

REYIHANGULI ALIMUJIANG

PRESCHOOLERS' UNDERSTANDING OF CULPABILITY

Bilkent University 2020

THE DEVELOPMENT OF PRESCHOOLERS'
UNDERSTANDING OF CULPABILITY

A Master's Thesis

by

REYIHANGULI ALIMUJIANG

Department of Philosophy

İhsan Doğramacı Bilkent University

Ankara

June 2020

To the silent dedication of my family

THE DEVELOPMENT OF PRESCHOOLERS' UNDERSTANDING
OF CULPABILITY

The Graduate School of Economics and Social Sciences
of
İhsan Doğramacı Bilkent University

by

REYIHANGULI ALIMUJIANG

In Partial Fulfillment of the Requirements for the Degree of
MASTER OF ARTS IN PHILOSOPHY

DEPARTMENT OF PHILOSOPHY
İHSAN DOĞRAMACI BILKENT UNIVERSITY
ANKARA

June 2020

ABSTRACT

THE DEVELOPMENT OF PRESCHOOLERS' UNDERSTANDING OF CULPABILITY

Alimujiang, Reyihanguli

M.A., Department of Philosophy

Supervisor: Asst. Prof. Dr. Nazım Keven

Co-Supervisor: Asst. Prof. Dr. Jedediah Wilfred Papas Allen

June 2020

The effect of outcome and intention on children's moral judgment is much debated in developmental literature. Combining this debate with norm understanding literature we investigated children's ability to understand intentions and norm violators' culpability in a given situation. 4-year-olds were randomly assigned to two different groups that involve a puppet either intentionally violates a norm or violates the norm based on a false belief. Further, we investigated other sociocognitive abilities that may influence children's understanding of intentions and culpability. Theory of mind Battery (Wellman & Liu, 2004), Change in location False Belief Task (Wimmer & Perner, 1983) and standard Executive Functions task (Zelazo, 2006) are tested in

each participant. Regarding protest behavior during transgression and tattling behavior after transgression as salient markers of children's understanding of intention and culpability, the result shows that children tattle more when a norm violation is done with an intention to violate the rule as compared to the same norm violation that is done based on a false belief about the situation. In other words, 4-year-olds do understand intentions and decide transgressors culpability by incorporating the mental state of the transgressor. Moreover, children who could not pass the explicit false belief tasks show the same pattern in tattling while the ones who passed at least one of the explicit false belief tasks show no difference in tattling among conditions. Taken together, 4-year-olds understand and comprehend intention information especially when they do not have explicit false belief understanding.

Keywords: Culpability, False Belief, Intention, Norm Violation, Theory of Mind

ÖZET

ANAOKUL ÇAĞI ÇOCUKLARINDA SUÇLULUK ATFETME ANLAYIŞININ GELİŞİMİ

Alimujiang, Reyihanguli

Yüksek Lisans, Felsefe Bölümü

Tez Danışmanı: Dr. Öğr. Görevlisi Nazım Keven

Eş Danışman: Dr. Öğr. Görevlisi Jedediah Wilfred Papas Allen

Haziran 2020

Davranışlardaki kasıtlılık ve davranışların sonucu ile çocukların ahlaki yargıları arasındaki ilişki gelişim psikolojisi literatüründe sıkça tartışılmıştır. Bu konulara ek olarak norm anlayışını da ele alan çalışmamız, çocukların oyunun kuralını kasıtlı ya da kasıtsız olarak ihlal eden kişinin suçunu anlama yetisini ve bunu takip eden tepkilerini araştırmaktadır. 4 yaşındaki çocuklar, kuklanın oyun kuralını birinde kasıtlı olarak, diğerinde ise yanlış inanç sonucu ihlal ettiği iki farklı gruba rastgele atanmıştır. Çocukların kasıtlılık ve suçluluk anlama becerisi, oyun kuralının ihlali ardından bulundukları sözel itiraz ve geveleme tepkileri ile ölçülmüştür. Bununla birlikte, çocukların karşılaştığı davranışın kasıtlılığını ve davranışta bulunan kişinin suçlu olma durumunu anlamasında etkisi olabilecek alan-genel becerileri

değerlendirilmiştir. Bu beceriler, Zihin Kuramı Görevleri Ölçeği (Wellman & Liu, 2004), Beklenmeyen Yer Değişikliği Görevleri (Wimmer & Perner, 1983), ve Yönetici İşlevler Testi (Zelazo, 2006) ile her katılımcı için ölçülmüştür. Sonuçlar, oyun kuralının kasıtlı olarak ihlal edildiği durumda çocukların, yanlış inanç yüzünden bozulmasına kıyasla, daha çok gevelediğini göstermektedir. Bu da 4 yaşındaki çocukların, normu ihlal eden kişinin zihin durumu ve davranışın kasıtlılığını değerlendirerek suçlulukla ilgili karar verdiğine işaret etmektedir. Ek olarak, Belirgin Yanlış Kanı testini geçemeyen çocuklarda kasıtlı kural ihlali grubu yanlış inanç grubuna göre daha çok geveleme davranışında bulunulmuştur, ancak testi geçen çocuklarda iki grup arasındaki bu geveleme farkı görülmemiştir. Tüm sonuçlar birlikte değerlendirildiğinde, özellikle belirgin yanlış kanı becerileri gelişmemiş 4 yaşındaki çocukların davranışların kasıtlı olup olmadığını anladığı görülmektedir.

Anahtar Sözcükler: Kasıtlılık, Norm İhlali, Suçluluk, Yanlış İnanç, Zihin Kuramı

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Asst. Prof. Dr. Nazım Keven for his endless support and encouragement. His willingness to teach and learn simultaneously has allowed me to express what I know without constraint; His scholarship and steady focus on a given subject has allowed me to explore my own areas of interest freely. He enabled this thesis project quite literally and I learned so much from him during this process, both academically as well as research life in general.

I owe a special thanks to my co-supervisor Asst. Prof. Dr. Jed Allen. He has been present in every critical point of my academic life for the last 6 years. His structured lectures are the reason that I was interested in developmental psychology in the first place. His support and supervision throughout the years, including this dissertation project, has helped me to advance my research and self-learning ability.

I would also like to thank Asst. Prof. Dr. Hande Ilgaz who included me into Bilkent Developmental Lab to learn the most practical parts of empirical research. I'm grateful for her involvement and invaluable feedback at the beginning of this work.

I express my profound gratitude to Assoc. Prof. Dr. Simon Wigley. I could not wish for a more engaging and attentive department chair if I tried. His encouragement and pervasive presence are the essential foundations of this project.

Of course my deep appreciation goes to our research team: Dilara Erecek, Egenaz Kiraz, İrem Kılıç and Özge Sabuncu. This project would not be possible without their time, interest and commitment. I learned valuable skills from each one of them. I also thank the parents and our little participants for their contribution to our project.

I'm very fortunate to receive the warmest friendship from Bartu Karakurt with whom I have experienced the good, the bad and the ugly. Without his challenge, moral support, appreciation and comfort for the past 2 years, this project would not have been possible.

My parents are the most important contributors to anything worthwhile that I've ever done. Thank you for your silent sacrifice.

Last but not least, I thank my one and only little brother who has been a father, a mother, an older brother, true friend and a companion to me. Thank you for taking care of me in my most vulnerable moments.

TABLE OF CONTENTS

ABSTRACT.....	i
ÖZET.....	iii
ACKNOWLEDGEMENTS.....	v
TABLE OF CONTENTS.....	vii
CHAPTER 1: HISTORICAL BACKGROUND.....	1
1.1 Thesis Overview.....	2
1.2 Historical Background.....	3
1.3 Socio-cognitive Capacities Related to Intention and Moral Judgment.....	7
1.3.1 Understanding of Intention and Executive Functions	7
1.3.2 Understanding of Intention and Theory of Mind	8
CHAPTER 2: LITERATURE REVIEW.....	11
2.1 Moral Norm Violation and Conventional Norm Violation in Norm Literature.....	12
2.2 Tattling in Norm Violation Context.....	14

2.3 Protest in Norm Violation Context.....	17
2.4 Design of the Current Study.....	18
CHAPTER 3: METHOD.....	21
3.1 Participants.....	21
3.2 Materials.....	23
3.2.1 Demographic Form.....	23
3.2.2 Theory of Mind (ToM) Battery by Wellman & Liu (2004)...	23
3.2.2.1 Diverse Desires Task.....	24
3.2.2.2 Diverse Beliefs Task.....	24
3.2.2.3 Knowledge Ignorance/Knowledge Access Task.....	25
3.2.2.4 Content False Belief Task.....	26
3.2.2.5 Hidden Emotions Task.....	27
3.2.3 Change in Location False Belief Task.....	28
3.2.4 Norm Violation Task.....	29
3.2.4.1 Coding of Norm Violation Task.....	31
3.2.5 Executive Functions Task (EF)	34
3.3 Procedure	36
3.4 Coding and Reliability.....	38
CHAPTER 4: RESULTS.....	40
4.1 Preliminary Analysis.....	40
4.2 Analysis Plan.....	42
4.3 The Effect of Condition on Understanding of Culpability.....	42

CHAPTER 5: DISCUSSION.....	50
5.1 General Discussion.....	50
5.2 Conclusion.....	60
REFERENCES.....	62
APPENDIX A: DEMOGRAPHIC FORM.....	68

LIST OF FIGURES

Figure 1. Tattling Across Conditions.....	43
Figure 2. The Effect of Condition on Tattling Behavior Among Those Who Passed False Belief Tasks.....	45
Figure 3. The Effect of Condition on Tattling Behavior Among Those Who Failed False Belief Tasks.....	45

CHAPTER 1

HISTORICAL BACKGROUND

Drawing upon theoretical and empirical literature that investigated the influence of intentionality in moral development, this thesis explores culpability assessment in preschoolers in the context of conventional norm violation. Specifically, it examines whether preschoolers have a reason to reduce culpability if a norm violation takes place based on a false belief as opposed to a situation where norm violation is due to an intention to break a rule. Moreover, this project uncovers the extent to which children's understanding of intentionality is influenced by mental states and social cognitive abilities such as Theory of mind and Executive Functions.

In chapter 1, I provide a thesis overview and then present the detailed account of culpability and intention in moral judgment literature. The purpose of this chapter is to provide a sufficient historical background for our empirical research. Chapter 2 introduces a body of empirical literature that is closely related to the particular design and measurements of current study. I refer back to this chapter when discussing our study results. Chapter 3 is methods section that contains detailed information about our study design, participants, task procedures and coding. Chapter 4 is our empirical study results. Finally, chapter 5 includes in-depth discussions about our main results in the context of existing literature.

1.1 Thesis Overview

In everyday life, adults' actions are interpreted in terms of outcomes and intentions that underlie the actions (Ames & Fiske, 2017; Nobes, Panagiotaki & Pawson, 2009). Within this framework, intentions have privilege when adults evaluate someone's culpability (Wu, Hua, Yang, & Yin, 2018). For instance, a person who intentionally causes damage is considered more culpable as compared to a person who accidentally causes the same harm. The intention here is understood as the intention to do harm, or intention to break a rule when there is no harm involved.

Moreover, harm was traditionally related to wrongdoings that have moral significance, therefore intention-outcome interactions were empirically studied under the context of moral judgment both in adults (Gray, Schein & Ward, 2014) and in developmental literature (Cushman, Sheketoff, Wharton, & Carey, 2013; Nobes et al., 2009; Nobes, Panagiotaki, & Bartholomew, 2016). It has been implied that people take intention or outcome into account with a varying proportion in their moral judgment. Consequently, relevant empirical research have emerged in developmental psychology that investigated an outcome-to-intent shift in children's moral judgment. One of the claims regarding the shift was that young children cannot focus on the intention of an action, so they base their moral judgments on the outcome (Nobes et al., 2016). After significant development in middle childhood, children can make the shift of outcome-based judgments to intention-based Judgments. The context that researchers induced children's moral judgment was a

norm violation context, therefore, literature on moral norm violation and norm violation as a whole became more relevant to understand children's moral reasoning.

This thesis combines the literature from children's moral judgment and norm understanding and investigates preschoolers' understanding of culpability in conventional norm violation context.

1.2 Historical Background

Historically, Piaget (Piaget, 1965) tested children's moral understanding by presenting a simple rule-breaking scenario to a child and asked "who is naughtier?". Children were provided with a short story in which there were two children named John and Henry. John broke 15 cups while not knowing that there was a tray full of cups behind the door. Henry on the other hand broke 1 cup while his mother was out and he was using this opportunity to get some jam. When asked which one is naughtier, younger children (6-year-olds) were more likely to say that the one who broke 15 cups accidentally was naughtier than the one who also accidentally broke 1 cup with a bad intention of trying to get some jam. Another similar story was that one child tried to help his father and made a large ink stain accidentally while the other child made a small ink stain deliberately. When asked who is naughtier, younger children (6- to 7-year-olds) typically said the one who makes the large ink stain was naughtier (Piaget, 1965, p.118). In other words, children did seem to ignore the existence of bad intentions (i.e. trying to have some jam in the absence of his

mother) in cup story and accidental or deliberate nature of the consequences in ink story. Younger children typically adopted consequentialist reasoning and they assigned objective responsibility that was based on concrete magnitude of the harm that was done.

To clarify, the word “naughtier” is interpreted as “morally worse” to indicate children’s moral judgment. Researchers who use elicited-response tasks adopted words such as “good/bad” or “like/dislike” to demonstrate various degrees of moral judgment in young children starting from 4 years of age (Margoni & Surian, 2016; Imamoglu, 1975). But fundamental idea is that when evaluating something as worse, the role of intention to do harm is ignored in young children. Moreover, researchers questioned Piaget’s methodology of pairing good intention with worse harm (i.e. large ink stain) and bad intention with mild harm (i.e. small ink stain) which subsequently led them to think that perhaps intent-based evaluation occurs much earlier than Piaget predicted.

On the one hand, there are researchers who claim that children can make intent-based evaluation much earlier than Piaget predicted. They base their research on 3- to 8-year-olds and claim that even young children can take intention as underlying information in moral judgment. For example, Nobes, Panagiotaki & Pawson (2009) used stories to test the degree to which young children (3- to 4-year-olds) appreciated intention. After control questions were asked to make sure children did understand

the stories and can remember them, researchers asked several test questions such as “Is it OK for X to do that?”, “How good/bad?” (on a visual rating scale) and finally a justificatory question “Why is it good/bad for X to do Y?” (Nobes et al., 2009). They discovered that only two young children based their judgment on outcome and all the other participants based their judgments primarily on intention. Researchers’ conclusion undermined the idea that children primarily use outcome when making moral judgment. While it was arguable that the extent to which a toddler or infant can make a moral judgment as such, researchers carefully controlled for intention salience and recency effects and found that young children exhibited sensitivities for intentions (Nobes et al., 2016; Nobes, Panagiotaki, & Engelhardt, 2017; Vaish, Carpenter, & Tomasello, 2017; Van de Vondervoort & Hamlin, 2018).

On the other hand, some researchers supported primarily outcome-based moral judgments in 4-year-olds. For instance, Cushman et al. (2013) investigated children’s moral judgment pattern with vignettes that were followed by questions such as “Did [the character] want [the relevant outcome to occur]?”, and “Did [the character] actually [produce the relevant outcome]?” (Cushman et al., 2013). They found that 4-year-olds could not understand whether the actions in the stories were performed unintentionally or on purpose. 5-year-olds on the other hand “exhibited perfect performance in terms of distinguishing two scenarios”. The sensitivity to intention then increased with age after 6, but it stayed relatively stagnant between the ages of 4 and 5. There were more studies that supported a later onset of intent-based moral judgment as Piaget did (Helwig, Zelazo, & Wilson, 2016; Killen, Lynn Mulvey,

Richardson, Jampol, & Woodward, 2011; Margoni & Surian, 2016; Zelazo, Helwig, & Lau, 1996). Although some of these studies focused on moral action rather than moral reasoning in middle childhood, the conclusions are still relevant because most of them used vignettes that asked questions about children's judgment about the situation. With vignettes, we are still obtaining children's explicit moral reasoning as opposed to understanding what they would do in an actual situation.

That said, the debate about outcome-to-intent shift has become a debate about whether young children understand intentions and the age at which children acquire the ability to understand intentions in their moral judgments. Piaget however, never claimed that young children do not understand intentions. In his book, *The Moral Judgment of the Child*, Piaget (1932/1965) said the following when discussing subjective responsibility (intention-based) and moral realism (outcome-based) perspectives that influence children's moral judgment

These two attitudes may co-exist at the same age and even in the same child, but broadly speaking, they do not synchronize. Objective responsibility diminishes on the average as the child grows older, and subjective responsibility gains correlatively in importance (p. 133)(Piaget, 1965).

Moral realism perspective and subjective responsibility perspective could co-exist in children because even though young children understood the intentions when

provided with a story that was related to their own lives, children did not change their judgments: “These answers show what fine shades even some of the youngest children we questioned could distinguish and how well able they were to take intentions into account” (pp. 130–131) (Piaget, 1965). Piaget was clarifying that younger children can understand intentions but this understanding of intentions did not necessarily influence their final moral judgment. Sometimes, even though they understood the intentions, they made a judgment from moral realists’ perspective. The conclusion seems as though understanding intention is not enough to decide someone’s culpability. From a developmental perspective, younger children have a limited capacity to incorporate different aspects such as intentions and consequences of the same event to make a comprehensive judgment. Therefore, the question is not necessarily about whether children understand intentions, but how well they utilize or incorporate several different aspects together, of which one could be more salient than the other, and decide others’ culpability as a result of coherent evaluation.

1.3 Socio-cognitive Capacities Related to Intention and Moral Judgment

It is not difficult to observe that developmental changes outside of moral domain are essential to incorporate various aspects of an event. However, to our knowledge, few studies investigated the link between development of moral judgment and other developmental milestones such as executive functions and theory of mind.

1.3.1 Understanding of Intention and Executive Functions

Apart from the ability to take intentions into account, children have to develop the ability to hold several aspects of a certain situation in mind (i.e. working memory), ability to switch between intentions and outcome if necessary (i.e. flexibility in thinking) and the ability to suppress their instinctive reflex about the situation (i.e. inhibitory control). That is, children have to develop Executive Function to integrate intention and outcome when making a moral judgment (Zelazo et al., 1996).

Moreover, judging one's culpability is a complicated process that includes information integration and coordination. Rules that govern intentions and outcomes separately need to be coordinated and integrated by a higher-order rule to achieve the inclusivity/complexity of judgment (Zelazo et al., 1996). This coordination process is similar to that of standard Dimensional Change Card Sort (DCCS) task. To integrate means to have the flexibility to decide which dimension to use in a given situation. Similarly, the child needs an understanding of intention and outcome in a situation separately, and s/he also has to have the ability to integrate relevant pieces of information that are produced in that specific condition (Zelazo et al., 1996).

1.3.2 Understanding of Intention and Theory of Mind

Theory of mind is another socio-cognitive capacity that's needed when deciding someone's moral responsibility, especially in terms of intention (Chandler, Sokol & Hallett, 2001, as cited in (Cushman et al., 2013). In an attempt to reconstruct our conceptual understanding of moral judgment, Cushman (2015) deconstructed intentional actions into three elements: mental states such as beliefs and desires,

performed actions, and consequences of those actions (Cushman, 2015). According to Cushman, when someone commits a wrong-doing, there are three essential areas that we base our judgment. One is referring to the person's mental states such as beliefs or desires to conclude whether the person has committed this action intentionally. Intention to break a rule or to do harm belong in this category. The second component is the performed action in which norms and social regulations are brought in to evaluate the action itself. The third component is that the consequences of action which moral domain is claimed to be originated. That said, one of the conclusions about the intentional component of action is that intentionality of a performed action heavily depends on a person's mental states such as beliefs or desires and, potentially, on other social cognitive abilities (Cushman et al., 2013).

False belief task has been used as one of the primary tasks to test children's theory of mind. Killen et al. (2011) conducted a study in which they used false belief tasks from standard theory of mind battery. Children were also exposed to separate norm violations where one condition was straightforward harm and the other condition was an accidental moral norm violation that was based on a false belief (e.g. a violator that, without knowing its owner, threw away the character's property). Almost all children regarded all moral transgressions as wrong. But the ones who passed the standard false belief task were able to appreciate the accidental nature of transgression when it was based on a false belief (Killen et al., 2011).

Chandler et al. (2001) for example explored the probable link between the ability to represent other people's minds and moral judgment (Chandler et al, 2001, as cited in (Cushman et al., 2013). They claimed that in order to make a comprehensive moral judgment, the agent had to be able to represent other people's intentions, beliefs, and desires first. This supports the idea that it would be difficult to decide others' culpability in moral transgression, if the child did not have sufficient theory of mind ability.

The studies about ToM and moral judgment were mostly focused on morally relevant ToM and study participants were mostly in their middle childhood. Although there was one study that included 3.5-year-olds, researchers like Killen and Cushman who explored potential conceptual issues about morally-relevant Theory of mind and moral judgment mainly focused on children around 8 years of age. However, we want to study younger children, therefore, theory of mind ability in preschoolers is a separate issue that is not related to moral judgment.

CHAPTER 2

LITERATURE REVIEW

So far, I reviewed conflicting accounts of children's intention understanding and its' relationship to theory of mind and executive functions in moral judgment literature. However, children do not necessarily have to make their own judgment about moral rules at an early age because moral rules are treated as somewhat sacred, even by adults. On the contrary, assessment of rule-based violations are more prevalent at an early age. Assuming that moral reasoning and non-moral reasoning are not fundamentally distinct processes, children's reasoning of rule-based violations could be a pre-cursor of their later moral reasoning. Thus, children's culpability understanding in conventional norm violation context is essentially related to their moral understanding and is worth exploring in depth. Besides, moral judgment is empirically studied in the context of moral norm violation. Rule-based conventional norms is the regular comparison to moral norms in philosophy and psychology, and conventional versus moral norm violation topic is frequently studied in norm literature.

Therefore, in this chapter, I will first discuss developmental literature that used different norm violation settings. Then I will present two measures that are frequently used to assess children's understanding of norms. Combining the findings

from norm literature and previous moral judgment literature, I will introduce the design of our current study.

2.1 Moral Norm Violation and Conventional Norm Violation in Norm Literature

The literature on children's norm understanding studied younger children as compared to moral judgment literature. Therefore, it is important to clarify what constitutes moral norm violation and conventional norm violation for younger children (i.e. 2- to 5-year-olds). Hardecker, Schmidt, Roden & Tomasello (2016) defined moral norm violation as “matters of harm, fairness, and rights” such as stealing the reward after a play or hitting another child. Conventional norm violation, on the other hand, was about “coordination of social interactions” which was considered as more flexible and less severe such as breaking the game rule (Hardecker, Schmidt, Roden & Tomasello, 2016).

Children can differentiate between moral and conventional norm violations in a reliable manner at 3 or 4 years of age (Hardecker et al., 2016; Smetana, Ball, Jambon, & Yoo, 2018). Several studies revealed that 3-year-olds blamed moral norm violators more as compared to conventional norm violators (Josephs, Kushnir, Gräfenhain, & Rakoczy, 2016; Vaish, Missana, & Tomasello, 2011; Rakoczy, Warneken, & Tomasello, 2008). Research even showed that at 2 years of age, children reacted to norm violations by their treatment of the norm violators. Specifically, when provided with a choice of someone who violated a conventional norm and a moral transgressor, they allocated fewer cookies to moral transgressors

indicating that they were aware of the difference between conventional norm violation and moral norm violation (Smetana et al., 2018).

Similar studies have been conducted in cross-cultural contexts. For example, a study in a Turkish sample by Tunçgenç, Hohenberger, and Rakoczy (2015) claimed that starting from 2 years of age, children's norm understanding emerges and it fully develops at 3 years of age (Tunçgenç, Hohenberger & Rakoczy, 2015). Another study was conducted on Chinese preschoolers (ages around 4.5 and 6) from Hong Kong regarding moral, social-conventional, and personal events. Children treated moral transgressions as more serious and more wrong. Further, as compared to social-conventional violations and personal event violations, children viewed moral violations as the most wrong even in the absence of authority (Margoni & Surian, 2020).

Conventional norm violations were used less often in isolation as compared to moral norm violations. To our knowledge, there were not many studies that demonstrate children's verbal or non-verbal reactions to conventional norm violations (i.e. simple change of game rule). One study showed that children protested during a simple change of game rules and did not approve or even actively intervened during a rule-based conventional norm violation (Rakoczy et al., 2008).

But more importantly, conventional norm violation was not used as a setting to measure children's other capacities such as intention understanding, punishment

understanding and so forth. Presumably, because a simple game rule violation such as wrong sorting may not trigger a significant/dramatic enough response in a toddler as compared to hitting someone or damaging someone's property. We aim to contribute to this specific area of the literature where conventional norm violation is used as the context in which we study children's culpability assessment and intention understanding. Utilizing our knowledge that we obtained from moral judgment literature and norm violation literature, there were two measurements that we adopted and discussed extensively: tattling behavior and verbal protest.

2.2 Tattling in Norm Violation Context

Tattling is a distinctive behavior that is observed in children in norm violation context. It is defined as "the reporting to a second party of a third party's counter-normative behavior" (Ingram & Bering, 2010). The "counter-normative behavior" here may refer to collectively unacceptable behavior or implicitly unacceptable to the individual who reports it. Although educators regard this behavior as something reprehensible, academics have found it useful to study children's understanding of norms.

Tattling behavior is mainly studied in two different ways. Researchers either looked at whether children tattled more or less depending on their assigned conditions, or whether the content of tattling changes across conditions (e.g. tattling about the partner or about the apparatus) (Heyman, Luke & Lee, 2016; Ingram & Bering,

2010; Kachel, Svetlova, & Tomasello, 2018; Loke, Heyman, Forgie, McCarthy & Lee, 2011).

In terms of developmental trajectory, tattling behavior was reported to emerge around 18 months of age (Dunn & Munn, 1985). Starting from this age children tattled about actions that potentially induced parental discipline. For example, children tattled about their siblings' physical aggression such as hitting or property damage (Bak & Ross, 1996). By the time children reach 3 years of age, they started to tattle about any type of rule violation even if it did not directly affect themselves (Yucel & Vaish, 2018, Vaish et al., 2011). For instance, Vaish et al. (2011) found that 3-year-olds tattled significantly more when a puppet destroyed an absent victim's picture as compared to when he destroyed a blank paper.

Several studies also showed that children reliably tattle to enforce norms. For example, Yucel & Vaish (2018) constructed the situation in such a way that the child could not be blamed for the transgression. The reasoning was that if children only tattle for self-serving motives, then tattling would not be necessary. However, they found that 3-year-olds tattled significantly more in harm condition than in control condition. Thus, they concluded that tattling was used to enforce norms. Another study that investigated 4- to 11-year-olds' policing behavior when interacted with adult norm violators found that 27% of children spontaneously tattled while nearly all of them tattled when asked (Heyman et al., 2016). As a result, tattling offered a

unique way for a child to understand, incorporate, and even enforce social rules and norms (Ingram & Bering, 2010; Engelmann, Herrmann & Tomasello, 2016).

More broadly, tattling was considered as a precursor of gossip, tattling behavior shared the evaluative aspect of gossip, but tattling had several distinct features that set it apart from gossip (Ingram, 2014). One of them is that unlike gossip the norm violator does not have to be absent in order for a child to tattle. In other words, when engaging in overt tattling behavior, children do not seem to care whether the violator is present or not. Gossip on the other hand is a covert expression that conventionally happens in the absence of the violator. Another distinctive feature of tattling behavior is that the report (i.e., tattling) is usually directed towards an authority figure whereas gossip is usually directed towards a peer (Ingram & Bering, 2010).

The transition of overt reporting to covert reporting, i.e. tattling to gossip, was considered as a developmental achievement in human individuals, and the transition was explained through patterns in aggression. Ingram (2014) presented an evolutionary developmental account of aggression that was influenced by indirect reciprocity and reputation management. Given low levels of intra-group physical aggression and large group sizes of human society, one of the most effective ways to transfer the relevant information was through linguistic transmission. By this framework, even though very young children engaged in physical aggression such as hitting, the impulses of physical aggression evolved into indirect verbal aggression such as tattling where they learn that this reporting behavior can be used to punish,

enforce others to comply in social situations and even manage their own reputations. In other words, children were socialized into repressing their direct aggression into indirect aggression via language. As children socialized more, the social cost of overtly tattling in front of the norm violators became more salient. Children had to consider their own reputation, how others would perceive them, and the possibility of future retaliation, if they snitch on others openly (Ingram & Bering, 2010; Ingram, 2014; Loke et al., 2011). Therefore, greater negative consequences were attached to the behavior and the overt indirect aggression turned into a more subtle and covert verbal aggression — gossip.

2.3 Protest in Norm Violation Context

Besides tattling, researchers have been using verbal/normative protest as one of the main indicators of preschoolers' norm understanding. Protest was broadly defined as any verbal expression that aims to prevent puppet from performing the norm violation (See section 3.2.4.1 for more details and examples). Protest is different from tattling in that protest usually occurs during transgression while tattling happens in the aftermath of norm violation. Empirical studies about moral judgment elaborated protest behavior in moral norm violation settings. For instance, 3- to 5-year-olds were shown to protest when a moral norm was being violated such as stealing or breaking a third party's property (Vaish et al., 2011; Schmidt, Rakoczy & Tomasello, 2012). However, children also protested in a conventional norm violation. For example, in an arbitrary game called daxing, 2-year-olds verbally protested when a puppet broke the rule of the game (Rakoczy et al., 2008). The underlying theme that unites protest behavior in moral norm violation (i.e. when there is harm) and in

conventional norm violation (i.e., when there is no harm) is the legitimacy that is embedded in protest behavior (Kachel et al., 2018). Whether there is a moral rule or a simple game rule that is violated, the parties that are involved already assume a similar ground that is shared by both partners. Therefore, this condition legitimizes protest behavior whenever there's a norm violation regardless of its consequences being harmful or harmless.

Further, Schmidt et al. (2012) demonstrated that 3-year-olds protested equally against in-group members and against out-group members in moral norm violation.

However, when there was a rule-based norm violation, they protested more against in-group members as compared to out-group members (Schmidt et al., 2012). This showed that at such a young age, children were able to use protest to differentiate types of norm violation. It also showed their ability to differentiate in-group and out-group members by the degree to which they protest. Thus, verbal protest is a sensitive measure that is indicative of children's understanding of norms. Our study involves a rule-based norm violation, so we included protest as one of our salient measures.

2.4 Design of the Current Study

To briefly summarize, children's moral judgment was usually studied under the umbrella of norm violation where moral norm violation and conventional norm violation were compared and occurred together frequently. However, to our knowledge, conventional norm violation was not studied in-depth as a setting for children's other nuanced capacities such as intention understanding.

Therefore, considering what we have learned so far, we used a conventional norm violation setting in which a puppet made a conventional norm violation (i.e. does a wrong sorting). Moreover, we tried to bring intention understanding from moral literature into norm context, but children that were studied in moral literature were much older as compared to those in norm literature. As a result, our task manipulation, which was adopted by the studies that focused on older children, might be too difficult for 4-year-olds. Therefore, we included theory of mind and executive functions primarily for methodological reasons rather than conceptual reasons. Our study procedure was the same as Study 2 in Hardecker et al. (2016). Our study was a 2 (intentional vs. unintentional) x 2 (4-year-olds vs. 6-year-olds) between-subjects design in which participants were randomly assigned to two different conditions. One condition was based on a puppet committing an intentional rule-based norm violation. In the other condition, the puppet made the same norm violation based on a false belief.

We measured children's protest and tattling behavior, because tattling seemed to be a distinctive measure in norm violation literature among other measures such as protest, preventative behaviors, and emotional agitation during norm violation. Further, each participant was tested for executive functions and theory of mind. We aim to explore children's understanding of intention and their assessment of culpability. We mainly hypothesized that children who have better performance in false belief tasks would have a bigger difference in their tattling and protest behaviors across conditions. Moreover, we predicted that 4-year-olds will show a

bigger condition effect in their protest and tattling behavior as compared to 6-year-olds. Details of our study design and specific hypotheses are found in the upcoming sections.

CHAPTER 3

METHOD

3.1 Participants

Thirty-seven Turkish preschoolers participated in the study. At the time of testing, they were all attending preschools in Turkey's capital city Ankara. Out of these 37 participants ($M = 54.35$, $SD = 2.762$, range = 49 - 59 months), 19 of them were male and 18 of them were female.

Our plan was to collect data from at least 80 children that belong to two different age groups; thus, it was a 2 (intentional vs. false belief) x 2 (4-year-olds vs. 6-year-olds) between-subjects design. However, because of the COVID-19 pandemic, our data collection was terminated abruptly when we collected 37 data from 4-year-olds and three data from 6-year-olds. As a result, we only had 2 groups to analyze and 4-year-olds' age in month was used as a continuous variable. Moreover, two of our male participants were twins and both of them were not responding throughout entire testing session. They also did not recognize colors that were presented at the beginning of testing. Therefore, we excluded their data from the statistical analysis. Among the remaining 35 participants, one participant's SES data was missing. Meaning, in some analyses where SES was the relevant variable, our sample size was reduced to 34 (17 males and females each).

Mothers were between age of 24 - 42 ($M = 33.8$, $SD = 4.9$, $N = 34$). One mother's age was missing from our data. Mothers' education levels ranged from middle school to a master's degree and most of the mothers had a university degree (71.4%). Among mothers, 22.9 % of them had high school or middle school degrees. The remaining 5.71 % had either a masters degree or Ph.D. degree. Fathers's age were between 30 - 50 ($M = 36.5$, $SD = 4.4$, $N = 35$). Their education level ranged from high school to a masters degree and most fathers also had a university degree (68.6%). Among fathers, 22.9% of them had a high school degree and 8.6% had masters degree. Family income was as follows: 3 families had income between 1000-2999 TL/Month; 3 families had income between 3000-4999 TL/Month; 10 families had income between 5000-6999 TL/Month and 10 families had income between 7000-9999 TL/Month; 8 families had income more than 10000 TL/Month. While we were collecting our data, which was from December 2019 to March 2020, the Confederation of Turkish Trade Unions (Türk-İş) announced the starvation line as 2219 TL for a four-member family. The starvation line was defined as the minimum amount of money a family needs to have enough and balanced nutrition. Given the starvation line of 2219 TL, 82.4% of the participants that participated in our study were coming from a household where the monthly earning was 5000 TL or above. Moreover, more than half of our participants (52.9%) come from a household where monthly income was 7000 TL or above. Therefore, our sample was considered as upper-middle-income level. Most of the participants had one or more siblings (52.9 %). Number of adults who were taking care of the child in the household ranged from 1 - 4 people ($M = 2.00$, $SD = .426$).

We recruited all participants by communicating with pre-schools to advertise and distribute parental consent forms. We proceeded to test children only after we obtained written consent from both parents. We also asked for child's verbal consent before each testing session begins. Children were given stickers as gifts after testing. All testing materials and procedures were approved by Bilkent University Ethics Committee and Turkish Ministry of National Education.

3.2 Materials

3.2.1 Demographic Form

Demographic form included 13 questions that were related to educational background of both parents, household income, perceived household income, sibling number, sibling's gender, sibling's birthdate, period that child spent in any form of preschool/daycare center, and number of adults who care-take the child at home (See Appendix A).

3.2.2 Theory of Mind Battery by Wellman & Liu (2004)

ToM battery was designed to measure whether children were able to comprehend, predict other people's desires, beliefs, emotions, and intentions that can be different from one's own and act accordingly as a result. This battery included diverse beliefs, diverse desires, knowledge ignorance, content false belief, and hidden emotions tasks (Wellman & Liu, 2004). Further, our norm violation study design had a specific demand on participants' false belief understanding. Therefore, we used Wimmer and

Perner's (1983) "Maxi Task" in addition to the existing content false belief task in Wellman and Liu's ToM battery.

3.2.2.1 Diverse Desires Task

This task was designed to reveal whether the child was able to understand the fact that other people might have different desires than that of his/her own. In this task, children saw a gender-matched picture of a child with pictures of a carrot and a cookie. The experimenter said: "Here's Osman/Duygu. He/she is hungry and he/she wants something to eat. Here're two things he/ she can eat, carrot and cookie, which one do you like the most? carrot or cookie?" After the child chose one of them, the experimenter said: "well it's a great choice! However, Osman/Duygu does not like carrot/cookie (same as the child's choice), In fact, he/ she likes carrot/cookie the best (the experimenter said Osman/Duygu likes the thing that the child did not choose). Now it's mealtime! Osman/Duygu can choose only of them, only one. Which one do you think he/she chooses? Carrot or cookie?". If the child answered in a reverse manner meaning "carrot" and "cookie" or "cookie" and "carrot", then he/she got 1 full point. If the child's answer was not reversed, i.e., "carrot" and "carrot" or "cookie" and "cookie", then the child got 0. In total, reverse coding got 1 full point in diverse desires task and the highest point a child can get was 1 while the lowest points the child can get in this task was 0.

3.2.2.2 Diverse Beliefs Task

This task aimed to investigate children's ability to comprehend that when an object's location was unknown, other people may have different beliefs than their own about the location of that object. In this task, children saw a gender-matched picture of a child with pictures of a car and bushes. The experimenter said: "Here's Tolga/Dilek. He/she wants to find her/his cat. His/her cat is either in the bushes or under the car. What do you think? Where's the cat? Is it in the bushes or under the car?". After the child chose one of them, the experimenter said: "well it's a great choice! However, Tolga/Dilek thinks that his/her cat is in the bushes/under the car (the experimenter said Tolga/Dilek thinks that the place that participant did not choose is where the cat is). Now where do you think Tolga/Dilek will look for the cat? In the bushes or under the car?". If the child answered in a reverse manner meaning "bushes" and "car" or "car" and "bushes", then he/she got full points. If the child's answer was the same for both questions, i.e., "bushes" and "bushes" or "car" and "car", then the child gets 0. This task was coded the same with diverse desires task in which reverse coding got 1 full point and the highest and the lowest points were 1 and 0 respectively.

3.2.2.3 Knowledge Ignorance/Knowledge Access Task

This task assessed whether participants over-attribute their knowledge to others who were ignorant of the situation. In this task, the experimenter shook an unidentified box in front of the child and asked : "What do you say is in the box?". After the child guessed and provided an answer, the experimenter said: "Let's open it and see what's in it. Wow, there's a key in the box". Then the experimenter closed the box and asked the child: "what was in the box?" as a control question. Upon getting the right answer "key" from child, experimenter proceeded: "Now comes Baris/Merve. Baris/

Merve has never seen what's in the box. Well, does Baris/ Merve knows what's in the box?", after the child's answer the experimenter asked: "Did Baris/Merve see what's in the box?". To get full point 1, the child had to answer "no" to both questions. If the child answered "yes" to both or "yes" to one and "no" to the other question, the child got 0.

3.2.2.4 Content False Belief Task

This task aimed to assess a child's ability to hold a true belief about the content of a box while appreciating a different belief of a person who was ignorant of the unexpected content. To get full point in this task, the target question and memory question should be answered as "crayon/chocolate" and "no" respectively. Any other answer was considered as wrong and the child got 0 which was the lowest point that a child can get in this task. At the beginning of the task, experimenter brought out a Crayon box and said: "Look, here's a Crayon box. What do you think is inside this box?". If the child did not say anything or said "I don't know", the child was encouraged to guess from the pictures of crayons on the outside of the Crayon box. Upon getting the answer, the experimenter said: "Let's open and see what's inside... Wow it's really a bear inside!" while holding the bear. After a while, the experimenter put the bear back into the box and ask a memory control question of which the answer was bear. If the child got it wrong then the experimenter showed what was inside again. The experimenter showed a picture of another gender-matched child (girl or boy) and said "Now here comes Can/Canan. Can/ Canan has never ever seen inside this Crayon box. So, what does Can/Canan think is in the box? Crayon or a bear?". Whether the child answered it right or wrong, no feedback was

given on the answer. Then a memory control question was asked: “Did Can/Canan see inside this box?”.

3.2.2.5 Hidden Emotions Task

This task evaluated children’s ability to understand the fact that other people’s actual feelings could be different from what they show on their faces. It used storytelling in which 2 test questions were asked after the story. In this task children were familiarized with three different facial expressions: “happy”, “sad” and “normal”. Experimenter told the child: “I’m going to tell you a story about a child. In this story, the child may feel happy, sad, or neither happy nor sad — normal.” The experimenter pointed to different faces accordingly while saying this and then asked: “Now, can you show me which face is sad? which face is normal? which face is happy?”. This was to test whether s/he had learned to identify the faces. Then experimenter said: “Now let’s begin the story. After telling the story, I’m going to ask you what the child actually feels and what emotion her face shows. The emotion that the child actually feels may be the same or may differ.” A gender-matched child’s picture without a face was presented right next to the facial expressions card and experimenter started the story: “This story is about Mert/Melis. Mert/Melis’s aunt just came back from a trip. Before she goes on the trip, she promised that she will bring a toy to him/her. But, instead of a toy, she brought an outfit. However, Mert/Melis does not like outfits. What Mert/Melis wanted was a toy. But, Mert/Melis has to hide what he/she truly feels, otherwise, if Mert/Melis’s aunt finds out about how he/she truly feels, she will not buy him/her anything anymore”. There’re several control questions before the experimenter asked the test questions. After telling the story, the experimenter

asked: “What did Mert/ Melis’s aunt buy for him/her?”. If answered correctly, experimenter then asked: “What would Mert/ Melis’s aunt do if she finds out about his/her true feelings?”. If the child gave a wrong answer, the story was told again to make sure that the child knew the story. Upon receiving the correct answer which indicated that the child understood the story, experimenter asked the first test question: “Well then, while receiving the outfit from his/her aunt, what did Mert/ Melis truly feel? happy, sad or normal?”, and the second test question: "Well then, while receiving the outfit from his/her aunt, how did Mert/Melis try to appear? happy, sad, or normal?”. In order to get the highest point, the child had to answer the first test question as “sad” and the second test question as “happy”. No partial points were given, all the other answers got 0.

Overall in ToM battery, children can get a maximum of 5 points and the lowest possible score was 0.

3.2.3 Change in Location False Belief Task

This well-known “Maxi” task was developed by Wimmer and Perner in 1983 to measure children’s ability to hold a true belief, while being able to predict that a person who was ignorant of that information would act according to his/her own belief rather than acting according to the true information. After a story, three questions were asked from the participant; Last question was the test question and the first two were control questions. Children were expected to answer “cupboard” to the third question in order to get 1 full point which was also the highest point they can get in this task; the lowest point a child can get was 0. Five card set was used to

illustrate the short story. The experimenter narrated: “This story is about Burak/Özge and his/her mother. This is Burak/Özge and this is his/her mother. And this is their kitchen. Burak/Özge’s mother gave a piece of chocolate to Burak/Özge (showing the first card). But, Burak/Özge wanted to play outside so he/she put the chocolate in cupboard. Later Bura/Özge went out to play with his/her favorite ball (showing the third card). While Burak/Özge was playing outside, his/her mother said to herself: “Hmm, it’s too hot in cupboard. I may as well put it in refrigerator so that it won’t melt”, then she took the cake out of the cupboard and put it in the refrigerator (showing the fourth card). Later on, Burak/Özge’s mother went to rest in her bed”. The first control question asked: “Where did Burak/Özge put his/her cake before he/she went out to play?”. The second control question was: “Well, where is the chocolate now?”. The third question was the test question that asked: “Well, where will Burak/ Özge look for when he/she comes back and wants to eat his/her chocolate?”. If the child answered “cupboard” then 1 full point was given, the answer “refrigerator” got 0.

3.2.4 Norm Violation Task

This task explored children’s capacity to decide someone’s culpability. The task procedure was the same with Hardecker et al. (2016) Study 2 (Hardecker et al., 2016). In this study, the experimenters entered the room, and after a short introduction about the puppeteer and puppet, a warm-up puzzle game was played to create a familiar and comfortable environment for the participant. The aim of this puzzle game was that the participant can feel comfortable to intervene when transgressions happen on the part of the puppet. In warm-up puzzle game, several

cut-in-half animal pictures were presented. The participant and the puppet were asked to match one half of the picture of the animal with its' other half. The puppet committed an instrumental mistake in the puzzle game while experimenter 1 (E1) turned away so that the participant had the opportunity to intervene. Then E1 introduced the testing game: "catch a cardboard fish and place it in the compartment of the cardboard tower while taking turns with the puppet until the compartments are full". E1 opened the first drawer from the bottom and took one white ball and a white cube out and taught the participant and puppet "the white ball goes into the yellow compartment of the box and the white cube goes into the blue compartment of the box". E1 then told participants and the puppet that they will get a sticker as a reward for each sorting phase, if sorted according to the rule. E1 reset the game and let the participant play. After the child sorted correctly, E1 said: "yay, you just won a sticker! Now it is Ayşe/Ali's turn, let's see if s/he can play". Puppet opened the second drawer with the same content and demonstrated that s/he understood the rule of sorting. The participant and the puppet are also given 2 small boxes to put their stickers in. Then 2 test trials were performed. During testing phase, E1 played with the child in the absence of the puppet and when it was puppet's turn, puppet came back and E1 went out and waited outside of the room for 10 seconds. Puppet committed a sorting mistake during this 10 seconds. E1 came back and asked how the game went. Afterward, E1 reset the game and the second test trial was performed with the same procedure.

Intentional Condition: The puppet left the room while E1 was resetting the game.

After the participant completed his/her turn, the puppet came back. The puppet said:

“Hmm, does it go here or there, I don’t care. I’m going to put this in here and this in here” and made a sorting mistake (ball into the blue compartment and cube into the yellow compartment). Then he put the sticker into his small box which was placed within reach of the participant on the desk so that it was convenient for the participant to intervene if he/she wanted.

False-Belief Condition: The puppet left the room while E1 was resetting the game. While the puppet was outside, E1 explained that the rules of the game had changed, from now on the balls go into blue box and the cubes go into yellow compartment. After the participant completed his/her turn according to the new rule, puppet came back. The puppet said “Hmm, I know how to play this game, I am going to put this in here and this in here” and played the game according to the old rule (balls go into the yellow box and the white cubes go into the blue box) thereby making a sorting mistake according to the new rule. Then s/he put the sticker into his/her small box which was placed within reach of the participant on the desk so that it was convenient for the participant to intervene if he /she wanted.

3.2.4.1 Coding of Norm Violation Task

The specific time frame for coding specific measures were critical. We separated a single trial into three time frames: before norm violation in which E1 taught the child and puppet how to play the sorting game, during norm violation in which the puppet

made the sorting mistake and after norm violation in which the child had a chance to report what happened during transgression to E1.

Protest behavior that we adopted from Yucel & Vaish (2018) was coded during the norm violation period until the experimenter entered the room. Protest behavior was defined as any verbal intervention or hints of verbal intervention that aimed to oppose puppet's norm violation. When a child engaged in normative protest, the child got 3 points which was the highest point in this category. Normative protest was when the child interferes with the process by using normative language or by reference to the rule. For instance, "No, you cannot do that", "No, that's wrong", "You should have played it like this". Moreover, we looked for keywords such as "changed!", "no longer...", and "unfair". Imperative protest got 2 points in which the child expressed simple disagreement about the situation or commanded the puppet to stop its' action without using any reference to the rule. For example, "No!", "Don't put it there!". Not having any normative content was the key feature of imperative protest. Hints of protest got 1 point. Protest behavior that could not be assigned to either normative protest or imperative protest went into "hints of protest" category. For instance, things that were said with a protesting tone of voice "hey!" or questions such as "why are you doing that?", or irrelevant statements (i.e. irrelevant to the norm violation situation), e.g. "but I like blue colors". Moreover, if the child did both normative protest and imperative protest, the child got 3 points for the highest form of protest that s/he displayed which was normative protest. Finally, child got 0 points if s/he did not engage in any protest. Therefore, the highest point a child got in this area was 3. Further, there were 2 trials for a single child. As a result, an individual

child had 2 separate protest scores: trial 1 protest score that ranged from 0-3 and trial 2 protest score that also ranged from 0-3. Preliminary statistical analysis on trial difference showed that there was no significant trial difference in protest scores, therefore, we computed an average protest score for an individual child and used this score in the following analyses.

Tattling measure was adapted from Yucel & Vaish (2018) to fit our scenario. First of all, we added spontaneity as an extra measure. Yucel & Vaish (2018) did not have spontaneity as a separate measure since they did not ask anything to induce tattling after the norm violation. However, in Hardecker et al. (2016) study 2 procedure that we used directly, E1 came back to the room after puppet did the wrong sorting and asked the child how the game went. What child reported as an answer, before the beginning of the next trial, was coded as tattling. However, we observed in our pilot study that some children spontaneously tattled before E1 asked how the game went. Therefore, we coded spontaneity as a separate sub-measure of tattling. Meaning, if the child started to report what happened as soon as E1 entered the room as opposed to answering the question of “how the game went?”, then the child got 1 point under the name of spontaneity. If the child tattled as an answer to the question “how the game went?”, then the child got 0 points in spontaneity category.

Tattling behavior itself was then divided into 3 categories: clear tattling, hints of tattling, and no tattling. This was coded after norm violation happens: right after E1 enters the room and before the beginning of the second trial. If the child told the experimenter about puppet's actions “Ayşe/Ali did...or said...”, then it was classified as clear tattling and the child got 2 points. If the child only talked about the sorting

game not working out without referencing to the puppet, or simply said “S/he couldn’t win”, then it was considered as hints of tattling and the child was given 1 point. If the child did not say anything to report the transgression or simply answered “nice” when E1 asked how the game went, then the child got the lowest score 0 for not tattling at all. If the child said all of the above, the child got 2 points for the highest form of tattling that s/he engaged in.

Some children were shy and they were tattling about what happened when the experimenter started to introduce the second trial, in which case the report was still considered as tattling. It is coded into the suitable category and the child got a corresponding score.

3.2.5 Executive Functions Task (EF)

Dimensional Change Card Sort Task (DCCS) (Zelazo, 2006) was applied in our experiment. According to Zelazo and Muller (2002) this task measured children’s flexibility in thinking and inhibitory control (Zelazo & Müller, 2002). In this task, children were presented with two different boxes with visible prototypes attached to them and the game included shape (car vs. elephant) and color (red vs. blue) dimensions. Previous studies consistently showed that the order of the shape and color game did not contribute to any difference in result, therefore we did not use counterbalancing between games. The order was as follows. First, color game with 2 introduction/practice cards proceeded with 6 test cards. Then the shape game with 6 test cards without any practice cards. Finally, the border game with 2 practice cards

and 12 test cards. During the color game, the experimenter introduced the game saying: "We are going to play a game. This is the color game. In color game, we put all the red cards over here and only blue cards go over there." Then the experimenter sorted 2 cards as a practice. Then said: "now it's your turn, do not forget, we're playing the color game. In color game, we put all the red cards over here and only blue cards go over there." The experimenter pulled out a card and said: "this is a red/blue card, where does it go?". As it did throughout all the practice cards and test cards during the games, the switch rule was reiterated by the experimenter before presenting each card. After successfully finishing at least 4 out of 6 test cards, the child continued to be tested for the shape game without any practice. The experimenter said: "Now we are going to switch. We are not going to play the color game anymore. We are going to play the shape game. In shape game, we put cars over here and only elephants go over there. Remember, cars over here and elephants over there, this is the shape game.". E1 then put out a card and asked: "this is a car/elephant, where does it go?". After successfully finishing at least 4 out of 6 test cards, the experimenter proceeded to the next game which was the border game. Border game was an alternate dimension game in which the child was told to play color game if s/he saw a card with black border and shape game if s/he saw a card without a black border. The experimenter said: "Now we're going to play a more complicated game, it's called the border game. In border game, we play color game if we see a black border on the card like this in which case this is a card with border and it is blue/red, therefore this card goes here. Sometimes you will see cards without a black border and we play shape game in this case. This card does not have a border, and it is an elephant, therefore it goes there." After showing two cards as an

example, the experimenter proceeded with the border game until the game was finished. The border game considered was terminated if the child made more than 4 mistakes out of 12 sorting. To get a full point which was 3 within the range of 0-3 in EF task, the child had to get 1 in each of the pre-switch (i.e. color game), post-switch (i.e. shape game), and alternate dimensions (i.e. border game) respectively. Each card was assigned 1 as a sub-score and in order to get 1 full point in color and shape game, the child had to collect 4 sub-scores out of 6 full sub-score. The lowest sub-score and the lowest game score that each child could get is 0. The sub-score range for border game was 0-12 and the child had to collect at least 8 sub-scores to get 1 full point for the border game. Three games were in conditional relation to each other, meaning if the child couldn't pass any game, the subsequent game was not played and EF task was terminated.

3.3 Procedure

Parents' prior written consent and children's verbal consent were obtained before testing begins. Testing took place in a quiet room in preschools where camera equipment was already set up by the experimenters. Each child was randomly assigned to either intentional condition or false belief condition before testing took place. This study was a between-subjects design such that each child was only tested for one of the two conditions. Other tasks such as Theory of mind Battery, Change in Location False Belief Task, and Executive Functions Task were common to all participants regardless of the conditions that they were assigned to.

Tasks and sub-tasks within a task were not counterbalanced. The reason that we did not counterbalance between Theory of mind Battery, Norm Violation Task, and Executive Functions Task was that we tried different orders in the pilot study and children were bored and were not willing to play when we adopted any other orders. According to our observation, this was mostly due to the verbal descriptive nature of Theory of mind Battery and repetitive nature of Executive Functions Task. Norm Violation Task on the other hand, included exciting factors for the child such as puppet and fishing game. This specific order of tasks enabled children to stay engaged until the end of a testing session.

We also did not counterbalance within Theory of mind Battery, Norm Violation Task, and Executive Functions Task. The reason for not counterbalancing within Theory of mind Battery was that Wellman and Liu did not randomly placed subtasks. The subtasks show developmental progress. Passing a task did mean that a child was able to pass all the previous tasks. Although there is a cultural difference in the sequence of diverse beliefs and knowledge access tasks (Zhang, Shao & Zhang, 2016; Shahaeian, Peterson, Slaughter & Wellman, 2011; Selcuk, Brink, Ekerim & Wellman, 2018) such that children passed knowledge access task before passing diverse beliefs task, the rest of the subtasks kept the same developmental sequence. Therefore, counterbalancing within ToM Battery was not necessary. In Norm Violation Task, there were no parallel sub-tasks to be counterbalanced. Several phases exist, a warm-up puzzle, fishing game, and sorting game. They all have specific purposes the way that Hardecker et al. (2016) arranged them. Besides, we were uncertain whether the manipulation was age-appropriate. We attempted to keep

other factors constant, hence the step-by-step replication of the subtask order. In terms of Executive Functions Task, counterbalancing cannot happen between line game and the previous two games because being able to play line game presupposed that the child passed the preceding two games. In other words, there was a developmental progress embedded in the structure. Therefore, counterbalancing was only reasonable between pre-switch and post-switch games, meaning pre-switch can be a color game or a shape game. However, Zelazo, who designed the DCCS task, performed counterbalancing of shape game and color game and did not find order effect regardless of which (color/shape) game was presented first (Zelazo, 2006). Therefore, we did not counterbalance within Executive Functions Task.

Children had to be capable of identifying several essential colors to be able to participate in all the tasks. Therefore, E1 presented three primary colors red, yellow, blue as an initial check. Then E1 administered Theory of mind Battery, Change in Location False Belief Task, Norm Violation Task, and Executive Functions Task, in that order. Experimenter 2 pretended that she was writing her assignment when, in fact, she coded Theory of mind and Change in Location False Belief Task. When Norm Violation Task started, E2 expressed that she was bored of doing her assignment and asked to join the game. Upon finishing as a puppeteer, E2 stated that she had fun and wanted to return to her assignment. E2 then coded Executive Functions data that marked the end of a testing session. Each test lasted for 17-20 minutes for an individual child.

3.4 Coding and Reliability

Theory of mind, Change in Location False Belief Task, and Executive Functions task were coded by Experimenter 2/puppeteer during each testing session. Our Norm Violation Task, on the other hand, was coded by the author alone due to the pandemic. To ensure coding reliability, we wrote down a step-by-step coding guideline that was created by adopting and adjusting the measures from Hardecker et al. (2016), Kachel et al. (2018), and Yucel & Vaish (2018). Moreover, a single dependent variable was coded throughout all the participants, then the next dependent variable was coded during a different coding session. For instance, protest in the first trial for all the participants was coded in a single session and then protest in second trial for all the participants were coded in another coding session before the author moved on to code tattling behavior. Finally, the data was coded twice in two different periods with a reasonable time interval so that any inconsistency due to the experimenter bias could be corrected.

CHAPTER 4

RESULTS

4.1 Preliminary Analysis

First of all, we created a composite z-score for Socio-Economic-Status (SES) by combining mothers' education level and monthly income of the family. A general correlation analysis showed that there is no correlation between fathers' education level and family income ($r = .14, p = .15$); Rather, there's a significant correlation between mothers' education level and family income ($r = .53, p < .001$). Therefore, we calculated separate z-scores for mothers' education level and family income and produced a composite z-score of SES for each participant. We used this composite score in our later analyses.

Secondly, our bivariate correlation analysis revealed that gender is correlated with Theory of mind diverse desires task ($r = .20, p = 0.23$), ToM total score ($r = .26, p = .09$), and false belief total score that combines content false belief score in ToM battery and change in location false belief task score ($r = .20, p = .15$). Further, gender was significantly correlated with content false belief task in ToM battery ($r = .33, p = .04$). We did a partial correlation controlling for gender and this analysis had

similar results with the original bivariate correlation analysis that included all the variables. Therefore, gender is removed from rest of the analyses.

Thirdly, our Dependent Variables were mostly significantly correlated with each other which suggests that these variables do measure different aspects of the same thing. Further, later analyses show that even though these DVs are significantly correlated, they are not dependent on each other, which confirms the independence of our measurements.

Finally, we used non-parametric tests for protest and tattling in our later analyses for following reasons. Normality test revealed that protest score and tattling score were both bimodal distributions respectively ($M=1.78$, $SD=1.31$, $N=35$; $M=1.56$, $SD=1.24$, $N=35$). Since we have hypothesized that there is an effect of condition on DVs, we divided protest score by condition to see if the bimodal distribution of scores were due to condition effect. To clarify, phrases such as “condition effect” or “the effect of condition” in this section and in following sections simply refers to the fact that objectively speaking, children have reasons to reduce culpability of others in false belief condition. So, if they have different reactions to intentional condition versus false belief condition, then we say that there is a condition effect on a certain dependent variable. Dividing protest score into intentional condition protest score vs. false belief condition protest score will potentially produce two normally distributed protest score. The result remained bimodal for protest behavior. Similarly, we divided tattling by condition and tattling score also remained bimodal. Moreover, we

also have hypotheses about condition difference on children who have false belief understanding and those who do not, we again divided children by false belief score categories to examine whether the bimodal distribution transforms into normal distribution. The result still remained bimodal.

4.2 Analysis Plan

We investigated our main research question regarding the effect of condition on protest and tattling. Non-parametric tests were used because of the bimodal distribution of our main DVs. Specifically, Kruskal-Wallis and Mann-Whitney U tests were used to compare the effect of condition on norm-violation-related DVs. Then we focused on our second research question which is about the interaction between condition and social cognitive markers such as Theory of mind and Executive Functions. Finally, we attempted to control for variables such as SES, sibling numbers etc. to trace the source of significant differences between conditions on our DVs.

4.3 The Effect of Condition on Understanding of Culpability

The main aim of this study is to examine whether children are sensitive to the reasons to reduce others' culpability. Specifically, whether children's protest and tattling behavior change depending on conditions that they were assigned to.

For the first Kruskal-Wallis Test, test variables (DV) were protest, tattling. Our grouping variable was condition in which our minimum value was 0 and maximum

value was 1. We made sure that assumptions of Kruskal-Wallis test and Mann-Whitney U test were not violated by using Levene's Test for Equality of Variances. Kruskal-Wallis test result showed that there's a statistically significant difference only in tattling behavior across 2 conditions (Intentional condition, $n=17$; False Belief condition, $n=18$), $\chi^2(1, n=35) = 4.48, p = .034$ with a mean rank tattling score of 14.53 for false belief condition and 21.68 for intentional condition. Separate median values in two conditions for tattling behavior were not available. Instead, tattling scores above median and below median for two conditions were computed (Figure 1 below). A follow-up Mann-Whitney U test was not needed because we only have two groups to compare and it is obvious that the difference between these two groups are statistically significant regarding tattling behavior.

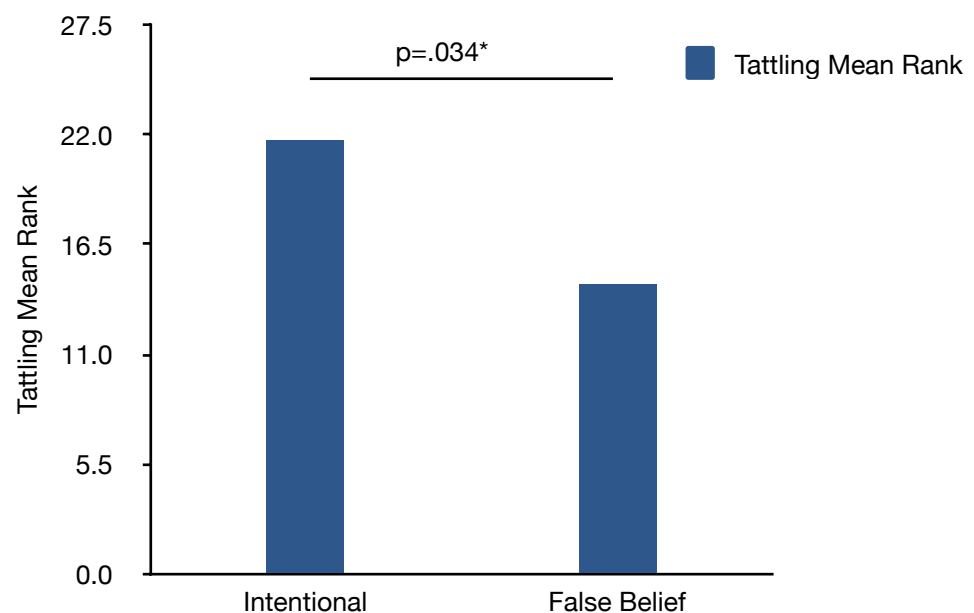


Figure 1. Tattling Mean Rank Across Conditions

Further, we divided participants into 2 groups: those who passed neither content false belief task nor change in location false belief task versus those who passed either one or both of these false belief tasks. The purpose here is to understand whether the significant condition difference in tattling score is driven by children's false belief understanding, because the expected condition difference assumes that children have an implicit understanding of false belief. Therefore, we selected participants who passed either one or both false belief tasks and did a Mann-Whitney U test by using tattling as DV and condition as IV. The result revealed that for those who passed one or both false belief task, tattling behavior did not differ among intentional condition ($Md = 2.5$, $n = 9$) and false belief condition ($Md = 2.5$, $n = 9$), $U = 34.000$, $z = -.591$, $p = .56$, $r = .14$ (Figure 2). Contrary to our expectation, for those who did not pass any of the false belief task, tattling behavior did differ significantly between intentional condition ($Md = 1.5$, $n = 8$) and false belief condition ($Md = 1$, $n = 9$), $U = 19.500$, $z = -2.177$, $p = .01$, $r = .50$ (Figure 3). This result shows that the effect of condition on tattling behavior is stronger in those children who failed both false belief tasks.

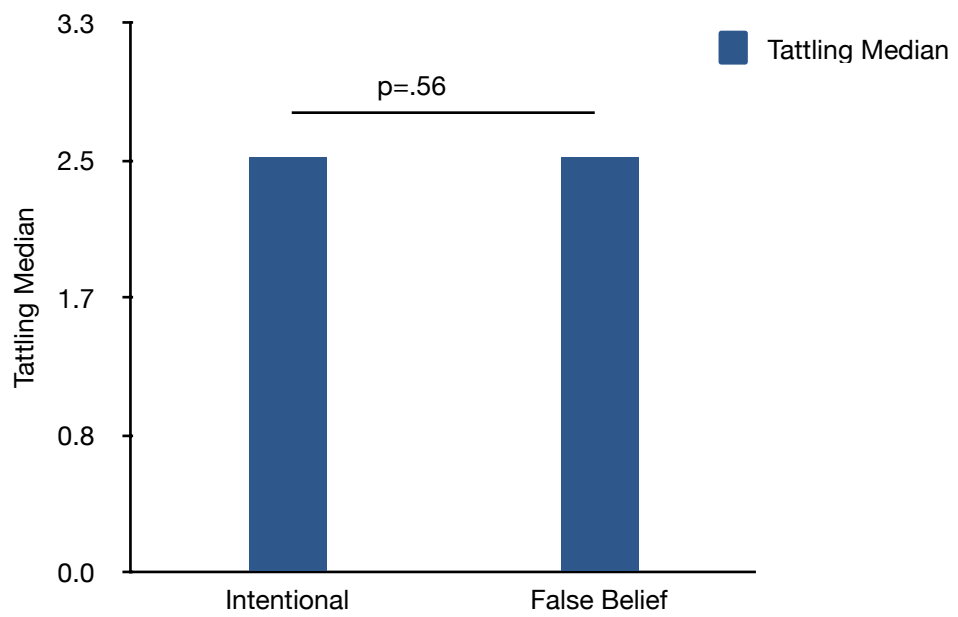


Figure 2. Tattling score across conditions among those who passed at least one of the False Belief

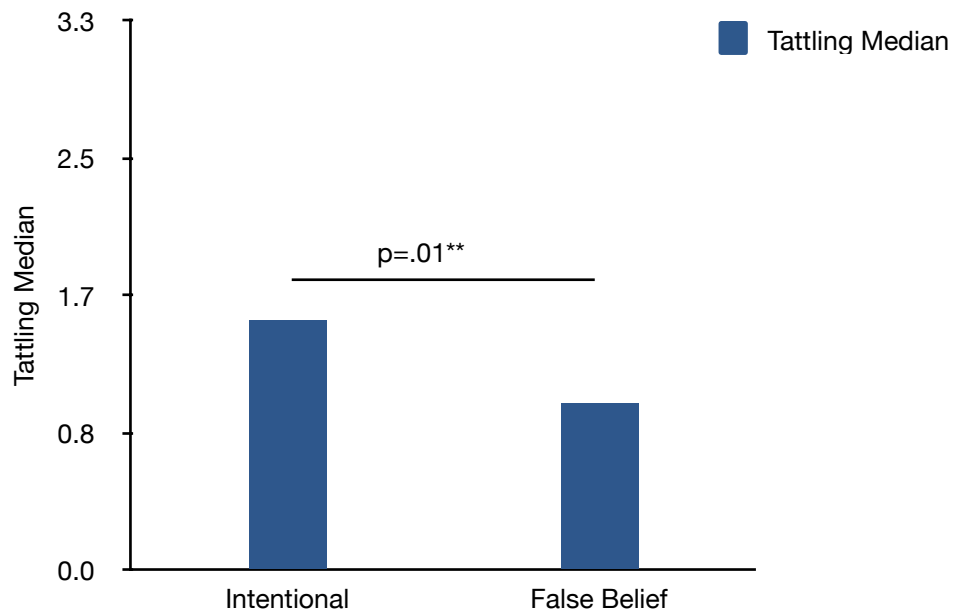


Figure 3. Tattling score across conditions among those who failed both False Belief Tasks

The composite score that is used in pass/fail false belief categories are 0, 1 and 2 which indicates failing both, passing either one and passing both false belief tasks respectively. We included scores 1 and 2 into pass category and 0 into fail category. Our original plan was to look at the ones who got 0 and 2 as separate groups, in which case those who passed would have full blown explicit false belief understanding. If the tattling behavior across conditions is still not significant among those who have full explicit false belief understanding, then we would be able to make a stronger claim about the significant effect of condition on tattling behavior. However, we only had 4 participants who passed both false belief tasks 3 of whom were in false belief condition and other 31 participants either failed both false belief tasks (scored 0) or passed one of them (scored 1). Further, participants were also divided by condition, so it would be statistically flawed to make such division and have 1 child in intentional condition and 3 children in false belief condition. Therefore, we opted for a less strict categories of passing: at least passing either content false belief task or change in location false belief task. As a result of the pass/fail division we opted for, we had 8-9 children in a single group.

When we apply normality tests to our DVs, Theory of mind total score was normally distributed. Executive Functions, even though skewed, is still normally distributed. Therefore, we used parametric analyses for this section to explore condition effects on our DVs. When using Theory of mind total score and Executive Functions task scores as DVs and condition as IV, an independent samples t-test reveals that Theory of mind does not differ among intentional ($M = 2.53$, $SD = 1.07$) and false belief

condition ($M = 2.67$, $SD = 1.24$). Executive Functions also do not significantly differ across intentional ($M = 10.82$, $SD = 5.57$) and false belief ($M = 12.39$, $SD = 5.16$) condition.

In order to explore the mediating effect of Theory of mind and Executive Functions on protest and tattling behavior across conditions, we created 2 categorical variables for Theory of mind and Executive Functions. Similar to our analyses regarding total False Belief scores, we separated ToM scores into high and low by using the median. Because ToM scores are normally distributed, mean scores are close to median scores, so we chose an actual score (i.e. median) as a cutoff point for high and low scores. Kruskal-Wallis test shows that, although non-significant, those who have high scores in ToM magnifies the effect of condition on tattling $\chi^2(1, n = 22) = 3.18$, $p = .07$) with a mean rank tattling score of 9.09 for false belief condition ($n = 11$) and 13.91 for intentional condition ($n = 11$). In comparison, those who have low scores show no difference in tattling score across intentional condition ($n = 6$) and false belief condition ($n = 7$) $\chi^2(1, n = 13) = .796$, $p = .37$) with a mean rank tattling score of 6.14 for false belief condition and 8.00 for intentional condition. In other words, although non-significant with such small sample size, having a higher or lower score in ToM battery does exhibit a trend in which those with higher score are likely to bring out the condition difference in tattling behavior even more. The same analyses were applied to Executive Functions score and results were the same with ToM such that there is no significant effect of condition on EF. Moreover, this result remained even among those who have higher executive functions ability $\chi^2(1, n = 17) = 2.73$, $p = .09$) with a mean rank of 11.36 for false belief condition and 7.35 for intentional

condition. There's also a mild trend towards a bigger condition difference in tattling behavior among those who have higher score in Executive Functions.

There were a few secondary independent variables such as SES, number of siblings and number of adults that take care of the child in household. Number of adults is almost uniformly distributed except a few outliers, meaning there is almost no variance, therefore, we did not analyze this variable. Nearly half the participants had no siblings and the other half mostly had 2 siblings, therefore we used the same method as false belief; We divided children into 2 categories: those who don't have siblings and those who do. The result of Kruskal-Wallis test shows that those who do not have siblings tattle significantly more in intentional condition as compared to false belief condition $\chi^2(1, n = 17) = 4.32, p = .038$ with a mean rank tattling score of 6.38 for false belief condition ($n = 8$) and 11.33 for intentional condition ($n = 9$). Moreover, those who do not have any sibling do not differentially protest across conditions. We did the same Kruskal-Wallis test for those who have 2 or more siblings, but we did not find any significant result on tattling and protest behavior among conditions. Composite z-score for SES is normally distributed, so we divided it into high and low scores by the mean. We did separate Kruskal-Wallis test for those who have SES scores above the mean and those who have SES scores equal to or below the mean. Tattling behavior again proved itself to be significantly different across conditions among those who have above the mean SES scores. Kruskal-Wallis test shows that those whose SES scores are above the mean tattle significantly more in intentional condition as compared to false belief condition $\chi^2(1, n = 17) = 5.70, p = .01$ with a mean rank tattling score of 6.91 for false belief condition ($n = 11$) and

12.83 for intentional condition ($n = 6$). We did Kruskal-Wallis test for those who have SES scores equal to or below the mean and there is no difference in tattling and protest across conditions.

CHAPTER 5

DISCUSSION

5.1 General Discussion

When adults engage in a transgression, the same person is held more responsible and more culpable if he does an intentional violation as compared to an accidental violation. Current research aims to examine preschoolers' understanding of intention that is potentially revealed by their understanding of culpability. Moreover, building upon the limited existing literature on the relationship between children's understanding of intentionality and theory of mind and executive functions ability, we sought to investigate the extent to which children's understanding of intentionality is influenced by social cognitive abilities such as theory of mind and executive functions. Overall, we found that 4-year-olds express their understanding of intentions behind a norm violation by differentially tattling about the norm violator. Meaning, they tattle significantly more when the transgression is done intentionally as compared to when it's done unintentionally (i.e. due to a false belief). Interestingly, the difference in tattling across condition disappears as they start to have some explicit false belief understanding. One interpretation could be that because of domain general cognitive development, 4-year-olds start to understand the social cost of tattling and they tattle less as a result.

The result concerning our main research question is consistent with literature. Our research, despite the small sample size, has found a significant effect of condition in tattling behavior. Specifically, depending on people's (puppet's) intention or lack thereof to violate the game rule, children tattled significantly more in intentional condition as compared to false belief condition. This shows that children understand intentionality such that they tattle differently when a violation is done deliberately as compared to a violation that is done based on false belief (non-deliberate). This result is probably due to the fact that tattling is a counter-normative report by definition and when there is a clear violation of game rule children tattle to indicate their understanding of the situation. This is also shown by Heyman et al. (2016) where almost all 4- to 11-year-olds that are involved in the study tattled about the transgression (Heyman et al., 2016). Moreover, children's willingness to tattle more or less is an implicit measure of their culpability assessment. For example, in a study, 3-year-olds tattled significantly more in harm condition as compared to no harm condition (Yucel & Vaish, 2018). Meaning, children tattle uniformly when there's a transgression and they tattle to show their sensitivities to the difference between harm and no harm when a norm is equally violated. In line with these studies, our main result indicates that 4-year-old Turkish children tattle to differentiate others' intention to break the rule from unintentionally breaking the rule, even when there is no active harm involved.

Contrary to our expectation, children who have no competence in false belief are tattling differently according to conditions they're in. They tattle less when the

person does the transgression based on false belief as compared to when the person does the transgression deliberately. We originally hypothesized that if children's tattling response to the norm violation is due to their understanding of intentionality and false-belief-based nature of two conditions, the ones who passed explicit false belief tasks show a greater difference between conditions as compared to those who failed in both false belief tasks. However, our result shows the opposite, that the significant effect of condition on tattling does not belong to the ones who passed one of the explicit false belief tasks but comes from those who failed both false belief tasks.

This finding could be interpreted from social cognitive development perspective while bearing our original age-related hypothesis in mind. We hypothesized that 4-year-olds differentially tattle in intentional condition as compared to false belief condition and this condition effect on tattling would disappear in 6-year-olds. Because as children socialize more, they consider their reputation and the possibility of future retaliation, if they snitch on others openly (Ingram & Bering, 2010; Ingram, 2014; Loke et al., 2011). A natural conclusion could be that children stop tattling as they age. However, Yucel and Vaish (2018) found that the difference between younger (4- to 7-year-olds) and older children (8- to 11-year-olds) is not that older children tattle less, but younger children are more likely to tattle without being prompted. In other words, spontaneity seems to be a significant marker of snitching that brings negative consequences. As a result, the amount of tattling/reporting does not change, but spontaneity does as children get older. Passing at least one of our false belief task shows more advanced socio-cognitive abilities as compared to

failing both false belief tasks. Meaning, children who can pass one of the explicit false belief tasks or both are more developed socially and cognitively such that they start to show effects of socialization. Given that our tattling score is a composite score of tattling and spontaneity, we can conclude that we're seeing the expected age effect even within 4-year-olds. Those who failed both false belief tasks are the cognitively younger ones who show their understanding of intention through their tattling behavior. Those who pass at least one of the false belief tasks are the cognitively older ones who do not show a condition effect. However, there is not enough empirical evidence to fully support this socialization explanation. In future analyses, we can separate out the spontaneity score from tattling composite score in 4-year-olds and add 6-year-olds' data to see if spontaneity diminishes with age as Yucel and Vaish (2018) found in their wider age range. Another intriguing phenomenon is that we observed 2 of our participants whispered and pointed to the boxes to show the experimenter what puppet did instead of tattling like their other counterparts. This whispering seems to be a sign of covert aggression/gossip where the child does not blatantly snitch on the puppet. Both children passed both of our false belief tasks and their theory of mind and executive functions score are at the higher end, showing their socio-cognitive competence. Therefore, in accordance with Ingram's (2014) claim that tattling is a precursor of gossip and overt verbal aggression (i.e. tattling) transforms into more covert verbal aggression (i.e. gossip), we might find that 6-year-olds who have a better understanding of false belief will gossip instead of overtly tattling like 4-year-olds.

Contrary to our hypothesis on verbal protest, we did not find any significant condition effect with this DV. Children did not differentially protest depending on the conditions that they are assigned to. This result is consistent with Hardecker et al. (2016). They also did not find any significant result regarding protesting behavior (Hardecker et al., 2016). It is probably due to the fact that both conditions include a rule violation and a mere rule violation is enough to induce a protest behavior. Unlike tattling where it might not be necessarily warranted if the person is not aware of the new rule (i.e. false belief condition), protest is warranted during transgression in both conditions. Moreover, in Hardecker et al. (2016) where a conventional norm is violated in one condition and a moral norm is violated in the other, one would expect protest behavior to be significantly different across conditions. However, they found no difference and they argued that protest behavior measures a general discomfort that cause one to protest which is in fact consistent with studies in norm understanding literature where they report that even 3-year-olds protest in norm violating situations (Rakoczy et al., 2008; Vaish et al., 2011). Ultimately, puppet is misbehaving/doing the wrong sorting and protest seems to be more closely related to the objective fact of puppet's rule violation. Our result on protest shows that children do not differentially protest across intentional and false belief conditions. Thus, we tentatively conclude that normative protest may not be closely related to intentionality understanding while tattling proved itself to be a more sensitive measure of intentionality understanding. Perhaps spontaneity component of tattling is what makes tattling substantially different from protest because all children reliably tattle when asked directly, yet spontaneity was the sensitive mark between younger children and older children (Yucel & Vaish, 2018). Protest among preschoolers on

the other hand seems to be a uniform reaction towards the objective environment during the transgression.

We expected an interaction between theory of mind and condition and executive functions and condition, we did not find any significant result. There is no statistically significant interaction between theory of mind and condition, nor there is a statistically significant interaction between executive function and condition. However, There is a noticeable trend regarding interaction between Theory of mind and condition. Those who have high Theory of mind scores tend to tattle more in intentional condition as compared to false belief condition. The same tendency exists for Executive Functions in that children who have high score in Executive Functions tend to magnify condition difference with their tattling behavior (See results section for high-low score division details of theory of mind and executive functions). This trend reflects our original hypothesis where we predicted a mediating effect of ToM and EF on condition. The basis of this hypothesis is that children need certain level of ToM and EF competency for our task manipulation to work. The result being not significant shows that certain degree of ToM and EF understanding is not a precondition for our task manipulation to be effective. However, the trend is in opposition with false belief result. While children who have no false belief understanding show a bigger condition difference in tattling, children who have high scores in ToM and EF show greater condition effect in their tattling behavior.

The reason that we compared ToM and EF related interaction with false belief related interaction was that a) false belief competency is a subset of Theory of mind competency b) we isolated false belief score from ToM total score and ran separate analyses since we thought children especially need certain level of false belief understanding among other Theory of mind abilities for our task to work. Therefore, the relation of ToM competency to tattling behavior was explored through two separate routes that directed towards opposite ways: the unexpected interaction between false belief and tattling and the expected interaction between ToM and tattling.

As difficult as it is to reconcile this conflicting result, there are reasons to assume that one of them is a more reasonable direction. Regarding the interaction between false belief and condition, our task manipulation is essentially a false belief manipulation, not a Theory of mind manipulation. Children's performance on false belief tasks are more reflective of their sensitivity to our task manipulation. Moreover, our non-significant result on the interaction between ToM and condition shows that Theory of mind competence is not a necessary precondition for our task to work. Therefore, the significant result we found regarding false belief and tattling among condition is more in line with our initial concern of incorporating false belief (potentially ToM related) tasks in the beginning. We also have reasons to consider the ToM related interaction route as more reasonable. There are five tasks that measure children's ToM ability in ToM Battery and it is a more sensitive measure as compared to two false belief tasks one of which is from the ToM Battery. Especially if we want to use socialization explanation for the unexpected false belief interaction

that is discussed earlier. The battery gives us more comprehensive picture of children's socio-cognitive ability as compared to false belief tasks alone. Further, it is reasonable to admit that cognitively older ones (i.e. those who scored high in ToM) are more sensitive to our manipulation such that they showed their understanding of manipulation through tattling noticeably more in intentional condition as compared to false belief condition. The fact that the interaction between Executive Functions and condition exhibiting the same trend with ToM adds another reason for us to consider this route: perhaps ToM direction (as opposed to FB direction) is more reasonable because this route is more in line with existing literature on relationship between ToM and EF (Carlson, Clexton & Moses, 2015).

The conflicting result could potentially be explained better when we test several continuous age groups. We were able to test a single age group, 4-year-olds. However, children's performance on ToM will probably show a better developmental trajectory when tested among 3-, 4-, and 5-year-olds at the same time. Moreover, 4 years of age is the most volatile age in terms of social cognitive development. More research that include older children with several age groups need to be conducted to examine the overall trajectory of condition effects on socio-cognitive development.

Moving away from our hypotheses, sibling number and SES seem to influence children's tattling but not protest behavior. Namely, children who do not have any sibling tattle significantly more in intentional condition as compared to false belief condition. Moreover, children who have higher SES score also tattled significantly

more in intentional condition. We also did not find a correlation between SES and sibling number in this sample. Detailed literature and more focused analysis are needed to interpret and incorporate the sibling and SES related results into our main discussion. Further, we had to exclude the preschool duration data from our analyses, because it was filled out in inconsistent manner. The other two variables, i.e. sibling number and number of adults who take care of the child are only used as covariates. More substantial analyses have to be done in this area to provide a full spectrum of variables that influence children's judgment of others' culpability and their norm understanding.

There is a conceptual reason that we gathered these relationship-related informations. Piaget (Piaget, 1965) argued that the salient relationship dynamic children experience influence their understanding of norms. He argued that children construct their social norms via two different types of relationship: that of constraint and that of cooperation. The relationship between children and their parents is a typical constraint relationship in that adults tend to prescribe certain rules that already exists and expect children to adopt/obey them. Piaget defined this relationship "relations of constraint, whose characteristic is to impose upon the individual from outside a system of rules with obligatory content, and relations of cooperation whose characteristic is to create within people's minds the consciousness of ideal norms at the back of all rules" ((Piaget, 1965), p. 395). Untouchable and somewhat sacred ideal norms are accepted and maintained through relations of constraint. Peer relationship on the other hand is a typical cooperative relationship in that mutual respect are in the center of this type of relationship. Rules are often negotiated

among children to include the best interests of the majority involved. Inherent coercion of moral rules does not seem to exist in cooperative relationships. Piagetian claim is that young children are immersed in constraint relationship where “generally speaking adults deal very harshly with clumsiness” (p. 131). Therefore, even though children understood the intentions in those stories of “who is naughtier?” they tend to be outcome-based moral realists in their moral judgment. From this perspective, changes in children’s major relationships influence their understanding of social norms and potentially affect other salient aspects such as intentionality when a transgression happens. Our study was not held in moral norm violation context. However, relationships that Piaget described seem to persist over a person’s life and probably influence areas other than moral domain. Therefore, more research is needed in this regard.

There were several limitations to our study. First limitation was that we were not able to collect and analyze all the data according to our original study plan due to the pandemic. If we collect rest of the data from 6-year-olds, we can interpret tattling behavior and its relation to false belief and theory of mind more thoroughly. In fact, a larger sample that includes 4-, 5- and 6-year-olds will allow us to see the developmental trajectory of culpability understanding. Moreover, our main dependent variables were not normally distributed, which prevented us from doing more diverse statistical analyses. With a larger sample, our dependent variables have greater probability of being normally distributed, which will enable us to analyze tattling and protest behavior with more flexibility. Secondly, our sample was considered to be from upper-middle-income level, so we cannot necessarily

generalize our results to a larger population. However, future research could choose a more representative sample from Turkish population and contribute to the existing literature with a cultural comparison.

5.2 Conclusion

Despite these limitations, we combined norm violation literature with children's intention understanding and their judgment of culpability in our study to have a more comprehensive appreciation for developmental trajectory of children's intention understanding and culpability decision. To our knowledge, only a limited number of studies have attempted to explore this area extensively. The results of the current study suggest that 4-year-olds in Turkish sample do appreciate the difference between intentional norm violation and unintentional norm violation that's based on a false belief. 4-year-olds show their understanding of the difference between intentional and unintentional norm violation through their tattling behavior.

Unexpectedly however, this appreciation of the difference between intentional transgression and transgression based on false belief mainly comes from the ones who do not have explicit false belief understanding. Current study also showed a cross-cultural consistency in protest response in that condition does not have an effect on 4-year-old Turkish Children's protest behavior.

This study is one of the first to include social cognitive milestones such as Theory of mind and Executive Functions to the developmental literature in intention and culpability understanding, but the findings need to be re-examined with a larger and

more diverse sample. More research should be conducted to explore the developmental flow of intention and culpability understanding starting from 3 years of age to late childhood.

REFERENCES

- Ames, D. L., & Fiske, S. T. (2017). HHS Public Access. *Physiology & Behavior*, 176(3), 139–148. <https://doi.org/10.1016/j.physbeh.2017.03.040>
- Bak, I. M., & Ross, H. S. (1996). I'm Telling! The Content, Context, and Consequences of Children's Tattling on Their Siblings. *Social Development*, 5(3), 292-309. <https://doi.org/10.1111/j.1467-9507.1996.tb00087.x>
- Carlson, S. M., Claxton, L. J., & Moses, L. J. (2015). The relation between executive function and theory of mind is more than skin deep. *Journal of Cognition and Development*, 16(1), 186-197. <https://doi.org/10.1080/15248372.2013.824883>
- Cushman, F. (2015). Deconstructing intent to reconstruct morality. *Current Opinion in Psychology*, 6, 97–103. <https://doi.org/10.1016/j.copsyc.2015.06.003>
- Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intent-based moral judgment. *Cognition*, 127(1), 6–21. <https://doi.org/10.1016/j.cognition.2012.11.008>
- Dunn, J., & Munn, P. (1985). Becoming a Family Member: Family Conflict and the Development of Social Understanding in the Second Year. *Child Development*, 56(2), 480-492. doi:10.2307/1129735
- Engelmann, J. M., Herrmann, E., & Tomasello, M. (2016). Preschoolers affect others' reputations through prosocial gossip. *British Journal of Developmental Psychology*, 34(3), 447–460. <https://doi.org/10.1111/bjdp.12143>
- Gray, K., Schein, C., & Ward, A. F. (2014). The myth of harmless wrongs in moral cognition: Automatic dyadic completion from sin to suffering. *Journal of experimental psychology. General*, 143(4), 1600–1615. <https://doi.org/10.1037/>

- Hardecker, S., Schmidt, M. F. H., Roden, M., & Tomasello, M. (2016). Young children's behavioral and emotional responses to different social norm violations. *Journal of Experimental Child Psychology*, 150, 364–379. <https://doi.org/10.1016/j.jecp.2016.06.012>
- Helwig, C. C., Zelazo, P. D., & Wilson, M. (2016). *Children's Judgments of Psychological Harm in Normal and Noncanonical Situations* Author (s): Charles C. Helwig, Philip David Zelazo and Mary Wilson Published by: Wiley on behalf of the Society for Research in Child Development Stable URL: [http://w.72\(1\), 66–81](http://w.72(1), 66–81).
- Heyman, G. D., Chiu Loke, I., & Lee, K. (2016). Children spontaneously police adults' transgressions. *Journal of Experimental Child Psychology*, 150, 155–164. <https://doi.org/10.1016/j.jecp.2016.05.012>
- Imamoğlu, E. (1975). Children's Awareness and Usage of Intention Cues. *Child Development*, 46(1), 39-45. doi:10.2307/1128831
- Ingram, G. P. D. (2014). From hitting to tattling to gossip: An evolutionary rationale for the development of indirect aggression. *Evolutionary Psychology*, 12(SPECIALISSUE.2), 343–363. <https://doi.org/10.1177/147470491401200205>
- Ingram, G. P. D., & Bering, J. M. (2010). Children's tattling: The reporting of everyday norm violations in preschool settings. *Child Development*, 81(3), 945–957. <https://doi.org/10.1111/j.1467-8624.2010.01444.x>
- Josephs, M., Kushnir, T., Gräfenhain, M., & Rakoczy, H. (2016). Children protest moral and conventional violations more when they believe actions are freely chosen. *Journal of Experimental Child Psychology*, 141, 247–255. <https://doi.org/10.1016/j.jecp.2016.06.012>

doi.org/10.1016/j.jecp.2015.08.002

Kachel, U., Svetlova, M., & Tomasello, M. (2018). Three-Year-Olds' Reactions to a Partner's Failure to Perform Her Role in a Joint Commitment. *Child Development, 89*(5), 1691–1703. <https://doi.org/10.1111/cdev.12816>

Killen, M., Lynn Mulvey, K., Richardson, C., Jampol, N., & Woodward, A. (2011). The accidental transgressor: Morally-relevant theory of mind. *Cognition, 119*(2), 197–215. <https://doi.org/10.1016/j.cognition.2011.01.006>

Loke, I. C., Heyman, G. D., Forgie, J., McCarthy, A., & Lee, K. (2011). NIH Public Access. *Bone, 23*(1), 1–7. <https://doi.org/10.1038/jid.2014.371>

Margoni, F., & Surian, L. (2016). Explaining the U-shaped development of intent-based moral judgments. *Frontiers in Psychology, 7*(February), 1–6. <https://doi.org/10.3389/fpsyg.2016.00219>

Margoni, F., & Surian, L. (2020). Conceptual continuity in the development of intent-based moral judgment. *Journal of Experimental Child Psychology, 194*, 104812. <https://doi.org/10.1016/j.jecp.2020.104812>

Nobes, G., Panagiotaki, G., & Bartholomew, K. J. (2016). The influence of intention, outcome and question-wording on children's and adults' moral judgments. *Cognition, 157*, 190–204. <https://doi.org/10.1016/j.cognition.2016.08.019>

Nobes, G., Panagiotaki, G., & Engelhardt, P. (2017). *October 2016, revised 12. October 2016*.

Nobes, G., Panagiotaki, G., & Pawson, C. (2009). The influence of negligence, intention, and outcome on children's moral judgments. *Journal of Experimental Child Psychology, 104*(4), 382–397. <https://doi.org/10.1016/j.jecp.2009.08.001>

Piaget, J. (1965). *The moral judgment of the child*. London: Routledge and Paul.

- Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The Sources of Normativity: Young Children's Awareness of the Normative Structure of Games. *Developmental Psychology*, 44(3), 875–881. <https://doi.org/10.1037/0012-1649.44.3.875>
- Schmidt, M. F., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition*, 124(3), 325–333. <https://doi.org/10.1016/j.cognition.2012.06.004>
- Selcuk, B., Brink, K. A., Ekerim, M., & Wellman, H. M. (2018). Sequence of theory-of-mind acquisition in Turkish children from diverse social backgrounds. *Infant and Child Development*, 27(4), e2098. <https://doi.org/10.1002/icd.2098>
- Shahaeian, A., Peterson, C. C., Slaughter, V., & Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Developmental Psychology*, 47(5), 1239–1247. <https://doi.org/10.1037/a0023899>
- Smetana, J. G., Ball, C. L., Jambon, M., & Yoo, H. N. (2018). Are young children's preferences and evaluations of moral and conventional transgressors associated with domain distinctions in judgments? *Journal of Experimental Child Psychology*, 173, 284–303. <https://doi.org/10.1016/j.jecp.2018.04.008>
- Tunçgenç, B., Hohenberger, A., & Rakoczy, H. (2015). Early Understanding of Normativity and Freedom to Act in Turkish Toddlers. *Journal of Cognition and Development*, 16(1), 44–54. <https://doi.org/10.1080/15248372.2013.815622>
- Vaish, A., Carpenter, M., & Tomasello, M. (2017). *Young Children Selectively Avoid Helping People With Harmful Intentions* Author (s): Amrisha Vaish , Malinda Carpenter and Michael Tomasello Source : *Child Development* , Vol . 81 , No . 6

(NOVEMBER / DECEMBER 2010), pp . 1661-1669 Published by : Wiley .

81(6), 1661–1669.

- Vaish, A., Missana, M., & Tomasello, M. (2011). Three-year-old children intervene in third-party moral transgressions. *British Journal of Developmental Psychology*, 29(1), 124–130. <https://doi.org/10.1348/026151010X532888>
- Van de Vondervoort, J. W., & Hamlin, J. K. (2018). Preschoolers focus on others' intentions when forming sociomoral judgments. *Frontiers in Psychology*, 9(OCT), 1–17. <https://doi.org/10.3389/fpsyg.2018.01851>
- Wellman, H. M., & Liu, D. (2004). Theory of mind task scale. PsycTESTS Dataset. <https://doi.org/10.1037/t50435-000>
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128. [https://doi.org/10.1016/0010-0277\(83\)90004-5](https://doi.org/10.1016/0010-0277(83)90004-5)
- Wu, X., Hua, R., Yang, Z., & Yin, J. (2018). The influence of intention and outcome on evaluations of social interaction. *Acta Psychologica*, 182(616), 75–81. <https://doi.org/10.1016/j.actpsy.2017.11.010>
- Yucel, M., & Vaish, A. (2018). Young children tattle to enforce moral norms. *Social Development*, 27(4), 924–936. <https://doi.org/10.1111/sode.12290>
- Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, 1(1), 297–301. <https://doi.org/10.1038/nprot.2006.46>
- Zelazo, P. D., Helwig, C. C., & Lau, A. (1996). Intention, Act, and Outcome in Behavioral Prediction and Moral Judgment. *Child Development*, 67(5), 2478–

2492. <https://doi.org/10.1111/j.1467-8624.1996.tb01869.x>

Zelazo, P. D., & Müller, U. (2002). Learning to Be an Education Leader: How a Web-Based Course Meets Some of the Challenges of Leadership Training and Development. *Scholar-Practitioner Quarterly*, 3(1), 25–38.

Zhang, T., Shao, Z., & Zhang, Y. (2016). Developmental steps in theory of mind of typical Chinese children and Chinese children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, 23, 210-220. <https://doi.org/10.1016/j.rasd.2015.10.005>

APPENDIX A

DEMOGRAPHIC FORM

1. Adınız ve soyadınız:

2. Yaşınız: Cinsiyetiniz: ☐ Kadın ☐ Erkek

3. Eşinizin Yaşı: Eşinizin Cinsiyeti: ☐ Kadın ☐ Erkek

4. Çocuklarınız/çocuğunuzla olan yakınlığınız:

☐ Anne ☐ Baba ☐ Diğer:

5. Kaç çocuğunuz var?

☐ 1 ☐ 2 ☐ 3 ☐ Diğer:

6. Çocuğunuzun/çocuklarınızın cinsiyeti: Çocuğunuzun/çocuklarınızın doğum tarihi:

_____ → _____
_____ → _____
_____ → _____

7. Çalışmamıza katılan çocuğunuz tam/yarım zamanlı olarak bir eğitim kurumuna gitmeye ilk hangi tarihte başladı?

8. Eğitim durumunuz nedir?

☐ Okuryazar değil

☐ İlköğretim

☐ Ortaokul

☐ Lise

☐ Üniversite

☐ Yüksek Lisans

☐ Doktora

☐ Diğer: _____

9. Eşinizin eğitim durumu nedir?

☐ Okuryazar değil

☐ İlköğretim

☐ Ortaokul

☐ Lise

☐ Üniversite

☐ Yüksek Lisans

☐ Doktora

☐ Diğer: _____

10. Evinizin aylık gelir düzeyi:

- ☐ 1.000 TL'den az
☐ 1.000 TL- 2.999 TL
☐ 3.000 TL- 4.999 TL
☐ 5.000 TL- 6.999 TL
☐ 7.000 TL- 9.999 TL
☐ 10.000 TL'den fazla

11. Evinizin aylık gelir düzeyini nasıl değerlendirirsiniz?

- ☐ Düşük ☐ Orta Seviyede ☐ İyi seviyede ☐ Çok iyi seviyede

12. Toplam olarak evde çocuğunuzla ilgilenen kaç yetişkin yaşamaktadır?

- ☐ Bir ☐ İki ☐ Üç ☐ Üçten fazla