*	.51*	.51*	.13	.08	.84*	.51*	.45*	.43*	-		
12. Time 2	-.15	.40*	.35*	.14	.13	-.01	.47*	.79*	.78*	.61*	.50*	-	
13. Time 3	-.16	.42*	.43*	.32*	.26*	.15	.44*	.72*	.89*	.72*	.41*	.79*	-
14. Time 4	-.22	.47*	.47*	.41*	.19	.14	.52*	.74*	.83*	.88*	.48*	.67*	.78*

Note. N = 63

Linear vs. curvilinear change in types of aggression

Table 2.4

Chi-square difference test for lineal and curvilinear growth curve models in direct and indirect aggression.

Type of model	χ^2	df	Scaling correction for MLR[a]	p	Model comparison	Scaled χ^2 difference	df
Direct aggression							
Unconstrained	16.57	7	6.56	.02			
Linear	14.01	9	7.87	.12	1 vs. 2	.20 n.s.	2
Curvilinear	10.39	7	6.63	.16	2 vs. 3	.39 n.s.	2
Indirect aggression							
Unconstrained	34.99	7	4.01	.00			
Linear	28.73	9	4.98	.00	4 vs. 5	.75 n.s.	2
Curvilinear	19.93	7	4.41	.00	5 vs. 6	1.25 n.s.	2

[a] Given the asymmetric nature of the distribution of aggression scores, robust maximum likelihood estimators were used, which required this scaling correction for chi-square difference tests.

Table 2.5

Classroom effects on aggression trajectories, differentiated by sex and type of aggression.

Models	Girls				Boys			
	Direct		Indirect		Direct		Indirect	
	High / Low	Increasing / Low	High / Low	Increasing / Low	Increasing / Low	Increasing / Desisting	Increasing / Low	Increasing / Desisting
Model 1. Perceived norm								
SES	-.12	-.51	.29	.20	.02	.31	.22	.07
Age	.40	.10	.69*	.17	.60*	.52*	.17	.35
Norm	1.06*	.48	1.04	.68	.88	.10	.95	-1.69
Cohesion	-.13	-.18	-.01	-.05	-.10	-.15	-.10	-.10
Interaction	.42†	-.21	.37*	.08	.11	-.47†	.20	-.29
Model 2. Aggression of most accepted								
SES	-.10	-.48	.25	.23	.12	.56	.18	1.05*
Age	.42	.21	.65*	.20	.67*	.46	.20	-.30
Norm	.24	.07	.23	.22	.25*	.28	.23*	.41*
Cohesion	-.06	-.21*	.05	-.09	-.12	-.22*	-.08	.60
Interaction	.06	-.03	.02	.09	.02	.59	.01	-.40*
Model 3. Aggression of most popular								
SES	-.19	-.57*	.20	.16	-.02	.54*	.11	.37
Age	.43	.22	.68*	.20	.61*	.37	.21	.07
Norm	.14	-.06	.18*	.21*	.25*	.21	.24*	.06
Cohesion	-.06	-.26*	-.01	-.09	-.14*	-.19*	-.09	-.08
Interaction	-.01	-.07	.01	.01	-.02	.17	-.05	-.09
Model 4. Aggression and acceptance r								
SES	-.27	-.51	.16	.14	.01	.37	.13	.55
Age	.52	.20	.63*	.24	.61*	.37	.19	-.39
Norm	.70	.13	.80	1.32	1.82	2.49	1.16	3.84*
Cohesion	-.07	-.22	.01	-.09	-.12	-.20*	-.09	-.27
Interaction	.52	-.22	.28	.58*	.28	-.17	.13	-1.59*
Model 5. Aggression and popularity r								
SES	-.19	-.61*	.12	.04	-.09	.39	.09	.31
Age	.39	.12	.62*	.10	.49*	.24	.12	-.14
Norm	.05	-.52	.59	1.49	1.59	2.21	1.42	1.73
Cohesion	-.06	-.35*	.00	-.09	-.15*	-.26*	-.10	-.15
Interaction	-.27	-1.08	-.13	-.12	-.24	-.55	-.50†	-.80

Note. Estimates for age and SES correspond to odds-ratios. Estimates for Norms, Cohesion and the Interaction are regression unstandardized coefficients.

† p < .10; * p < .05; significance levels correspond to one-tailed tests, except for interactions, where p-values correspond to two-tailed tests.
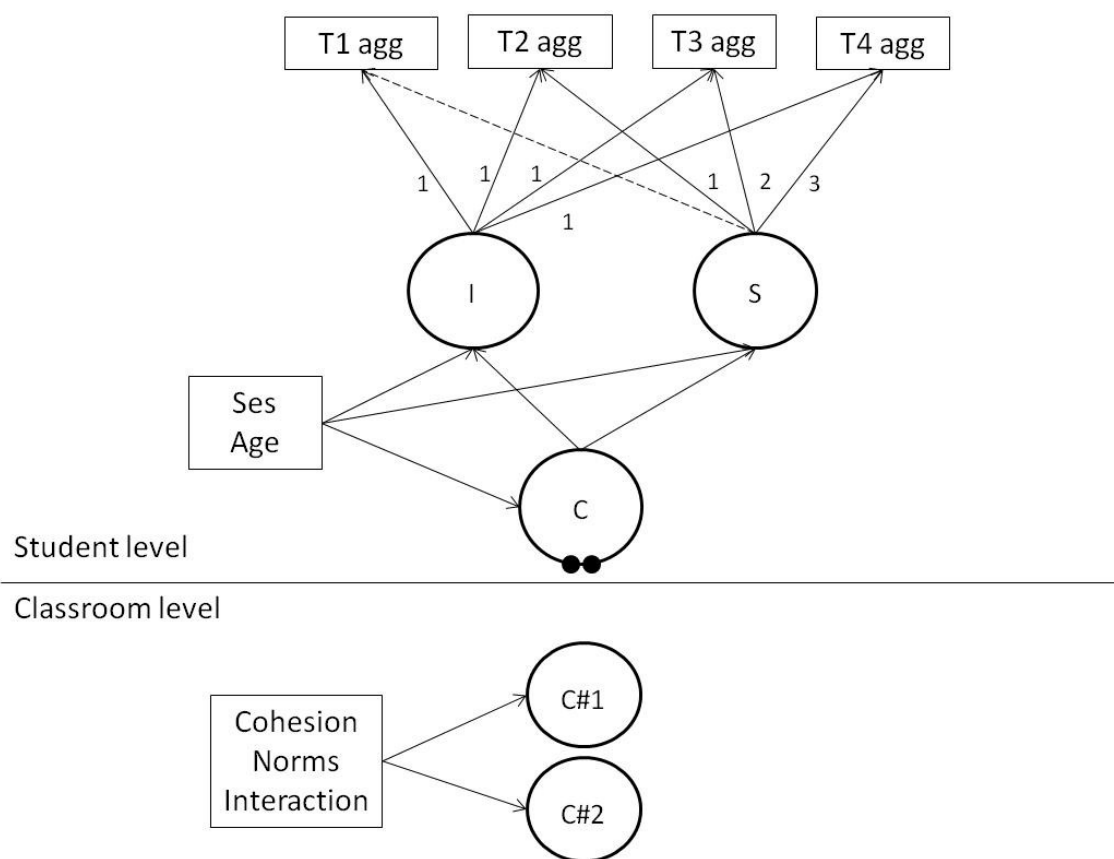
Figure 2.1. Multilevel growth mixture model of aggression development as explained by classroom norms, cohesion and the interaction between both, and controlling for SES and age.
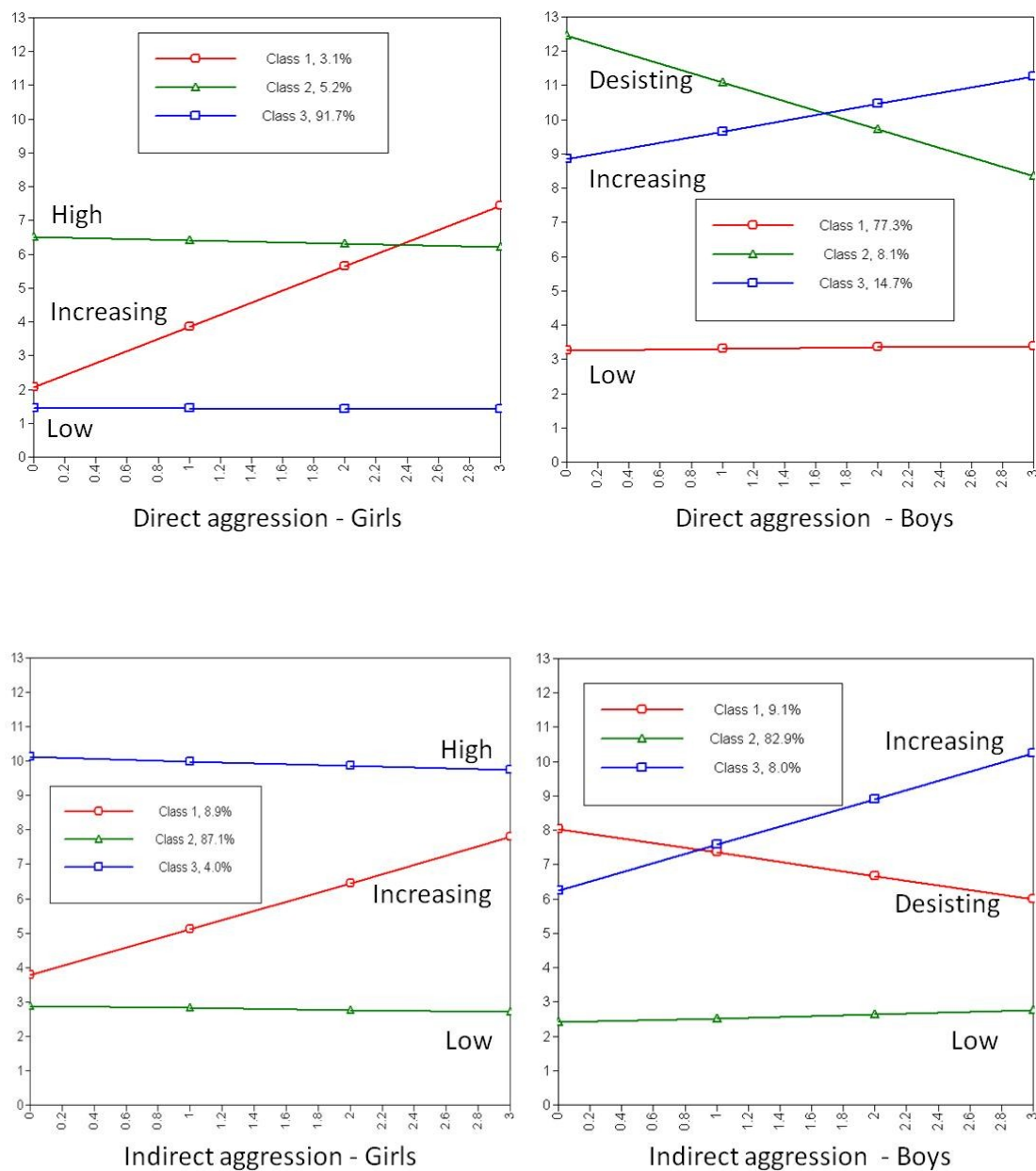
Figure 2.2. Trajectories of change in direct and indirect aggression over the course of the school year for boys and for girls.

Chapter 3. Predicting changes in classroom aggression-related norms: The role of

teachers' normative beliefs and support

Theory and research show that the behaviour of individuals is affected by group norms (Miller & Prentice, 1994). For instance, it has been found that children who develop in classroom contexts where aggression is tolerated tend to become more aggressive themselves (e.g., Velásquez, Bukowski, & Saldarriaga 2010b). Given the power of the classroom context to affect the behavioural outcomes of individual children, there is a need for research to identify the factors that are antecedent to classroom characteristics. Studies have typically approached the classroom context from a very narrow frame, specifically as either as a predictor or as a moderator of individual processes. Only a limited number of studies have conceptualized classrooms as dynamic social systems that have particular characteristics and outcomes that are distinct from mere aggregates of phenomena at the level of the individual student. (e.g., Sinclair & Fraser, 2002). Having a clearer understanding of the factors that affect the emergence of classroom characteristics is a clear pre-requisite for changing school environments to make them more effective contexts for the promotion of healthy development. Based on this, the goal of the present study is to assess changes in the classroom environment over the course of the school year and to identify the influence that teachers have on these changes. The specific goal of the study was to examine the way in which classroom aggression-related norms change over the course of the school year and the role of teachers' beliefs and support in the prediction of those changes.

In order to understand the dynamics of the social structure of the classroom one needs to regard the features of these contexts as outcomes on particular processes,

Sinclair and Fraser (2002) present several ways in which classrooms have been studied as outcomes. One concerns the analysis of differences in the perceptions of the classroom members, for example, between students of different sex (e.g., Fisher, Fraser & Rickards, 1997), between teacher and students (e.g., Fisher & Fraser, 1983) or between real and ideal perceptions (e.g., Sinclair & Fraser, 2002). The other way is by looking at changes in classroom environments as a result of educational interventions (e.g., Batisstich, Schaps, Watson, & Solomon, 1996; Frederickson & Turner 2003; Greenberg, 2007; Kellam et al., 1998; Leadbeater, Hoglund, & Woods, 2003; Teh & Fraser, 1994). In spite of the value of these investigations, a lack of research examining natural dynamics in the classroom. We consider this type of research essential to be able to propose research-based interventions derived from the observations of classroom development.

Although little theory developed specifically to explain classroom processes, the ecological perspective (Bronfenbrenner, 1976, 1992, 2006) provides a general framework that is well suited to studying the classroom as a social system. The central feature of this model is recognition of the importance of the interaction between the individual and his/her surrounding ecologies. In the case of classrooms, two processes need to be acknowledged. On the one hand, classroom members (i.e., students and teachers) bring their own attributes that contribute to form the structure of these social groups. On the other hand, classrooms are embedded in and hence influenced by broader social systems such as schools, neighbourhoods, and communities. Accordingly, the formation and development of classrooms settings needs to be understood as a function of the interaction between its member's characteristics and the processes involved in higher order systems.

Arrow's social-psychological theory on small groups may also shed light as to how classroom settings develop (Arrow et al., 2004). This theory claims that the same basic concepts that are used to study individual change should be used to study group change, namely, change, stability and continuity. Applied to group processes, change is the alteration in the patterns of interaction within the group as a whole; stability is the process that maintains the group in a current state or in equilibrium; and continuity is the experience of change in a group, while maintaining its collective identity. It is our contention that the social system that develops within classrooms is subject to the interaction of forces that drive classrooms to be stable, to be transformed and to maintain their identity while adapting to environment demands.

Another important contribution of this theory is related to its "developmental" perspective on group processes. Specifically it claims that group processes will vary according to the phase of the group's development, that is, whether it is a period of formation, maintenance, or conclusion. According to Arrow et al. (2004), the life of a group is subject to permanent accommodations to internal and external forces that lead to either stability or change. As for the change forces, several considerations need to be taken into account for classroom processes. On the one hand, external forces may create demands resulting in the adaptation of the classroom group to reestablish a state of equilibrium. For example, school organization, activities, rules, and staff turnover, among others, can trigger different processes within the classroom that change its internal dynamics. Therefore, the demands that the school environment exert on classrooms are among the external forces that could explain classrooms changes. On the other hand, there can be internal forces that promote changes through transformations in

the patterns of affiliation within the group. They may also results from the actions of the teacher who, as the main leader of the classroom, might change his or her pattern of interaction with the students, generating transformations in the climate.

Taken together, these ideas suggest that classroom dynamics can be analyzed as a function of several underlying forces that drive them to change. In the present study, we focus on the role of teachers as internal agents of classroom change. Teachers' relationship with their students has previously been studied as a dimension of classroom climate (e.g., Esposito, 1999; Griffith, 1999; Howes, 2000; Wentzel, 1997). Given that the teacher is a member of the classroom group, we assess this relationship as part of the internal structure of the group. We propose that teachers generate transformations in the classroom by means of changing their pattern of interaction with the students. Based on teachers' leadership role, we expect them to have an important impact on the course of development of the classroom structure.

Classroom outcomes can take several forms. Consistent with the peer relations literature (Rubin et al., 2006) and with theories of small groups (Shaw, 1981), we focus on one particular aspect of the classroom peer group structure, namely group norms. Group norms have been defined as the acceptable and/or expected behaviour of the members of a social group (Shaw, 1981). Already there is evidence that classroom norms affect the behaviour of individual children. Previous research has typically examined the effects of descriptive (i.e., the typical behaviour in the group) and injunctive norms (i.e., standards of what is approved or disapproved in a social context) on aggression (e.g., Chang, 2004; Henry et al., 2000). From an ecological point of view (Baker, 1968) there is a need to take a more comprehensive view that includes feedback circuits within the

classroom system indicating which behaviours are appropriate in the group. This approach would consider the positive or negative consequences that the members of a group face in their social system when enacting certain behaviours. In fact, in previous research we found that when norms are made salient in the group (through the exhibition of aggression by high status peers) such norms appear to have a higher predictive value of students' development of aggression compared to other forms of norms (Velasquez & Bukowski, 2010). Accordingly, in this study we will look at changes in aggression-related norm salience.

As role models and authority figures, teachers may have an influence on how the classroom culture is defined. It is presumed that they can influence the meaning and value of the different social behaviours taking place in the classroom. As students perceive teachers expectations, they may adapt their behaviours to comply with them (Birch & Ladd, 1996). Two particular ways in which we hypothesize that teachers might exert an influence in aggression-related normative behaviours of the classroom are by (a) conveying their attitudes and beliefs regarding the acceptability of such behaviours and (b) modeling what they consider to be appropriate social behaviours, for example, social support and warmth.

Previous research has already suggested that teachers' tolerance of deviance influence the value and expectations that are defined in the classroom context (Kedar-Voivodas, 1983). Presumably, teachers' attitudes and beliefs are expressed in their strategies for enforcing or inhibiting certain behaviours in the classroom. This might explain why teachers' attitudes have been found to be associated with students' evaluations of their peers' social attributes and likeability (Hughes, Cavell, & Wilson,

2001). For instance, Chang (2003) showed that in classrooms where teachers disapprove of aggression, aggressive children tended to be less popular among their classmates. These findings suggest that the teacher plays an important role in the normative processes that occur in the classroom. However, these findings are based on concurrent associations between measures, and thus provide insufficient evidence for the potential that teachers might have in transforming the classroom culture over time. For this reason, this study looked at the influence of teachers' aggression-related normative beliefs on changes over time in classroom norm salience (i.e., aggression of popular children in the classroom).

Another way in which teachers may impact the acceptability of aggressive behaviours in the classroom is through their supportive behaviours. Previous research has provided evidence of the positive effects that teachers' warmth and support have on students' outcomes (Skinner & Belmont, 1993; Wentzel, 1994; Wubbels, Mieke, & Hooymayers, 1991). Applied to classroom normative processes, we hypothesize that, by modeling warm and caring behaviours for their students, teachers may convey an expectation for these positive behaviours, which may in turn discourage harmful and disruptive behaviours among classmates. Based on this, we expected that aggression-related norm salience would tend to decrease over time in classroom contexts where teachers modeled supportive behaviours. Although this association has been explored concurrently in a study by Chang (2003), the unique contribution of this study is to understanding the role of the teachers in the dynamics on the classroom social structure, by looking at changes in norm salience over the course of the school year.

In sum, the present study had three goals. First, we wanted to explore the dynamic nature of the classroom structure by examining changes over time in aggression-related norm salience. Although we hypothesized that such transformations exist, we did not have any prior hypotheses regarding the overall direction of those changes. Second, we set out to evaluate the role of teachers' aggression-related normative beliefs at the beginning of the school year as a predictor of the course of development of classroom norm salience across one school year. We expected that the level of aggression of popular children would tend to increase in classrooms where teachers tolerate this type of behaviours. Third, we examined the predictive value of teachers' support on changes in norm salience regarding aggression. We hypothesized that teachers' support would tend to discourage negative behaviours among students and thus be associated with decreasing aggressive behaviours among high status children (i.e., norm salience) in their classrooms.

Method

Participants

The participants of this study were 1595 boys and girls (mean age = 10.2 years; 53% male; 47% female) from fourth, fifth and sixth grade in nine schools in Bogotá, Colombia. Participants were enrolled in mixed-sex private schools drawn from low to high socio-economic status (SES) neighborhoods. Individual SES was assessed via parental reports about their household stratification, which is defined in Colombia by the local government and ranges from 1 (low) to 6 (high). A total of 63 classrooms and their respective home-room teachers took part in the study and classroom size ranged from 11 to 32 participants (mean size = 21.98). Of the participating teachers 14 were male and 49

were female. Their average age was 34.43 years. On average teachers had 10.97 years of teaching experience. In terms of their level of study, 60% had a university-level degree, 37% had completed postgraduate studies[5].

Procedure

School principals or coordinators were invited to consider their institutions for participation in the study. After administrative approval was received, the students in the schools were informed of the objectives and procedures of the study in their classrooms, and letters of information as well as parental consent forms were given to the students to be brought home to their parents. Only the participants who returned a consent form signed by their parents were included in the study. Using this recruitment procedure, the overall participation rate of the study per classroom was 79%. At the end of the study, participants received school supplies as an expression of our gratitude for their participation in the study.

Using a classroom administration format, students completed a multi-section questionnaire during a one-hour in-class session. Data from students were collected at four time points over the school year (which in Colombia runs from February to November), separated by approximately eleven-week intervals (Time 1 = March; Time 2 = May; Time 3 = August; Time 4 = October). For the Time 1 data collection, home-room teachers were also asked to fill out a self-report questionnaire.

Instruments

---

[5] 3% did not report their level of studies

Teachers' normative beliefs. Teachers' approval of aggression was assessed through self-report questions, taken from a questionnaire previously used by Daza and Mejia (2005) in Colombian samples. In this questionnaire the teachers were presented with a set of four hypothetical vignettes that involved aggression among students. They consisted of a situation and the reaction to the situation. The teachers were asked to rate their level of agreement with the reaction. The ratings were made on a five-point Likert scale (1 = total disagreement; 5 = total agreement). An example of one of the situations was: "Someone insulted Maria and for that reason Maria insulted that person back. Do you agree with Maria's behaviour?" These vignettes and questions were originally created in Spanish. The average of these four items was used as the teacher's approval of aggression score (Cronbach's alpha = .82).

Teachers' support. The extent to which students perceived that they could rely on their teachers for help was assessed with a questionnaire adapted from Brand and colleague's school climate questionnaire (Brand, Felner, Shim, Seitsinger & Dumas, 2003). Given that in these schools students were taught by several teachers, they were asked to rate how many of their teachers were supportive. Ratings were made with a five-point Likert scale (1 = none of their teachers; 5 = the majority of their teachers). An example of the items used in this scale was "How many of your teachers go out of their way to help students?" The classroom aggregated score of the perceptions of students across three items was calculated and then averaged across the items to obtain a classroom index of teachers' support (Cronbach's alpha = .77).

Classroom norm salience. An adaptation of Henry and colleagues' (2000) measure of norm salience was used for this study. Based on results from a previous study

(Velásquez et al., 2010b) aggression-related classroom norms were estimated by computing the mean level of aggression of the most popular children within each class. This calculation was based on overall indices of aggression and popularity derived from a peer assessment procedure and a measure of acceptance taken from a sociometric measure.

In the unlimited choice peer assessment procedure (for a description see Rubin et al., 2006) each student was given a list of all the participating children in their class and several characteristics and behaviours. They were asked to indicate which of their participating classmates fit each characteristic or behaviour on the list. Four items assessed direct aggression (e.g, someone who hits or pushes other people; Cronbach's alpha at Time 1= 0.94), two items assessed indirect aggression (e.g., someone who talks bad about others behind their backs to hurt them; Cronbach's alpha at Time 1 = 0.87) and one item measured perceived popularity (i.e., someone who is popular). Each child was given a score on each item indicating how often she/he had been nominated by her/his participating classroom peers. The average of the six aggression items was calculated to obtain an overall score for this construct at each of the four time points. The popularity score was the number of times the child was chosen for the "is popular" item.

Acceptance was assessed by means of a sociometric questionnaire (Bukowski et al., 2000; Newcomb & Bukowski, 1983). Students received a list with the name of each of the participants in their class. Same-sex peers and other-sex peers were grouped in two separate columns of names. Children were asked to indicate which of their peers was their first best friend, their second best friend, their third best friend, and any other peer who they considered to be their friends. They completed this procedure first for same-sex

peers, and then for other-sex peers. The index for acceptance was the number of positive nominations (i.e., being chosen as a friend) a child received from both same-sex and other-sex peers, regardless of the friendship ranking. To calculate indices of norm salience, students with popularity scores one standard deviation above their classroom mean were identified and their overall aggression scores were averaged for each class. Two measures of norm salience were then created for each classroom at each time point: a) the mean level of aggression of the children perceived as most popular, and b) the mean level of aggression of the most accepted children.

Nomination-based scores adjustment for classroom size differences. Considering the difference in the number of nominators among classrooms (range = 9 to 31) scores for aggression, perceived popularity and acceptance were corrected for classroom size (i.e., number of nominators). To maintain the original scale of the items (i.e. number of nominations received), a regression-based procedure was used, as opposed to standardization of the scores (for a review of the procedure see Velásquez et al., 2010a). First, we examined the extent to which changes in the classroom size had an effect on the number of nominations received for each variable. To do this, the linear and quadratic effects of the classroom size minus one (i.e. the person receiving the nominations) were used as predictors of scores at each time point. These scores were aggregated at the classroom level and analyzed with the classroom as the unit of analysis, to account for the nested nature of the data.

Results for acceptance showed that the variance explained by the classroom size was 8% at time 1, 14% at time 2, 22% at time 3 and 17% at time 4. Unstandardized *B* coefficients in the linear effect indicated that for each unit increase in the group size,

students' acceptance score at time 1 through time 4 increased by .141, .156, .195 and .178, respectively. In addition, *B* coefficients for the quadratic effect showed that for each squared unit increase, scores in acceptance decreased in .004, .009, .011, and .010, respectively from time 1 through time 4. Based on this, a standard group of 20 students (which represents the median) was selected as the reference group. For each unit increase in classroom size students' scores were reduced by its correspondent linear units (e.g., .141 for Time 1 scores) and increased in the respective squared units (e.g., $.004^2$ for Time 1 scores). Conversely, for each unit decrease in the classroom size, linear units were added and quadratic units were subtracted from the students' scores. Scores representing zero nominations were left untransformed.

A similar procedure was conducted for the aggression and perceived popularity scores. However, to make variables in the peer assessment questionnaire comparable to each other, an average number of nominations received for the 12 items of this questionnaire was calculated first. This value was included as the dependent variable in the regression analysis. These results indicated that, at each time point (i.e., 1 though 4) classroom size explained 39%, 24%, 20%, and 28% of the variance of peer assessment scores. In this case, the linear unstandardized *B* coefficients from Time 1 through 4 were .159, .125, .127, and 147, respectively, and the quadratic coefficients for each time point were -.006, -.003, -.004, and -.001, respectively. The peer assessment scores were then adjusted according to the steps described for acceptance scores.

Plan of analysis

Given that repeated measures were nested within classrooms, a multilevel analysis was conducted with the software HLM 6.08 (Bryk & Raudenbush, 1992-2009). The

purpose of these analyses was to model initial status and changes in classroom norm salience across the school year, and to test the predictive value of teachers' characteristics on the initial status and rate of change.

First, we created an unconditional model including only classroom norms as the outcome variable (i.e., scores in the repeated measures). Two models were tested separately for norm salience as measured by the level of aggression of highly popular children, and as measured by the level of aggression of highly accepted children. This allowed us to identify the proportion of variance within- and between-classrooms. Next, we added time as a predictor of the outcome, which was coded Time 1 = 0, Time 2 = 1, Time 3 = 2, and Time 4 = 3. This way we could identify the mean intercept (i.e., initial level of norm salience) and mean linear slope (i.e., change), for all the classrooms. Additionally, we tested whether or not these parameters were fixed or random, that is whether there was significant variability in the way classrooms started and in the way they changed over time. When random effects were found, teachers' normative beliefs and support were included as predictors of the intercept and time slope. For these analyses, socio-economic status and grade of the classroom were included as control covariates. All predictors were centered around their grand mean. Chi-square difference test were used to estimate whether a significant proportion of variance of the intercept and slope were explained by the predictors.

Results

Descriptives

Means and standard deviations of the teacher variables and norm salience across the four-time points, along with the correlations among these variables are described in

Table 3.1. Correlation coefficients indicate that measures of norm salience tended to be stable over time. In terms of teachers' variables, results show a positive association between teachers' favorable attitudes toward aggression and levels of norm salience as measured by both indices of popularity. In contrast, teachers' support was only associated with measures of norm salience based on perceived popularity.

Changes in norm salience based on perceived popularity

The unconditional model for norm salience as indicated by the level of aggression of the children perceived as most popular by their classmates showed that 31.9% of the variance was within-classrooms whereas 68.1% of the variance was between-classrooms. A second model that included time as a Level 1 predictor of the outcome showed that time explained 20.6% of the within-classroom variance. Results from this model revealed that time was a positive and significant predictor of the outcome ($B = .17$, $SE = .08$, $p < .05$). That is, there is a general tendency of popular children to become more aggressive over time in these classrooms. Random effects estimations revealed that both the intercept (i.e., initial level of the norm) and slope (i.e., rate of change of the norm) varied significantly across classrooms (intercept variance $= 2.20$, $X^2_{(62)} = 233.98$, $p < .05$; slope variance $= .15$, $X^2_{(62)} = 103.89$, $p < .05$).

Given that there was significant variability to be explained both in the intercept and the slope, teachers' normative beliefs and support were included as Level 2 predictors of the Level 1 parameters. Results from Table 3.2 show that initial levels of classroom norm salience as measured by the level of aggression of children perceived as most popular in the classroom was positively predicted by the teachers' normative beliefs about aggression. Further, the rate of change in the measure of classroom norm salience over

time was negatively predicted by teachers' support ($p < .04$, one-tailed). In other words, in classrooms where teachers tolerated aggression, children who were perceived as popular tended to be more aggressive at Time 1, and in classrooms where most teachers were perceived as supportive by their students, norm salience (i.e., aggression among popular students) tended to decrease over time. Changes in Tau revealed that the inclusion of these Level 2 predictors accounted for 30.3% of the variance of the intercept and 9.7% of the variance of the time slope. Chi-square difference test showed that these proportions of variance explained were significant ($\Delta X^2_{(4)} = 63.88$, $p < .05$, for the intercept; $\Delta X^2_{(4)} = 10.52$, $p < .05$, for the slope).

Changes in norm salience based on acceptance

The unconditional model that included norm salience as measured by the level of aggression of highly accepted children in the classroom showed that 30.4% of the variance was within- and 69.6% of the variance was between-classrooms. Adding time as a Level 1 predictor of the repeated measures in norm salience explained 20.8% of the within-variance. Results from this model indicated that on average, classrooms do not seem to change in this type of norm salience over time ($B = .08$, $SE = .06$, $p > .05$). However, there was significant variation across classrooms in norm salience at the beginning of the year and in the rate of change of this feature across the school year (intercept variance $= 1.90$, $X^2_{(62)} = 247.76$, $p < .05$; slope variance $= .08$, $X^2_{(62)} = 90.22$, $p < .05$).

Next, teachers' normative beliefs and support were added as predictors of the intercept and slope variability in norm salience. The coefficients in Table 3.2 indicate that teachers' normative beliefs positively predicted levels of norm salience at the

beginning of the year. That is, in classrooms where teachers approve of aggression in their students, highly accepted children tend to be more aggressive at the beginning of the year. None of the variables predicted changes in this measure of norm salience across time. Chi-square difference tests showed that a significant amount of variance was accounted for in the intercept ($\Delta X^2_{(4)}$ = 50.08, $p < .05$, for the intercept), while no significant amount of variance was accounted for in the slope ($\Delta X^2_{(4)}$ = 2.63, $p > .05$, for the slope).

Discussion

The goals of the present study were to explore the transformations that elementary school classrooms undergo in their social structure, particularly in aggression-related norms, and to examine the role of teachers' tolerance of aggression and support on predicting such changes. To accomplish these research goals, we assumed an ecological perspective (Arrow et al., 2004; Bronfenbrenner, 1976, 1992, 2006) whereby the dynamic nature of a social context such as the classroom is acknowledged and the characteristics of a particular member of the classroom, namely the teacher, are regarded as internal forces that drive the classroom structure towards change.

In terms of the dynamics of the classroom, our results show that overall changes in norm salience over the course of the school year depend on the type of popularity that is used as a measure of norm salience. While no mean changes were found in the level of aggression of highly accepted children, children perceived as popular increased in their level of aggression across time. It is possible that, at the group level, the perceived popularity measure is better at capturing who in the classroom is being acknowledged as a high-status member and hence, who is having their behaviours reinforced by the group.

In trying to explain these dynamics of the classroom, we found that teachers play an important role by influencing the culture of the classroom. It has been proposed that teachers implicitly or explicitly express their own standards of expected behaviours within the classroom (Kedar-Voivodas, 1983). Accordingly, it has been found that students tend to adapt their behaviours to comply with such expectations (Birsh & Ladd, 1998). Our findings were consistent with these ideas. In relation to aggression-related classroom norms, we proposed two ways in which teachers could reinforce or discourage norms that tolerate aggression. On the one hand, teachers may hold normative beliefs that vary in their level of tolerance towards these behaviours, and students may perceive such beliefs via teachers' treatment of aggressive children in the classroom. In particular, teachers might influence who becomes popular in the class by approving their behaviours, thus enhancing the power that high-status children have in influencing others by defining the group norm. In fact, we found that teachers' aversion to aggression was concurrently negatively related to the level of aggression of high-status children (i.e., norm salience). This was consistent with Chang's (2003) results. However, we were not able to predict changes in norm salience across the school year. This could suggest that teachers influence the classroom norms as they are emerging during the first weeks of school. It is important to note that the first time of measurement in this study occurred approximately 10 weeks after school started, and thus it is possible that teachers' attitudes may have affected norm salience development during a period of time that was no captured by this study.

The other mechanism whereby teachers may shape their classrooms' norm salience is by modeling appropriate behaviours. Accordingly, we assessed the extent to which

students perceived their teachers to be caring, warm and supportive. In line with previous research (Skinner & Belmont, 1993; Wentzel, 1994; Wubbels, Mieke, & Hooymayers, 1991), we found that teachers' support influenced the classroom climate and, by extension, the students in positive ways. Specifically, when teachers were perceived as supportive at the beginning of the school year this was associated with a decrease in classroom norms favoring aggression over the course of the year. This effect was found only for aggression norm salience measured with perceived popularity. These results suggest that, as teachers model positive interactions with their students, they set the standard of what is expected in the classroom in terms of social interactions. Consequently, peers will tend to ascribe popularity to those children who enact these appropriate behaviours, discouraging a norm that tolerates harm among the students. This finding is particularly important for informing interventions, because it illuminates the characteristics of the teachers that may prevent aggression.

Based on these results, three important contributions of the present study can be highlighted. First, this is the first study examining changes in the classroom normative structure from a developmental perspective. Most studies have regarded this context as a static predictor of individual outcomes. Even intervention studies concentrate on students' outcomes and not so much on changes in the classroom structure. When they do, classroom changes are mainly assessed as aggregates of students' behaviours (e.g., Leadbeater & Hoglund, 2003). Therefore, one of the main contributions of this study is to contrast concurrent with longitudinal data and to differentiate associations from predictions of change. Given that intervention efforts are designed to produce change, results from studies like this are more informative concerning the variables that should be

manipulated in experimental designs so as to generate transformations in the classroom settings.

The second contribution of this study was discerning how teachers' cognitions and behaviours may influence classroom normative processes. According to our findings, group development in the classroom appears to be more sensitive to teachers' expression of positive interactions than to their beliefs. It is possible that those normative beliefs may influence teachers' aggressive behaviours toward their students, which in turn shape the classroom culture. Therefore, it would be interesting to explore the effect of punitive or coercive forms of teachers' behaviour in further research.

The third contribution of this study surrounds the assessment of normative processes in the classroom. In situating our study within an ecological framework, we aimed to examine classroom processes that informed, in a dynamic way, which behaviours were expected or accepted. Theory and research in the peer relations literature indicate that status is granted to individuals in the group who exhibit acceptable behaviours (Coie & Kupersmidt, 1983) and that high-status individuals become group leaders who are looked up to by others in the group (Prinstein & Cillessen, 2003). As a consequence, high status individuals may become models for acceptable (normative) behaviours within the group. By assessing aggression in high-status children, we can gain a better sense of the behaviours that students are motivated to emulate in the classroom. Given that high-status can be assessed by means of actual and perceived popularity (Rubin et al., 2006), we included both forms of assessment in our classroom measures. Interestingly, we could only predict changes in classroom norm salience as measured with perceived popularity.

Consistent with previous research (Velásquez et al., 2010b), this might be an indication that perceived popularity is a better index of peer group level processes.

Some limitations of this study should also be noted. First, due to the methodological burden of collecting repeated measures at the level of the classroom, our sample size was relatively small (63 classrooms). For this reason, we may have lacked variability across classrooms in the parameters of interest and may thus have had limited power to detect statistically significant results. Second, assessing teachers' characteristics at the end of elementary school and the beginning of secondary school can be problematic. At this level, students are exposed to a variety of teachers and therefore deciding on who would be the most influential one is not an easy task. In most school contexts, the home-room teacher most frequently deals with behavioural problems in the classroom, which led us to focus on them for the assessment of their normative beliefs. However, in the case of teachers' support we wanted to obtain a general sense of students' perceptions of their teachers, so we asked them about consistency in support across all of their teachers. In terms of the value of these results for intervention purposes, this finding may be challenging to program design, as it implies that all teachers should be trained in supportiveness, which might be more difficult compared to just training home-room teachers. This consideration is important as it calls for more attention to the most appropriate strategies for measuring teachers' characteristics in schools contexts where students are exposed to several teachers.

In addressing the research questions of this study new questions emerge that are worth examining in further studies. The finding that teachers' normative beliefs were only associated with our Time 1 classroom norm salience led us to speculate that these

attitudes might have an effect in earlier weeks. Therefore, a closer examination of the changes that occur in the classroom structure as soon as the students start school might be of great value. In addition, there might be other ways in which teachers can exert influence on children's normative processes. For example, variations in the teachers' classroom management styles, especially those related to discipline or group activities may provide students with more information about what type of interactions are expected or accepted in the classroom. Hence, further examination of teachers' classroom management strategies deserves attention. Finally, we only examined normative processes as they were related to deviant behaviours in the classroom. Another interesting direction would be to examine normative processes regarding positive behaviours such as sociability and prosocial leadership.

In sum, this study contributed to our understanding of the dynamics of the classroom social structure. By using longitudinal data, we were able to provide evidence that these contexts change over time and vary in the way that they change. Our findings provide an initial approximation of how teachers, as main leaders in the classroom, might promote these changes. Specifically, this study suggest that teachers' consistency in the support that they offer to students can help in improving the classroom cultural environment, which in turn may prevent children from following maladaptive paths of development. This new insight into the dynamics of the classroom may help to enhance the effectiveness of research-based prevention programs.

Table 3.1

Means, standard deviations and correlations among the variables

Variable	1	2	3	4	5	6	7	8	9	10
1. T1 teachers' attitudes	-									
2. T1 teachers' support	.01	-								
Norm salience based on perceived popularity										
3. Time 1	.25*	-.27*	-							
4. Time 2	.43*	-.23*	.70*	-						
5. Time 3	.33*	-.21*	.61*	.76*	-					
6. Time 4	.33*	-.32*	.64*	.76*	.75*	-				
Norm salience based on acceptance										
7. Time 1	.28*	-.12	.57*	.52*	.49*	.52*	-			
8. Time 2	.31*	-.15	.43*	.62*	.64*	.55*	.67*	-		
9. Time 3	.28*	-.11	.43*	.58*	.72*	.62*	.60*	.76*	-	
10.Time 4	.30*	-.19	.43*	.56*	.66*	.68*	.64*	.75*	.76*	-
M	1.50	3.77	3.57	3.49	3.52	4.12	2.73	2.77	2.73	3.01
SD	.53	.28	1.91	1.82	2.15	2.47	1.66	1.84	1.88	2.03

Note. Sample size = 63 classrooms/teachers.

* $p < .05$

Table 3.2

Equation Coefficients of the Level 1 intercept and time slope of classroom norm salience

on Level 2 predictors

Variable	Unstandardized coefficient	SE	t	df	p
Norm salience as measured with perceived popularity					
Level 1. Intercept					
Intercept	3.42	.18	18.51	58	.00
Level 2					
Grade	.78	.25	3.16	58	.00
SES	.10	.22	.47	58	.63
Teachers' beliefs	1.25	.34	3.64	58	.00
Teachers' support	-.47	.76	-.62	58	.53
Level 1. Slope					
Intercept	.17	.07	2.32	58	.02
Level 2					
Grade	-.17	.11	-1.57	58	.12
SES	-.09	.08	-1.13	58	.26
Teachers' beliefs	.14	.12	1.12	58	.26
Teachers' support	-.54	.30	-1.79	58	.07
Norm salience as measured with acceptance					
Level 1. Intercept					
Intercept	2.69	.18	15.11	58	.00
Level 2					
Grade	.48	.25	1.92	58	.06
SES	-.40	.19	-2.14	58	.04
Teachers' beliefs	1.10	.31	3.52	58	.00
Teachers' support	-.10	.67	-.16	58	.87
Level 1. Slope					
Intercept	.08	.06	1.31	58	.19
Level 2					
Grade	-.07	.11	-.61	58	.55

Variable	Unstandardized coefficient	*SE*	*t*	*df*	*p*
SES	-.03	.06	-.47	58	.64
Teachers' beliefs	.07	.13	.50	58	.62
Teachers' support	-.27	.31	-.87	58	.39

General discussion

The present thesis provided evidence that the classroom setting represents a context of major importance for the development of a fundamental form of behavior, specifically aggression. Given the well-documented significance that peer relations have on children, the classroom should be regarded as a fundamental context for children's socialization.

Compared with other approaches that have regarded classroom characteristics as aggregates of individual behaviors, we see the classroom as an entity with its own properties that transcend those of its members. Specifically, we were concerned with the features of the classroom that can be observed from the patterns of interactions and transactions that develop among peers. This framework served as our point of departure to define the design of the studies that comprise this thesis in two ways. First, taking into account the social processes that develop in the classroom, we were able to propose a dynamic concept of normative processes and to evaluate the feedback circuits that take place within the classroom social system, in order to the behaviors that are tolerated in this type of peer group. In addition, this systemic view motivated our interest in examining at the classroom as a dynamic system that changes as a function of internal and external forces.

The main purpose of this research project was to take a comprehensive approach to understanding the role of the classroom context in child development. In particular, we had three inter-related goals. The first was to more successfully capitalize on the potential that the classroom has as a context for measurement. The second was to understand the effect that the classroom peer group environment had on children's social

development. The third was to explore and predict the dynamic changes of classrooms over the course of their lifetimes. These goals were accomplished by means of three separate studies.

Study 1. The classroom as a context of measurement

The purpose of this first study was to establish a basis for assessing individual and classroom characteristics in reliable ways, while capitalizing on the richness of the information available from peer evaluations. We were particularly interested in analyzing how peer-nomination-based measures can be used even in the presence of differences in the size of classrooms (number of nominators). Although several techniques have been developed to deal with this issue, important limitations were outlined. In particular we demonstrated that standardization techniques yield scores that are difficult to interpret given that they are relative to the overall functioning of individual classrooms, and thus wiping out the differences between classrooms and limiting the possibility to explore classroom context effects by wiping out the differences between classrooms. Also, we showed that although some researchers interested in maintaining between-groups differences have resorted to use proportion scores, our results invalidate a basic assumption of this technique: that the effect of classroom size is perfectly linear. Based on this, we presented a regression-based technique that overcomes these limitations.

Our results confirmed our expectation that the size of the nomination group is positively associated with sociometric and peer assessment scores. However, we were also able to show that such an effect is not constant across all classroom size differences. A non-linear effect indicated that as the size of the classroom increases, the linear effect

is reduced. Among all techniques available, the regression-based technique presented seems to be the best approach to control for these curvilinear effects.

In addition to this, we were able to present two strong advantages of the regression-based procedure. First, because the original metric of the scale is maintained (i.e., number of nominations received), the interpretation of any results derived from such scores is facilitated. Second, in contrast to standardization techniques, scores adjusted with this type of procedure allow for the estimation of group differences. Therefore, we were able to develop a powerful, efficient and practical tool capitalize on the advantages of classroom-based peer evaluations. In the context of this thesis, this methodological development was fundamental to establish the basis upon which individual and classroom measures for the subsequent studies were created.

Study 2. The classroom as a predictor of individual development.

The second of this set of studies was designed to test the effects of the classroom social structure (i.e., aggression norms and cohesion) on the development of students' aggressive behaviors. We framed our research questions on an ecological perspective that took into consideration the dynamic nature of the classroom structure to evaluate norms. In particular, we took the concept of social systems feedback circuits proposed by Baker (1968) as a frame of reference to assess normative processes in the classroom, specifically, norm salience.

Three results from this study deserve particular attention. First, compared to injunctive norms, aggression-related norms made salient by high-status children in the classroom were better at differentiating children who remained low in their levels of aggression from those who tended to increase their aggression as the school year passed

by. Second, these results highlight high-status children's power to establish classroom norms and influence others in their peer group. In terms of prevention programs guidelines, this result is important because it means that in making strategic decisions about the ideal targets for intervention, more attention should needs to be paid to high-status children. Third, our results call for a further examination of the significance of the group as a reference for conformity to norms. We examined this effect by testing the interaction between group norms and the level of cohesion of the classroom. Our findings suggest that this structural aspect of the classroom can have both positive and negative effects on aggression. This calls for a further exploration of the classroom context variables that might represent the extent to which the classroom is meaningful for students.

For the purposes of this thesis, this study also shows that norm salience, and in particular, changes throughout the school year, is fundamental to the study of classroom processes. A better understanding of this process may provide guidelines as to how positive changes can be generated in the classroom context. The next study addressed this issue.

Study 3. The classroom as a dynamic system.

When examining at normative processes from a systemic point of view, it is assumed that the system is subject to permanent forces that drive it towards change and stability. This study assessed how norm salience regarding aggression changed throughout the school year, and whether differences in change between classrooms changed could be explained by teachers' normative beliefs and students' perceptions of teachers' support.

To our knowledge, this study is the first nonintervention project to assess changes in the classroom environment. We were able to show that classrooms characteristics, measured as the patterns of interactions among students, have dynamic transformations during the life of the classroom group. We were particularly interested in determining why some classrooms increased over time in their level of tolerance towards aggression. We found that increases could be explained by students' perceptions of the teachers as unsupportive.

Following the same line of reasoning that has been applied to interventions targeted at the individual level, these results can inform interventions that aim to change the classroom context. In individual development research, for example, it has been found that social information processing predicts changes in children's aggressive behaviors (Weiss, Dodge, Bates & Pettit, 1992). As a result, intervention programs have included components aimed at modifying children's social cognitions to prevent aggression. A similar logic can and should be applied to studies at the level of the classroom. By understanding changes in classroom's structure and processes, we will be able to better identify the factors that may promote positive transformations in these settings, and therefore translate this valuable information into educational practices.

Summary

The three studies that comprise this thesis present a series of inter-related findings that contribute to the building of knowledge regarding the significance and role of classroom context for child development. We highlighted the value of the classroom as a context of measurement, as a context that can be transformed in positive directions.

It is our expectation that what we have learned from this project will serve as an inspiration to future research examining the role of the classroom context in child and youth development. Of special importance for us is to provide insight into how schools and teachers in particular can guide their educational practices to provide better environments for their students. As developmental psychologists this is our hope and main goal.

References

Allport, F. H. (1924a). *Social Psychology.* Boston : Houghton-Mifflin.

Allport, F. H. (1924b). The group fallacy in relation to social science. *Journal of Abnormal and Social Psychology, 19,* 60-73.

Arrow, H., Poole, M. S., Henry, K. B., Wheelan, S., & Moreland, R. (2004). Time, change, and development: The temporal perspective on groups. *Small Group Research, 35,* 73-105.

Asch, S. E. (1987). *Social psychology*. Oxford, England: Oxford University Press.

Baker, R. G. (1968). *Ecological psychology: Concepts and methods for studying the environment of human behaviour*. Stanford, CA: Stanford University Press.

Bandura, A. *(*1986)*. Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A., & Walters, R.H. (1963). *Social learning and personality development*. New York: Holt, Rinehart, & Winston.

Barth, J. M., Dunlap, S. T., Dane, H., Lochman, J. E., & Wells, K. C. (2004). Classroom environment influences on aggression, peer relations, and academic focus. *Journal of School Psychology, 42,* 115-133.

Bass, B. M. (1960). *Leadership, psychology, and organizational behavior*. Oxford, England: Harper.

Batisstich, V., Schaps, E., Watson, M., & Solomon, D. (1996). Prevention effects of the child development project: Early findings from an ongoing multisite demonstration trial. *Journal of Adolescent Research, 11,* 12-35.

Birsh, S. H., & Ladd, G. W. (1998). Interpersonal relationships in the school environment and children's early school adjustment: The role of teachers and peers. In J. Juvonen & K. R. Wentzel (Eds.), *Social motivation: Understanding children's school adjustment* (pp. 199-225). New York: Cambridge University Press.

Boivin, M., Dodge. K. A., & Coie, J. D. (1995). Individual-group behavioural similarity and peer status in experimental play groups of boys: The social misfit revised. *Journal of Personality and Social Psychology, 69*, 269-279.

Bollen K. A. and Curran P. J. (2006). *Latent curve models: a structural equation perspective*. John Wiley & Sons, NJ.

Brame, B., Nagin, D., & Tremblay, R. (2001). Developmental trajectories of physical aggression from school entry to late adolescence. *Journal of Child Psychology and Psychiatry, 42*, 503-512.

Brand, S., Felner, R., Shim, M., Seitsinger, A., & Dumas, T. (2003). Middle school improvement and reform: development and validation of a school-level assessment of climate, cultural pluralism, and school safety. *Journal of Educational Psychology, 95,* 570-588.

Bronfenbrenner, U. (1943). A constant frame of reference for sociometric research: Part I. Theory and technique. *Sociometry,6,* 363-397.

Bronfenbrenner, U. (1944). A constant frame of reference for sociometric research: Part II. Experiment and inference. *Sociometry,7,* 40-75.

Bronfenbrenner, U. (1976). The ecology of human development: History and perspectives. *Psychologia Wychowawcza, 19*, 537-549.

Bronfenbrenner, U. (1992). In Vasta R. (Ed.), *Ecological systems theory*. London, England: Jessica Kingsley Publishers.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology* (Vol. 1, pp. 793-828). New York: John Wiley & Sons.

Bryk, A. S., & Raudenbush, S. W. (1992-2009). *Hierarchical linear models: Applications and data analysis methods*. Sage Publications, Inc.

Bukowski, W. M., Brendgen, M., & Vitaro, F. (2007). Peers and socialization: Effects on externalizing and internalizing problems. In J. E. Grusec, & P. D. Hastings (Eds.), *Handbook of socialization: Theory and research* (pp. 355-381). New York: Guilford Press.

Bukowski, W. M., & Hoza, B. (1989). Popularity and friendship: Issues in theory, measurement, and outcome. In T.J. Berndt & G.W. Ladd (Eds.), *Peer relations in child development* (pp. 15-45). New York: Wiley.

Bukowski, W. M., & Sippola, L. K. (2001). Groups, individuals, and victimization: A view of the peer system. In J. Juvonen & S. Graham (Eds.), *Peer harassment in school: The plight of the vulnerable and victimized* (pp. 355-377). New York: Guilford.

Bukowski, W.M., Sippola, L., Hoza, B., & Newcomb, A.F., (2000). Pages from a sociometric notebook: An analysis of nomination and rating scale measures of acceptance, rejection, and social preference. In A. Cillessen & W. Bukowski (Eds.), *Recent advances in the measurement of acceptance and rejection in the peer system* (pp. 11-26). San Francisco, CA: Jossey-Bass.

Cartwright, D. (1968). The nature of group cohesiveness. In D. Cartwright & A. Zander (Eds.), *Group dynamics: Research and theory* (3rd. ed., pp. 91-109). London: Tavistock.

Cartwright, D., & Zander, A. (1953). *Group dynamics: Research and theory.* Evanston, IL: Row, Peterson.

Chang, L. (2003). Variable effects of children's aggression, social withdrawal, and prosocial leadership as functions of teacher beliefs and behaviours. *Child Development, 74,* 535-548.

Chang, L. (2004). The role of classroom norms in contextualizing the relations of children's social behaviours to peer acceptance. *Developmental Psychology, 40,* 691–702.

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology, 58*, 1015-1026.

Cillessen, A. (2009). Sociometric methods. *Handbook of peer interactions, relationships, and groups* (pp. 82-99). New York: Guilford Press.

Coie, J. D., Dodge, K. A., & Coppotelli, H. (1982). Dimensions and types of social status: A cross-age perspective. *Developmental Psychology, 18*, 557-570.

Coie, J. D., & Kupersmidt, J. (1983). A behavioural analysis of emerging social status in boys' groups. *Child Development, 54*, 1400-1416.

Cotterell, J. (1996). *Social networks and social influences in adolescence*. Florence, KY: Taylor & Frances/Routledge.

Crick, N. R., Wellman, N. E., Casas, J. F., O'Brien, M. A., Nelson, D. A., & Grotpeter, J. K., et al. (1999). Childhood aggression and gender: A new look at an old problem. In D. Bernstein (Ed.), *Nebraska Symposium on Motivation* (pp. 75 − 140). Lincoln: University of Nebraska Press.

Crosbie, P. V. (1975). *Interaction in small groups*. London & New York: MacMillan.

Daza, B.C., & Mejía, J.F. (2005). *Prueba SABER de Competencias Ciudadanas*. Bogotá: ICFES.

Durkheim, E. (1895). *Les régles de la méthode sociologique.* Paris : F. Alcan.

Esposito, C. (1999). Learning in urban blight: School climate and its effects on the school performance of urban, minority, low-income children. *School Psychology Review, 28,* 365-377.

Farrell, A., Sullivan, T., Esposito, L., Meyer, A., & Valois, R. (2005). A latent growth curve analysis of the structure of aggression, drug use, and delinquent behaviours and their interrelations over time in urban and rural adolescents. *Journal of Research on Adolescence*, *15*, 179-204.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behaviour: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Fisher, D. L., & Fraser, B. J. (1983). A comparison of actual and preferred classroom environments as perceived by science teachers and students. *Journal of Research in Science Teaching, 20,* 55-61.

Fisher, D. L., Fraser, B. J., & Rickards, T. (1997). *Gender and cultural differences in teacher-student interpersonal behavior.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Fraser, B. J. (1994). Research on classroom and school climate. In D. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 493-541). New York: Macmillan.

Frederickson, N., & Turner, J. (2003). Utilizing the classroom peer group to address children's social needs: An evaluation of the Circle of Friends intervention approach. *The Journal of Special Education, 36,* 234-245.

Gazelle, H. (2006). Class climate moderates peer relations and emotional adjustment in children with an early history of anxious solitude: A child × environment model. *Developmental Psychology, 42*, 1179-1192.

Gest, S. D., Graham-Bermann, S. A., & Hartup, W. W. (2001). Peer experience: Common and unique features of friendship, network centrality, and sociometric status. *Social development, 10,* 23-40.

Ghaith, G. M., & *Bouzeineddine*, A. R. (2003). Relationship between reading attitudes, achievement, and learners' perceptions of their jigsaw II cooperative learning experience. *Reading Psychology, 24*, 105-121.

Greenberg, M. T. (2007). *Current and future challenges in school-based prevention.* Symposium conducted at the meeting of the Society for Research in Child Development conference, Boston, U.S.

Griffith, J. (1999). School climate as "social order" and "social action": A multi-level analysis of public elementary school student perceptions. *Social Psychology of Education, 2,* 339–369.

Hare, A. (1976). *Handbook of small group research* (2nd ed.). New York: Free Press.

Henry, D. (2008). Changing classroom social settings through attention to norms. In M. Shinn & H. Yoshikawa (Eds.), *Toward positive youth development* (pp. 40-57). New York: Oxford University Press.

Henry, D., Guerra, N., Huesmann, R., Tolan, P., VanAcker, R., & Eron, L. (2000). Normative influences on aggression in urban elementary school classrooms. *American Journal of Community Psychology, 28,* 59-81.

Hoglund, W. L., & Leadbeater, B. J. (2004). The effects of family, school, and classroom ecologies on changes in children's social competence and emotional and behavioral problems in first grade. *Developmental Psychology, 40,* 533-544.

Honaker, J., King, G., & Blackwell, M. (2009). *Amelia II. A Program for Missing Data*. Retrieved in August 2009: http://gking.harvard.edu/amelia/docs/amelia.pdf

Howes, C. (2000). Socio-emotional classroom climate in child care, child-teacher relationships and children's second grade peer relations. *Social Development, 9,* 191-204.

Huesmann, R., & Guerra, N. (1997). Children's normative beliefs about aggression and aggressive behaviour. *Journal of Personality and Social Psychology, 72,* 408–419.

Hughes, J. N.. Cavell, T. A., & Wilson, V. (2001). Further support for the developmentl significance of the quality of the teacher-student relationship. *Journal of school psychology, 39,* 289-301.

Juvonen, J., & Galván, A. (2008). Peer influence in involuntary social groups: Lessons from research on bullying. In M. J., Prinstein, & K. A. Dodge (Eds.), *Understanding peer influence in children and adolescents* (pp. 225-244). New York: Guilford Press

Juvonen, J., & Ho., A. (in press). The motivating "pull" of bullies increasing antisocial behaviours across middle school grades. *Journal of Youth and Adolescence.*

Kedar-Voivodas, G. (1983). Impact of elementary children's school roles and sex roles on teacher attitudes: An interactional analysis. *Review of educational research, 53,* 415-437.

Kellam, S. G., Ling, X., Mersica, R., Brown, C. H., & Ialongo, N. (1998). The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behaviour into middle school. *Development and Psychopathology, 10,* 165–185.

Kelly, J., Ryan, A. M., Altman, B. E., & Stelzner, S. (2000). Understanding and changing social systems: An ecological view. In J. Rappaport & E. Seidman (Eds.). *Handbook of community psychology.* Dordrecht, Netherlands: Kluwer Academic Publishers.

Newcomb, A. F., & Bukowski, W. M. (1983). Social impact and social preference as determinants of children's peer group status. *Developmental Psychology, 19*, 856-867.

La Fontana, k. M., & Cillessen, A. H. N. (1998). The nature of children's stereotypes of popularity. *Social development, 7,* 301-320.

Leadbeater, B., Hoglund, W., & Woods, T. (2003). Changing contexts? The effects of a primary prevention program on classroom levels of peer relational and physical victimization. *Journal of Community Psychology, 31*, 397-418.

Leaper, C. (2000). The social construction and socialization of gender during development. In P. H. Miller & E. K. Scholnick (Eds.), *Toward a feminist developmental psychology* (pp. 126 – 152). New York: Routledge.

Levine J. M. (1983). Social comparison and education. In Levine, J. M. & Wang, M. C. (Eds.), *Teacher and student perception: Implications for learning.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Levine, J. M., & Moreland, R. L. (1990). Progress in small group research. *Annual Review of Psychology, 40,* 585-634.

Lewin, K. (1936). *Principles of topological psychology.* New York: McGraw-Hill.

Little, T. D., Jones, S. M., Henrich, C. C. & Hawley, P. H. (2003). Disentangling the 'whys' from the "whats" of aggressive behaviour. *International Journal of Behavioural Development, 27,* 122-133.

Lott, B. E. (1961). Group cohesiveness: A learning phenomenon. *Journal of Social Psychology, 55,* 275-286.

Lüdtke, O., Trautwein, U., Kunter, M., & Baumert, J. (2006). Reliability and agreement of student ratings of the classroom environment: A reanalysis of TIMSS data. *Learning Environments Research*, *9*, 215-230.

Martens, M., Page, J., Mowry, E., Damann, K., Taylor, K., & Cimini, M. (2006). Differences between actual and perceived student norms: An examination of alcohol use, drug use, and sexual behaviour. *Journal of American College Health*, *54*, 295 -300.

Martino, S., Ellickson, P., Klein, D., McCaffrey, D., & Edelen, M. (2008). Multiple trajectories of physical aggression among adolescent boys and girls. *Aggressive Behaviour*, *34*, 61-75.

Masten, A. S., Morison, P., & Pellegrini, D. S. (1985). A Revised Class Play method of peer assessment. *Developmental Psychology, 3*, 523-533.

McGrath, J., Arrow, H., & Berdahl, J. (2000). The study of groups: Past, present, and future. *Personality and Social Psychology Review, 4,* 95-105.

McKnight, P., McKnight, K., Sidani, S., & Figueredo, A. (2007). *Missing data: A gentle introduction*. New York: Guilford Press.

Mercer, S., McMillen, J., & DeRosier, M. (2009). Predicting change in children's aggression and victimization using classroom-level descriptive norms of aggression and pro-social behaviour. *Journal of School Psychology*, *47*, 267-289.

Merton, R. K. (1968). *Social theory and social structure.* (2nd ed.). Glencoe, IL: Free Press.

Miller, D. T., & Prentice, D. (1994). Collective errors and errors about the collective. *Personality and Social Psychology Bulletin, 20,* 541-550.

Murray-Close, D., Ostrov, J., & Crick, N. (2007). A short-term longitudinal study of growth of relational aggression during middle childhood: Associations with gender, friendship intimacy, and internalizing problems. *Development and Psychopathology*, *19*, 187-203.

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In Kaplan, D. (Ed.), *Handbook of Quantitative*

Methodology for the Social Sciences (pp. 345–368). Newbury Park, CA: Sage Publications.

Muthén, L. K. & Muthén, B. O. (1998-2008). *Mplus User's Guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.

Nagin, D., & Tremblay, R. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child Development*, *70*, 1181-1196.

Nagin, D., & Tremblay, R. (2005). Developmental trajectory groups: Fact or a useful statistical fiction? *Criminology: An Interdisciplinary Journal*, *43*, 873-904.

Newcomb, T. M. (1960). Varieties of interpersonal attraction. In D. Cartwright, & A. Zander (Eds.), *Group dynamics: Research and theory* (2nd ed., pp. 104-119). Evanston, IL: Row, Peterson.

Newcomb, A., & Bukowski, W. (1983). Social impact and social preference as determinants of children's peer group status. *Developmental Psychology*, *19*, 856-867.

O'Reilly, C.A., & Caldwell, D.F. (1985). The impact of normative social influence and cohesiveness on task perception and attitudes: A social information processing approach. *Journal of Occupational Psychology, 58,* 193-206.

Opp, K. (1982). The evolutionary emergence of norms. *British Journal of Social Psychology, 21*, 139-149.

Ostrov, J., Ries, E., Stauffacher, K., Godleski, S., & Mullins, A. (2008). Relational aggression, physical aggression and deception during early childhood: A multimethod, multi-informant short-term longitudinal study. *Journal of Clinical Child and Adolescent Psychology*, *37*, 664-675.

Ozer, E. J.,Weinstein, R. S., Maslach, C., & Siegel, D. (1997). Adolescent AIDS prevention in context: The impact of peer educator qualities and classroom environments on intervention efficacy. *American Journal of Community Psychology, 25,* 289–323.

Parker, J. G., Saxon, J., Asher, S. R., & Kovacs, D. (1999). Dimensions of children's friendship adjustment: Implications for studying loneliness. In K. J. Rotenberg & S. Hymel (Eds.), *Loneliness in childhood and adolescence* (pp. 201-224). New York: Cambridge University Press.

Parkhurst, J. T., & Hopmeyer, A. (1998). Sociometric popularity and perceived popularity: Two distinct dimensions of peer status. *Journal of early adolescence, 18,* 125-144.

Pettigrew, T. F. (1967). Social evaluation theory. In D. Levine (Ed.), *Nebraska symposium on motivation*. Lincoln, NE: University of Nebraska Press.

Prinstein, M. J., & Dodge, K. A. (2008). Current issues in peer influence research. In M. J., Prinstein, & K. A. Dodge (Eds.). *Understanding peer influence in children and adolescents* (pp. 3-13). New York: Guilford Press.

Prinstein, M.J., & Cillessen, A. (2003). Forms and functions of adolescent peer aggression associated with high levels of peer status. *Merrill-Palmer Quarterly, 49,* 310-342.

Ramsey, P. G. (1995). Changing social dynamics in early childhood classrooms. *Child development, 66,* 764-773.

Richer, S. (1976). Reference-group theory and ability grouping: A convergence of sociological theory and educational research. *Sociology of Education, 49,* 65-71.

Rimm-Kaufman, S. E., La Paro, K. M., Downer, J. T., & Pianta, R. C. (2005). The contribution of classroom setting and quality of instruction to children's behavior in kindergarten classrooms. *The Elementary School Journal, 105,* 377-394.

Rubin, D. B. (1987). Multiple imputation for nonresponse in surveys. New York: Wiley.

Rubin, K.H., Bukowski, W., & Parker, J.G. (2006). Peer interactions, relationships, and groups. In N. Eisenberg (Ed.), W. Damon (Series Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 571-645). New York: Wiley.

Salmivalli, C., & Voeten, M. (2004). Connections between attitudes, group norms, and behaviour in bullying situations. *International Journal of Behavioural Development, 28,* 246-258.

Schaeffer, C., Petras, H., Ialongo, N., Poduska, J., & Kellam, S. (2003). Modeling growth in boys' aggressive behaviour across elementary school: Links to later criminal involvement, conduct disorder, and antisocial personality disorder. *Developmental Psychology*, *39,* 1020-1035.

Schutz, W. C. (1966). *The interpersonal underworld.* Palo Alto, CA: Science and Behavior Books.

Shaver, K. G. (1977). *Principles of social psychology.* Cambridge, MA: Winthrop Publishing Co.

Shaw, D., Gilliom, M., Ingoldsby, E., & Nagin, D. (2003). Trajectories leading to school-age conduct problems. *Developmental Psychology*, *39*, 189-200.

Shaw, M. (1974). *An overview of small group behaviour*. Morristown, NJ: General learning press.

Shaw, M. (1981). *Group dynamics* (3rd ed.). New York: McGraw-Hill.

Sinclair, B. B., & Fraser, B. J. (2002). Changing classroom environments in urban middle schools. *Learning Environments Research, 5*, 301-328.

Skinner, E., & Belmont, M. (1993). Motivation in the classroom: Reciprocal effects of teacher behaviour and student engagement across the school year. *Journal of Educational Psychology*, *85*, 571-581.

Stormshak, E. A., Bierman, K. L., Bruschi, C., Dodge, K. A., Coie, J. D., & the Conduct Problems Prevention Research Group (1999). The relation between behaviour problems and peer preference in different classroom contexts. *Child Development, 70,* 169–182.

Teh, G., & Fraser, B. J. (1994). An evaluation of computer-assisted learning in terms of achievement, attitudes and classroom environment. *Evaluation and Research in Education, 8,* 147–161.

Tesser, A. (1988). In Berkowitz L. (Ed.), *Toward a self-evaluation maintenance model of social behavior*. San Diego, CA, US: Academic Press.

Thomas, D. E., Bierman, K. L., & The Conduct Problems Prevention Research Group (2006). The impact of classroom aggression on the development of aggressive behavior problems in children. *Development and Psychopathology, 18*, 471–487.

Tseng, V., & Seidman, E. (2007). A systems framework for understanding social settings. *American Journal of Community Psychology, 39,* 217–228

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327-352.

Underwood, M., Beron, K., & Rosen, L. (2009). Continuity and change in social and physical aggression from middle childhood through early adolescence. *Aggressive Behaviour*, *35*, 357-375.

Velásquez, A. M., Bukowski, W. M., & Saldarriaga, L. M. (2010a). *Adjusting for group size effects in peer nomination data*. Manuscript in preparation for publication.

Velásquez, A. M., Bukowski, W. M., & Saldarriaga, L. M. (2010b). The effects of classroom peer norms and cohesion on aggression trajectories: A short-term longitudinal study. Manuscript in preparation for publication.

Velásquez, A. M., Santo, J. B., Saldarriaga, L. M., & Bukowski, W. M. (2010). Context-Dependent Peer Victimization: Are physical and relational aggression tolerated differently in mixed-sex versus all-girl schools? *Merril Palmer Quarterly, 56,* 283-302.

Vaillancourt, T., Miller, J. T., Fagbemi J, Cote, S., Tremblay, R.E. (2007). Trajectories and predictors of indirect aggression: Results from a nationally representative sample of Canadian children aged 2–10. *Aggressive Behaviour, 33,* 314–326.

Vitaro, F., Boivin, M., & Bukowski, W. (2009). The role of friendship in child and adolescent psychosocial development. *Handbook of peer interactions, relationships, and groups* (pp. 568-585). New York: Guilford Press.

Weinstein, C. S. (1991). The classroom as a social context for learning. *Annual Review of Psychology, 42*, 493-525.

Weiss, B., Dodge, K. A., Bates, J. E. & Pettit, G. S. (1992). Some consequences of early harsh discipline: child aggression and a maladaptive social information processing style. *Child Development, 63*, 1321-1335

Wentzel, K. (1994). Relations of social goal pursuit to social acceptance, classroom behaviour, and perceived social support. *Journal of Educational Psychology, 86*, 173-182.

Wentzel, K. R. (1997). Student motivation in middle school: The role of pedagogical caring. *Journal of Educational Psychology, 89,* 411-419.

Williams, S., Conger, K., & Blozis, S. (2007). The development of interpersonal aggression during adolescence: The importance of parents, siblings, and family economics. *Child Development, 78*, 1526-1542.

Willms, J. D., & Somers, M. (2001). Family, classroom, and school effects on children's educational outcomes in Latinamerica. *School Effectiveness and School Improvement, 12*, 409-445.

Wright, J. C., Giammarino, M., & Parad, H. W. (1986). Social status in small groups: individual-group similarity and the social misfit. *Journal of Personality and Social Psychology, 50*, 523-536.

Wubbels, T., Mieke, B., & Hooymayers, H. (1991). Interpersonal teacher behaviour in the classroom. In B. J. Fraser & H. J. Walberg (Eds.), *Educational environments: Evaluation, antecedents and consequences* (pp. 141-160). Oxford, England: Pergamon Press.