UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE CIENCIAS

EVALUACIÓN DE CONDUCTAS DE RIESGO EN UN CONTEXTO COOPERATIVO

TESIS
QUE PARA OBTENER EL TÍTULO DE BIÓLOGO
PRESENTA:

Santiago Gracia Garrido

DIRECTOR DE TESIS:
Dr. Marcos Francisco Rosetti Sciutto

CIUDAD UNIVERSITARIA, CDMX
2018

HOJA DE DATOS DEL JURADO

1. **Datos del alumno**
   Gracia
   Garrido
   Santiago
   5541813271
   Universidad Nacional Autónoma de México
   Facultad de Ciencias
   Biología
   307527426

2. **Datos del tutor**
   Dr.
   Marcos Francisco
   Rosetti
   Sciutto

3. **Datos del sinodal 1**
   Dra.
   Robyn Elizabeth
   Hudson

4. **Datos del sinodal 2**
   Dr.
   Carlos Rafael
   Cordero
   Macedo

5. **Datos del sinodal 3**
   Dr.
   Arturo Carlos II
   Becerra
   Bracho

6. **Datos del sinodal 4**
   Dra.
   Lynna Marie
   Kiere

7. **Datos del trabajo escrito**
   Evaluación de conductas de riesgo en un contexto cooperativo
   86 p.
   2018

Agradecimientos

A mi comunidad Montessori y a la UNAM, por propiciar el pensamiento. Estas dos instituciones me han hecho sentir libre; siempre estaré agradecido por ello.

A Marcos Rosetti, por procurar un trato horizontal, por sus reflexiones, por nuestros diálogos sin freno, por su paciencia y porque siempre me hizo sentir acompañado durante este proceso. Marcos ha sido un eje central en mi formación académica e intelectual.

A Vicente Zarco, por la confrontación, por darme un empujón, por lo que hemos estado pensando juntos.

A Robyn Hudson, por ser el germen del grupo que hemos formado en el laboratorio, por siempre mantener un ambiente académico y a la vez, de camaradería; por su humildad intelectual.

A mi sinodales, por las sugerencias y observaciones críticas. Sus comentarios resultaron muy útiles para la concreción y conclusión del trabajo.

A los miembros del laboratorio, por el buen rollo, por generar un ambiente intelectualmente estimulante sin caer en pedanterías, por el sentimiento de unidad.

A mi familia: a mis padres, por su amor imperecedero; a mi hermano, por su respaldo, porque siempre me he sentido protegido por él.

A mis padrinos: Pati y Alex, por su presencia, por todas las muestras de cariño que he recibido desde mi infancia.

A mis amigos, mis compañeros de camino, ellos le dan sentido a mi vida. Especialmente agradezco a Leonardo, por las largas conversaciones que me espabilan y me alientan. Quisiera también expresar mi gratitud a Kev, por su tiempo y disposición durante el muestro; sin él no habría sido posible.

A las escuelas primarias Tlamatini, Ejercito Nacional y José Jesús Aguilera Palomino, por apoyar la investigación. Agradezco a los estudiantes que formaron parte del estudio.

Al Almagro, por su calidez y cotorreo, porque en ciertos momentos ha sido mi fuga.

TABLE OF CONTENTS

Evaluación de conductas de riesgo en un contexto cooperativo

1. Introducción

El riesgo constituye un aspecto ubicuo del comportamiento humano, desde decisiones cotidianas y en apariencia intrascendentes como la elección de llevar el paraguas, hasta decisiones cuyas consecuencias pueden traducirse en daños irreversibles (e.g. lesiones, enfermedades) o pérdidas definitivas (e.g. vidas, tiempo, recursos). Las conductas de riesgo son aquellas que suponen una mayor recompensa con una menor probabilidad de éxito. Un ejemplo concreto del planteamiento anterior es el cruce a pie de una vía transitada: esperar el alto y cruzar por el paso peatonal representa la opción segura; la opción arriesgada sería cruzar durante el flujo vehicular. Ésta última supone una ganancia en términos de tiempo, así como la probabilidad (en menor o mayor mediada) de ser atropellado.

En vista de lo anterior, las decisiones de riesgo y las conductas derivadas, implican la expectativa de beneficios potenciales en condiciones de incertidumbre, lo cual significa que la opción implementada puede conducir a múltiples desenlaces y al menos uno de estos resulta desfavorable. Este hecho da cuenta de la importancia de asumir riesgos, la cuestión no es evitarlos sino correr los pertinentes: el riesgo paga.

La complejidad del comportamiento humano debe de ser explicada a través de la interacción entre bases biológicas (características inherentes) y factores ambientales (características adquiridas). Las conductas de riesgo no son la excepción, ciertamente algunas variables biológicas se encuentran estrechamente asociadas a tales conductas (e.g. producción de testosterona e influencia de vías dopaminérgicas); sin embargo, resultan insuficientes para explicar la aversión o propensión al riesgo es su totalidad, por ello, es necesario contemplar la influencia cultural y las normas sociales vinculadas. Es un hecho irrevocable que las tendencias conductuales sexualmente dimórficas están sujetas a la interacción entre elementos inherentes y adquiridos.

Las diferencias entre sexos en conductas de riesgo se han estudiado ampliamente en contextos muy diversos, los deportes, finanzas, apuestas, relaciones sexuales y consumo de drogas son algunos de los más representativos. Tanto estudios observacionales como experimentales han mostrado que las conductas de riesgo son mayoritariamente efectuadas por hombres; sin embargo, se sabe poco sobre la manifestación de tales conductas en contextos cooperativos que involucren a ambos sexos, es decir, ¿qué ocurre cuando dos o más individuos interaccionan dentro de un escenario que supone disyuntivas e incertidumbre sobre los posibles desenlaces? ¿En qué sentido la propensión al riesgo de un individuo influye sobre la aversión al riesgo de otro? ¿Es posible que se exacerben ciertas conductas a consecuencia de la interacción entre estrategias contrastantes o más bien, se produce una forma de amortiguamiento?

El objetivo de la presente investigación fue i) evaluar las diferencias entre sexos en conductas de riesgo a partir de una tarea cooperativa realizada por diadas de un solo sexo o mixtas, así como ii) identificar si existen posibles efectos de modulación o amortiguamiento en las diadas mixtas (resultados intermedios). Por último, iii) determinamos en qué grupo de edad las diferencias conductuales asociadas al riesgo se manifiestan o se hacen notorias. Para explorar estas preguntas propusimos una metodología novedosa: *La tarea de construcción* *de torre* que fue aplicada en diadas de un solo sexo y en diadas mixtas en tres grupos de edad: 6, 12 y 18 años. La tarea consiste, en construir de manera conjunta, la torre más alta posible.

2. Metodología

Participantes

Se evaluaron 138 diadas divididas en tres grupos de edad y tres combinaciones de sexo: hombres, mujeres y mixtas. Los participantes de 6 y 12 años fueron reclutados de tres primarias públicas de la Ciudad de México, mientras que los de 18 años fueron reclutados de la Facultad de Ciencias de la UNAM.

Tarea de construcción de torre

Consiste en la construcción cooperativa de la torre más alta posible dentro de un intervalo de 10 minutos; para ello, se utilizaron bloques de madera de las siguientes dimensiones: 1.5 cm × 2.5 cm × 7.5 cm. La tarea es efectuada por dos participantes, cada uno recibe un conjunto de 54 bloques del mismo color (azul o rojo) y sólo pueden añadir bloques del color que les fue asignado.

Procedimiento

Se reclutaron estudiantes y se formaron diadas de manera aleatoria, posteriormente se llevaron al área experimental en donde recibieron las indicaciones pertinentes. Los participantes fueron instruidos de la siguiente manera: *(i) Tienen que construir, juntos, la torre más alta posible; es un trabajo en equipo. La tarea dura 10 minutos. (ii) Sólo pueden utilizar los bloques del color que les fue asignado. (iii) Si la torre se les cae pueden intentarlo* *cuantas veces quieran, siempre y cuando no haya terminado el tiempo. (iv) Si están satisfechos con su torre pueden acabar antes de tiempo.* No hubo restricciones, ni en las estrategias de construcción, ni en la comunicación entre participantes, ni en el número de re-intentos en caso de que la torre colapsara.

Codificación de descriptores conductuales

La codificación de la conducta se llevó a cabo mediante el software BORIS (Behavioral Observation Research Interactive Software), en el cual se estableció un etograma que incluye los descriptores conductuales que fueron interpretados como manifestaciones de riesgo, ya sea de propensión o aversión. Los descriptores que se emplearon para la determinación de las conductas de riesgo son los siguientes: altura máxima, ganancia de altura (cm) por bloque añadido, número de colapsos, latencia de inicio, latencia de reinicio, conclusión temprana (torre satisfactoria antes del tiempo límite), tasa de adición, proporción de piezas verticales.

Análisis estadístico.

Se realizaron análisis de varianza (ANOVAs) unifactoriales a través del software R para determinar si existe significancia estadística entre las tres combinaciones de sexo dentro de cada grupo de edad. Adicionalmente se efectuaron análisis equivalentes para determinar diferencias significativas entre grupos de edad pertenecientes a una sola combinación de sexo. Finalmente se llevaron a cabo comparaciones por pares mediante pruebas de Tukey (análisis post hoc).

3. Resultados

Aspectos generales

La construcción de torres fue una conducta generalizada durante la ejecución de la tarea y prácticamente todos los participantes colaboraron en el proceso de construcción. Los colapsos se produjeron de manera frecuente, el 68% de las pruebas tuvieron al menos un colapso. Las demoliciones y colapsos involuntarios fueron los tipos de colapso predominante. La comunicación verbal fue un aspecto extendido que ocurrió en todos los grupos de edad. El 41% de las diadas optaron por la conclusión temprana, lo cual sugiere que quedaron satisfechos con la torre construida antes de consumir el tiempo estipulado. Aunque las estrategias de construcción fueron diversas, los tipos predominantes de torre estuvieron conformados mayoritariamente por niveles de dos o tres bloques horizontales. Las piezas verticales fueron claramente minoritarias; se emplearon para construir arcos o estructuras semejantes.

Diferencias en combinaciones de sexo y grupos de edad

Encontramos diferencias estadísticamente significativas entre combinaciones de sexo pertenecientes a un solo grupo de edad: altura máxima (entre hombres y mixtas; entre hombres y mujeres) y ganancia de altura por bloque añadido (entre hombres y mujeres) en diadas universitarias; el número de colapsos difirió significativamente en diadas de 6 años (entre mujeres y mixtas). Entre grupos de edad pertenecientes a una sola combinación de sexos, encontramos diferencias significativas en la gran mayoría de los descriptores conductuales establecidos: altura máxima, ganancia de altura (cm) por bloque añadido, tasa de adición, proporción de piezas verticales, latencia de inicio y latencia de reinicio. La comparación entre edades fue útil para identificar posibles tendencias conductuales dentro de una determinada combinación de sexos.

4. Discusión

La presente investigación propone una metodología cuya finalidad es evaluar conductas de riesgo en un contexto social que involucra la participación y el consenso de dos individuos. Los objetivos primordiales del estudio fueron: i) la evaluación de diferencias entre sexos en la adopción de conductas de riesgo; ii) identificar posibles efectos de modulación en diadas mixtas; iii) e identificar en qué grupos de edad se manifiestan tales diferencias.

Ciertamente se encontraron diferencias significativas entre sexos, fundamentalmente en diadas universitarias. Aquellas conformadas por hombres construyeron torres más altas. Por otro lado, la ganancia de altura por adición fue superior en diadas de hombres. Considerando que, en el grupo de 18 años, estas fueron las únicas diferencias significativas entre sexos y que no hay evidencias sobre la superioridad de un sexo en lo que refiere a las aptitudes requeridas para desempeñarse en la tarea de construcción, podemos suponer que las diadas conformadas por hombres mostraron una mayor propensión al riesgo y que la interacción entre sexos (diadas mixtas) produjo un efecto de amortiguamiento: las diadas mixtas y las diadas de mujeres presentan valores similares, tanto en la mediana como en la dispersión. De acuerdo a lo anterior, la presencia de una mujer en la diada parece ejercer un efecto de modulación. Inicialmente se había determinado que los valores intermedios de diadas mixtas constituirían los casos de modulación de conductas de riesgo. A posteriori, se reformuló la determinación de los efectos de modulación ya que identificamos casos en que los resultados de diadas mixtas se asemejan en gran medida a uno de los grupos de diadas conformados por un solo sexo mientras que distan notablemente de las diadas del otro sexo. Finalmente, estos casos fueron contemplados como efectos de modulación, lo cual resulta aún más interesante pues se torna evidente la injerencia de un sexo sobre el otro.

En cuanto a las diferencias entre grupos de edad, la verticalidad, la tasa de adición, la altura máxima, la ganancia de altura (por bloque añadido) y las latencias fueron las más notables. Las diferencias estadísticamente significativas en dichos descriptores conductuales se relacionan con numerosos aspectos que se manifiestan durante la tarea de construcción: la disposición para cooperar, la coordinación, la planeación, la comunicación verbal, la percepción del riesgo (consecuencias) y la motricidad, son algunos de ellos. Estas diferencias muestran tendencias conductuales y/o cognitivas en a través del desarrollo.

De acuerdo a las manifestaciones conductuales derivadas de la tarea y a los resultados obtenidos, consideramos adecuado el uso de la Tarea de construcción de torre para la evaluación de conductas de riesgo en un contexto cooperativo. Dado que la mayoría de estudios sobre riesgo están basados en escenarios hipotéticos, autoreportes, encuestas u otras medidas indirectas, encontramos relevante que la tarea en cuestión contempla la posibilidad de efectuar conductas de riesgo que no comprometen la integridad física ni psicológica de los involucrados. Algunas características que denotan la versatilidad de la tarea son: la inclusión y evaluación de varios grupos de edad, la ejecución en ambientes o contextos variados (contexto independiente) y la facilidad de réplica. Al tratarse de una tarea recientemente desarrollada, es conveniente explorar y modificar las condiciones experimentales para robustecer los descriptores conductuales con la finalidad de aumentar su sensibilidad en la estimación del riesgo; también es posible implementar nuevos descriptores. La versatilidad de la prueba permite su adaptación a diversos escenarios que aborden preguntas sobre cooperación, competencia, capacidad inhibitoria, capacidades cognitivas y que incluyan arreglos de diadas con variantes de edad, de afiliación (e.g. in-group-out-group o parentesco) o con determinados perfiles psicopatológicos.

Conclusión

Aunque las conductas de riesgo y las decisiones subyacentes se han estudiado exhaustivamente, con especial énfasis en las diferencias entre sexos, poco se sabe sobre la manifestación del riesgo en escenarios que involucran la interacción y el consenso de las partes implicadas. Por ello, consideramos relevante la implementación de metodologías que permitan explorar conductas de riesgo en contextos cooperativos que simulen situaciones de la vida diaria.

Sex differences in risk-taking behavior on a cooperative task in children and young adults

Abstract

Risk-taking behavior has been studied in a variety of situations; however, risk-taking has been much less studied in cooperative contexts. Our aim in the present study was to assess sex differences in risk-taking behavior in children and young adults using a cooperative task. For this, male, female and mixed sex dyads were tested in a tower building task where participants had to pool their assigned wooden blocks in order to build the tallest tower possible within ten minutes. No evidence of sex differences was found in children dyads. However, among young adults, male dyads built taller towers and achieved larger height per block placed, whereas the presence of a female in the dyad resulted in a reduced tower height. This study highlights the importance of risk decisions made by consensus and introduces a novel method to evaluate risk behavior in a realistic social scenario.

Keywords:

Risk-taking behavior, sex differences, age differences, cooperation, dyadic interaction.

1. Introduction

Risk-taking entails the expectation of potential benefits under uncertainty, that is, the selection of the less certain, but potentially more rewarding option (Apicella, Carré, & Dreber, 2015). Risk-taking constitutes one of the most studied behavioral sex differences among humans. As shown by Byrnes, Miller, & Schafer (1999) in a meta-analysis of 150 observational and experimental studies, it is males who often show a stronger tendency to engage in risky behaviors. Consistent with this, more recent studies (Cobey, Laan, Stulp, Buunk, & Pollet, 2013; Pawlowski & Atwal, 2008) also support Byrnes's meta-analysis, showing a male tendency towards more risky behaviors. Nevertheless, there are exceptions reporting no differences between sexes (see review by Charness, Gneezy, & Imas, 2013) or reporting the opposite pattern, for instance in matrilineal societies (Gneezy, Leonard, & List, 2009).

While sex differences in risk-taking invite a biological explanation (Dreber, Apicella, & Eisenberg, 2009; Steinberg, 2008), it is clear that cultural and environmental aspects also play a role. Examples include effects of ethnicity, religion, socioeconomic status or education (Barsky, Juster, Kimball, & Shapiro, 1997; Donkers, Melenberg, & Van Soest, 2001; Fan & Xiao, 2005; Grable, 2000; Grable & Joo, 2004). Age may also be considered a variable of particular interest since this entails both inherent and acquired features. For instance, a raise in hormone levels represent an example of an inherent cause of changes in behavior with age (Apicella et al., 2008; Mehta, Welker, Zilioli, & Carré, 2015), whereas encouragement among male peers and consequent willingness to engage in risky situations may qualify as an instance of an acquired feature. This idea has been well expressed by Slovic (1966, p. 169): *"A prevalent belief in our culture is that men should, and do, take greater risks than women. For the child, a man's role is defined to a considerable extent in terms of courage"*. Thus, risk behavior may be best understood using a nature via nurture approach, or as the result of existing cultural pressures acting on inherent features (Booth & Nolen, 2012; Keltikangas-Järvinen, Räikkönen, Ekelund, & Peltonen, 2004).

Due to its general scope and complexity, risk-taking has been addressed from a wide range of disciplines including psychology, biology, medicine, economics and finances. Risk behavior permeates a large number of life scenarios: driving, gambling, unprotected sex, drug consumption, and even daily activities such as crossing busy roads (see meta-analysis by Byrnes et al., 1999) or adjusting arrival times at a bus stop in order to cut waiting time (Pawlowski & Atwal, 2008). Even though understanding risk entails some differences among the previously mentioned contexts, three basic elements are thought to be involved in all decisions concerning risk: options, outcomes, and uncertainties (Fischhoff & Kadvany, 2011). Thus, risk-taking entails the implementation of a goal-directed option which could result in more than one outcome of which at least one is undesirable (Furby & Beyth-Marom, 1992) or as Fischhoff & Kadvany (2011, p. 22) succinctly write "*Risks involve threats to outcomes that we value*".

Most studies addressing risk behavior have focused on situations relevant to adolescent and adult subjects. This is not surprising since measuring risk-taking in children can imply several confounding factors such as carelessness or lack of motor skills, and in some cases, studies are entirely based on economic decision-making tasks (for an example, see Paulsen, Platt, Huettel, & Brannon, 2011) rather than behavioral displays. One of the few studies that has evaluated risky behaviors in children is that of Ginsburg & Miller (1982), in which sex differences in risk taking were examined by observing 3 to 11 year-old children at four different situations at the San Antonio zoo: riding elephants, petting the burro, feeding animals and climbing the river embankment. In all situations significantly more boys than girls engaged in risk-taking behaviors. This result is consistent with Slovic's levers set up (1966), were children could pull levers to win M&M candy. Participants were allowed to continue playing the game, but they would lose all their cumulative payoff (candy), if they pulled the wrong lever (randomly assigned as the "disaster switch") in a further attempt. The contribution of these studies lies not only in making evident children's risk-taking behavior, but also in showing that differences between sexes emerge from very early ages.

In order to ensure the welfare and safety of participants when evaluating risk taking behavior, assessments are generally based on self-report (e.g. Pfefferbaum & Wood, 1994) and scales (e.g. Zuckerman, Kolin, Price, & Zoob, 1964; Zuckerman, 2007), hypothetical financial scenarios (e.g. Gneezy & Potters, 1997; Kahneman & Tversky, 1984) inferences drawn from datasets describing antisocial behavior and other factors leading to a sex disparity in mortality rates (e.g. criminal acts, car crashes, homicide, among others described in Moffitt, Caspi, Rutter, & Silva, 2001; Kruger & Nesse, 2004) or from observational studies (e.g. Pawlowski & Atwal, 2008).

In spite of numerous existing and elegant methodologies such as the ones previously mentioned, we set out to develop a behavioral ludic task to evaluate risk-taking: the Tower Building Task (TBT). The TBT involves using uniformly-size wooden blocks to build the tallest tower possible within a limited time. To do this, a pair of participants have to work together, each using only blocks of the color that was assigned to them. The number of blocks assigned to each participant is limited, and thus they need to cooperate to ensure a favorable outcome.

An important feature of the task is that it implements the possibility of a physically realistic negative outcome (the tower falling down) that is nevertheless without harmful consequences for the participants. This makes it especially useful for risk assessment in children as they are not directly testable in hazardous scenarios. It is important to distinguish hazard from risk. According to Kaplan and Garrick (1981) the former exists as a source, while the latter includes the likelihood of conversion of that source into actual delivery of loss, injury, or some form of damage. To illustrate this, we can think of a daily activity such as crossing the road – the road exists as a source, but the degree of risk depends on the choice of the pedestrian to cross either by jaywalking or at the zebra cross. Despite the physically harmless nature of the TBT, it fulfills the basic criteria of an operationally broad definition of risk consistent with the three fundamental elements or as Byrne (1999, p. 367) puts it "*…the implementation of options that could lead to negative consequences*". The TBT provides options that involve potential risk-taking within a safe environment, it is partially independent of linguistic ability, and can be applied in a wide range of age groups, each performing at their own level of skill.

Another important aspect of the TBT is its cooperative nature. The task used in the present study represents a simulation where members of a group influence each other's decision making and subsequent actions either towards or away from risky choices. Examples of scenarios where decisions are usually arrived at mutually include planning a trip among friends, having a baby or deciding whether to get a mortgage on a house. Clearly, in all cases a goal-directed option has to be implemented. Those involved might feel a degree of risk, as there is an investment (time, money or emotional strain) and a potential danger or loss (cost). Moreover, these situations often involve seeking consensus by the members of a group. As in such real-life situations, the TBT requires cooperative interaction operationally defined as the combined behavior of all participants is needed to attain the best possible outcome (Keller & Schoenfeld, 1950, p. 357-358).

1.1. Aims of the study

Our aim in the present study was to assess sex differences in risk-taking across development using a cooperative task. We expected males to engage more frequently in risky behaviors than females, but for this to be modified in mixed-sex dyads, resulting in scores intermediate to those for same-sex dyads of the two sexes tested separately. We also expected any such sex differences to become more apparent with increasing participant age.

2. Methods

2.1. Participants

We tested 138 dyads (276 participants) from three age-groups: six-year-old (6YO) and twelve-year-old children (12YO) from a public elementary school in Mexico City and university students (18YO) recruited on the main campus of the Universidad Nacional Autónoma de México (UNAM). We arranged participants into dyads according to one of three sex combinations: same-sex, (female [F] and male [M]) and mixed sex dyads (MIX). Each dyad participated in one test only. Dyad members, including university participants, were classmates, but were paired randomly by the experimenter based on their number in the attendance roll which was obtained beforehand. This recruitment procedure resulted in a slightly unequal number of dyads per group. A summary of the participants can be found in Table 1. A few (n=6) dyads were discarded due to experimental errors such as failure of equipment, or unclear instructions from experimenter (shown in parenthesis in Table 1).

Table 1
Description of sample

Age group	*Dyad*	*n (discarded)*	*Mean age (SD)*
6YO	F	14 *(1)*	6.32 (0.25)
	M	15	6.37 (0.3)
	MIX	15	6.37 (0.3)
12YO	F	15 *(2)*	11.31 (0.25)
	M	14 *(1)*	11.43 (0.47)
	MIX	14	11.47 (0.34)
18YO	F	17	18.71 (0.79)
	M	14 *(1)*	19.89 (1.53)
	MIX	14 *(1)*	18.86 (1.03)

2.2. Tower building task

The task consisted of having a single dyad build together the tallest tower possible using their assigned wooden blocks (Figure 1a). Building blocks (1.5 x 2.5 x 7.5 cm.) were adapted from the board game *Jenga* (Parker Brothers, Hasbro Inc, USA). Each member of the dyad received 54 blocks of a single color, either red or blue (Figure 1b). In order to keep the tower building processes cooperative, each participant was asked to only handle blocks of the color assigned to him or her. To eliminate the effect of small variations in the floor topography across trials, we instructed participants to build on a 50 x 50 cm board with a smooth melamine surface (Figure 1c). A large green hourglass was placed conspicuously beside the board to inform participants about the time remaining (Figure 1d). Dyads were also told they could keep attempting to build their tower until time on the hourglass had run out (10 minutes) or stop at any moment if satisfied with their result.

Figure 1. Experimental setup. Dyads had to attempt to build the tallest tower possible (a) using only blocks from the pile they were assigned (b) over a flat, uniform melamine surface (c) before time on the hourglass ran out (d).

Dyads received the following instructions verbatim: "*(i) The goal is to build together the tallest tower possible within 10 minutes; it is a team effort. (ii) You can only use the blocks of the color that was assigned to you, (iii) if the tower collapses, you can keep on building until time is up. Finally, (iv) if satisfied with the tower, you can stop building before the time finishes*". In Spanish: *(i) Tienen que construir, juntos, la torre más alta posible; es un trabajo en equipo. La prueba dura 10 minutos. (ii) Sólo pueden utilizar los bloques del color que les fue asignado. (iii) Si la torre se les cae pueden intentarlo cuantas veces quieran, siempre y cuando no haya terminado el tiempo. (iv) Si están satisfechos con su torre pueden acabar antes de tiempo.*

If participants asked about the building procedure (e.g. "*Can we put two blocks at the same time?*" or "*Can I talk to him/her?*") these were answered with standardized unspecific responses ("*You can build in any way you want*" and "*Yes, talking to each other is allowed*"). When participants infringed basic rules of the task, such as taking blocks from their partner or building a structure outside of the melamine surface, these were corrected by a brief verbal reminder from the experimenter.

When testing the youngest dyads, we provided a slower explanation using simpler terms, and made sure that they understood the instructions by asking them questions and having them explain the task in their own terms. We also took care to explain the purpose and ask about the function of the hourglass, which all participants seemed to understand successfully.

2.3. Procedure

Using an attendance roll, participants' names were called out loud and asked to approach the classroom door. After that, each dyad walked along with the experimenter towards the testing place where instructions were given to them. The task was performed by a single dyad and other participants were not allowed to watch the building procedure. Trials were conducted in either empty classrooms or in an unoccupied playground located a few steps from where the participants were recruited. All testing was done by the same male experimenter and all trials were conducted during school hours (9 to 14 hrs). Before the test, members of each dyad were asked their name and age. Afterwards, instructions were given and both participants took their place next to their assigned set of wooden blocks. All trials were filmed using a small video-camera (Handycam CX405, Sony Corporation, Tokyo, Japan) installed on a tripod about 3 meters from the participants. A whiteboard containing the trial and participants' information (date, number, age and sex combination) was filmed for a few seconds before telling participants they could start the task. After the task was concluded participants were further interviewed about their experience with the Jenga board game.

The recruitment processes and experimental procedure met bioethical requirements established by the Internal Review Board for Research with Human Subjects of the Instituto de Investigaciones Biomédicas, UNAM.

2.4. Behavioral coding

Video files of all trials were later analyzed using BORIS (Behavioral Observation Research Interactive Software; Friard & Gamba, 2016), where we coded the sequence of events within each trial encompassing block addition and collapses in time. These events were later transcribed into behaviors considered to be informative of participants' performance and risk-taking propensity, which are described in detail below.

2.4.1. Behavioral descriptors

Maximum height. Tallest tower achieved during the trial.

Height gain. Height gained per piece added (in cm).

Number of collapses. We distinguished three forms of collapse *a posteriori*, depending on the cause: a) involuntary, b) demolition and c) clumsiness. An involuntary collapse refers to the unwanted loss of tower height and was identified by meeting all following criteria: (i) a loss of at least a quarter of the length constructed, (ii) at least a quarter of the pieces that make up the tower must fall and (iii) the tower had to consist of at least 10 pieces. An involuntary collapse is not a risky behavior, but rather the negative consequence of the accumulation of risky choices leading to the loss of height. Therefore, it may be interpreted as a risk adoption descriptor. Demolitions occurred when dyad members, presumably unsatisfied with their performance, voluntarily collapsed the tower (either partially, by disassembling a part of it, or totally) in order to improve on their previous attempt. Since demolition behavior implies another attempt by discarding the built tower, it entails the implementation of an uncertain option; the main interpretation is that they do this with the purpose of improving performance. As long as uncertainty involves a potential loss it must be considered as a risk indicator (Kaplan & Garrick, 1981). Finally, collapses due to clumsiness resulted from events such as accidentally bumping the board or involuntarily striking the tower with any part of the body but in fact occurred in only 5% (n=7) of trials.

Latency to start. Time elapsed between the start of the test and the time when the first block was placed.

Latency to restart. Time elapsed after a collapse and the time participants started re-building a tower by placing the first new block. This could happen several times in the course of a single trial.

Early conclusion. The time elapsed between the moment participants placed the first block to the time they stopped building. Only dyads whose choice to stop construction happened 50 seconds before the allotted time were considered. An early conclusion suggests an unwillingness to engage in a new (uncertain) attempt, and thus it can be regarded as a form of risk avoidance.

Addition rate. Number of pieces added per unit of time. All pieces added were considered regardless of their placement on the board or as part of the structure. Moreover, all pieces were counted even if they did not contribute to an increase in height.

Color ratio. Proportion of blue to red pieces. We calculated this descriptor in order to assess the contribution of each member of the dyad.

Proportion of vertical pieces. Ratio of vertical to horizontal pieces throughout the entire building process. The larger the value, the greater the number of vertical pieces the tower contained in relation to the number of horizontal pieces. Since vertical positions result in a higher pay-off and reduced stability, they may be considered to be a risky choice. Horizontal additions could be placed in two ways, with one having a larger contact area but a lesser height gain.

Rehearsal. This occurred when one of the participants proposed the use of a particular arrangement of blocks, either verbally or by means of an example. The rehearsed structure did not compromise the focal tower; more than one rehearsal could take place during a trial.

2.5. Statistical analysis

Statistical analysis was done using R (R Core Team, 2017). We used one-way analyses of variance (ANOVAs) to evaluate the effect of sex on those behavioral variables which offered continuous values that resembled a normal distribution. Because of heterogeneity in variance between the age categories, we chose to compare performance between the sex combinations of each age group by means of one-way ANOVAs for each of the behavioral descriptors (except collapses). This was appropriate to the study design since our interest is in which age group sex differences significantly manifest, while allowing us to evaluate a possible modulatory effects in the mixed dyad group. Regarding age, we also conducted one-way ANOVAs considering age as categorical variable (independent variable). For the latency to re-start, which could occur more than once within a trial, we used repeated measures ANOVA. For the ANOVAs we report the F statistic, which is a ratio of the two variances compared, as well as the effect size (η^2), which we calculated using the function *eta_sq* from the package sjstats (Lüdecke, 2018). We also performed *post hoc* Tukey's Honest Significance Tests and report the confidence coefficient for the set. An exception was made for the total number of collapses and the number of towers, both of which are count data, and thus for these we fitted a generalized linear model (GLM) using a Poisson error and performed pairwise *post hoc* comparisons on the sex combination of the dyads in each age group using the *lsmeans* function of the lsmeans package (Lenth, 2016). For GLMs we report the estimate and its standard error, the z ratio (a similar measure to the t statistic) and the p value. All plots were done using the package *ggplot2* (Wickham, 2009). All tests were two tailed and significance was set to $p < 0.05$.

3. Results

3.1. General description of building behavior

Participants readily engaged in the construction of a tower. Almost all participants collaborated to build towers (91%, n=120), which varied in height (mean = 51.7, sd = 27.26, range = 4.5 – 171 cm; Fig. 2a) and clearly showed that blocks were added, at least for the two older groups, with the purpose of increasing tower height (Fig. 2b).

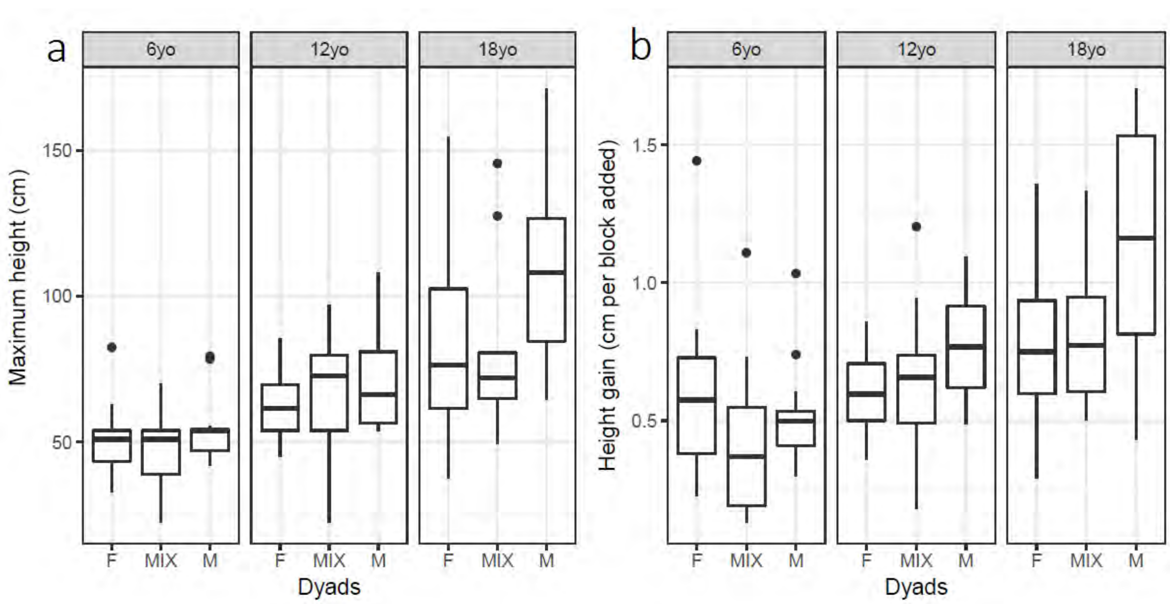


Figure 2. Boxplots of the maximum height (cm) and height gain (cm gained per block added) for each age group and sex combination. Horizontal line at the middle of the box marks median values, box limits represent the 1st and 3rd quartiles. Whiskers extend from the limits of the box to the smallest and largest value no further than 1.5 times the inter-quartile range. Outliers (dark circles) extend past this range.

Collapses happened relatively often (68% of trials had at least one collapse; Fig. 3a, see Fig. S1 of Supplementary Material for a breakdown of the frequency of each type of collapse), yet participants continued building, with a mean of 2.15 (sd = 1.26) towers per trial, which includes the first attempt (Fig. 3b).

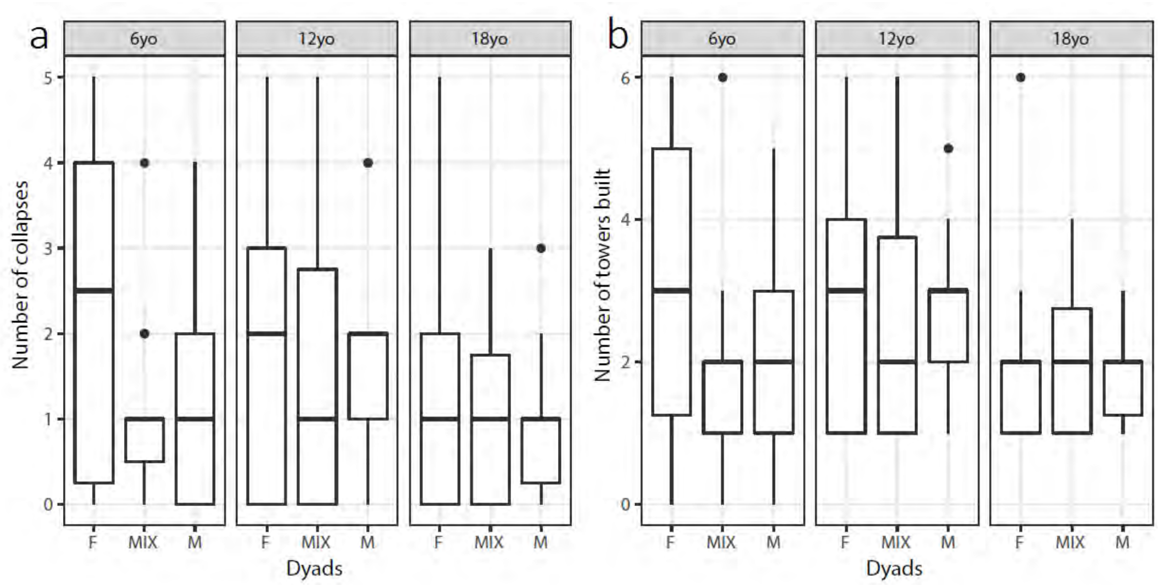


Figure 3. Boxplots of the total number of collapses and number of towers for each age group and sex combination. Horizontal line at the middle of the box marks median values, box limits represent the 1st and 3rd quartiles. Whiskers extend from the limits of the box to the smallest and largest value no further than 1.5 times the inter-quartile range. Outliers (dark circles) extend past this range.

All participants communicated verbally during the task, discussing turn taking, how to place the blocks or reaching consensus on when to stop. Some of these aspects were reflected in the time intervals. Participants started the task quite quickly (mean = 13 s, sd = 9.41; Fig. 4a) and when a collapse occurred, they rapidly went back to building (mean = 16.37 s, sd = 18.84; Fig. 4b, Table 2). Fewer than half (27%, n=36) of the dyads were satisfied with the tower they built and chose to end the activity before the allotted time.

When this happened they only chose to do so rather late in the test (on average after 438.23 s, sd = 134.89, or 73% of the allotted time had elapsed; Fig. 4c, Table 2).

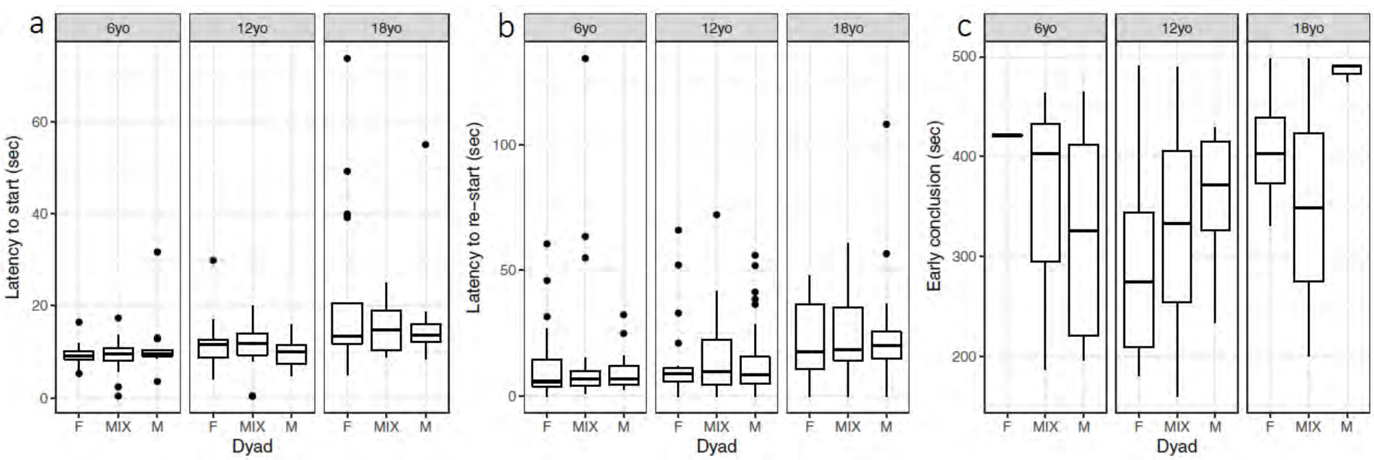


Figure 4. Boxplots of the latency to start (sec), latency to re-start (sec) and early conclusion times (sec) for each age group and sex combination. Horizontal line at the middle of the box marks median values, box limits represent the 1st and 3rd quartiles. Whiskers extend from the limits of the box to the smallest and largest value no further than 1.5 times the inter-quartile range. Outliers (dark circles) extend past this range.

Dyads added blocks at a mean rate of 0.38 blocks per second (sd = 0.12; Fig. 5a). Participation of both members of the dyad was evident from a mean color ratio of 1 (sd = 0.28; Fig. 5b) as well as from visual inspection of the construction sequences, which show a clear color alternation, even when turn taking did not always involve one block per turn. (see Fig. S2 of Supplementary Material for the full set of individual construction sequences). In general, building structures were diverse despite only 7% of the total pieces being placed vertically (Fig. 5c) and included structures such as arches (a horizontal piece over two vertical pieces) or vertical tunnels (alternating pairs of parallel horizontal pieces leaving a space in the middle, see Fig. S3 of Supplementary Material for the sample of building arrangements).

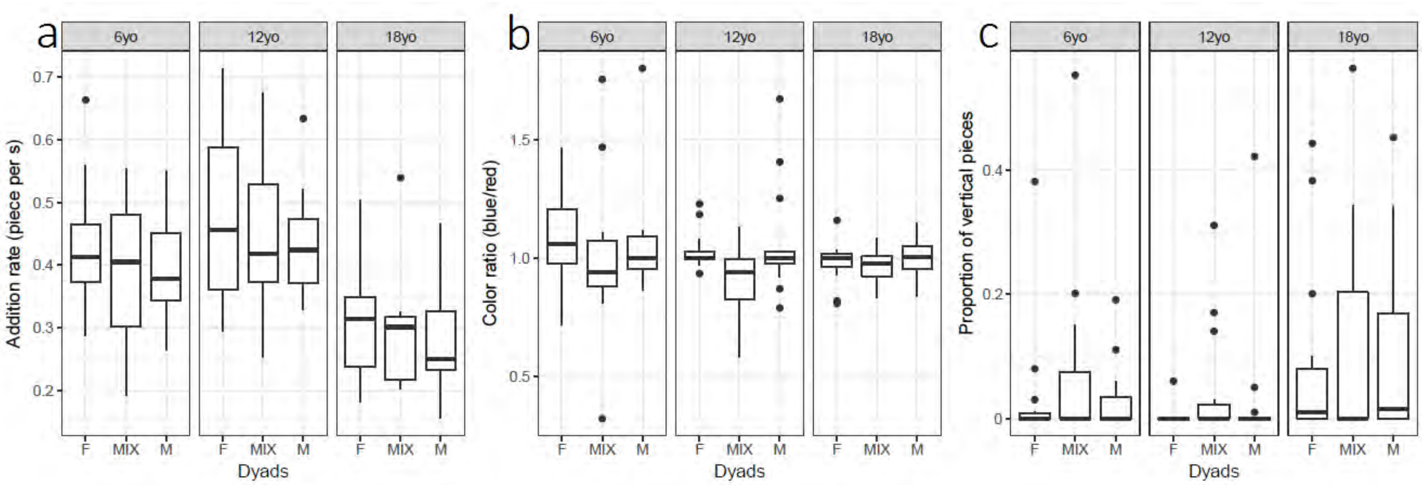


Figure 5. Boxplots of the addition rate (blocks added per sec), color ratio and proportion of vertical pieces for each age group and sex combination. Horizontal line at the middle of the box marks median values, box limits represent the 1st and 3rd quartiles. Whiskers extend from the limits of the box to the smallest and largest value no further than 1.5 times the inter-quartile range. Outliers (dark circles) extend past this range.

Regarding the current methodology, we encountered several unexpected results, particularly from the participants in the youngest age group, which are summarized in Table 2. As mentioned in the methods, we took particular care when explaining the instructions to these dyads and yet, a portion of the sample proceeded to build towers separately during the whole test (despite corrective comments from the experimenter), while several others deviated from the instructions at least at some point during the task and built separate towers. In some trials, particularly in the older groups, members of the dyad discussed and rehearsed structures before engaging in the task. Several dyads built towers composed of three adjacent horizontal pieces per level, with each different level alternating the direction of the blocks in the previous level, replicating the conventional Jenga structure. Finally, in some cases, and particularly in the youngest group, at least one of the subsequent building attempts after a collapse was almost identical to the previous tower.

Table 2
Construction peculiarities

			Trials where members of the dyad (n (%)):					
Age	*Dyad*	*n*	*Did not cooperate*	*Built separate towers*	*Discussed & rehearsed structures*	*Replicated the Jenga structure*	*Built identical tower after a collapse*	*Ended the task before time limit*
6yo	F	14	2 (14.29)	8 (57.14)	2 (14.29)	6 (42.86)	5 (0.36)	2 (14.29)
	M	15	2 (13.33)	6 (40)	1 (6.67)	6 (40)	4 (0.27)	3 (20)
	MIX	15	2 (13.33)	9 (60)	0 (0)	4 (26.67)	4 (0.27)	5 (33.33)
12yo	F	15	0 (0)	1 (6.67)	1 (6.67)	11 (73.33)	1 (0.07)	6 (40)
	M	14	0 (0)	2 (14.29)	5 (35.71)	8 (57.14)	2 (0.14)	6 (42.85)
	MIX	14	0 (0)	1 (7.14)	2 (14.29)	9 (64.29)	1 (0.07)	5 (35.71)
18yo	F	17	0 (0)	0 (0)	9 (52.94)	6 (35.29)	0 (0)	4 (23.52)
	M	14	0 (0)	1 (7.14)	8 (57.14)	3 (21.43)	0 (0)	2 (14.29)
	MIX	14	0 (0)	0 (0)	8 (57.14)	5 (35.71)	0 (0)	3 (21.42)

3.2. Performance according to age and sex

We found significant age differences within each sex combination for several behavioral descriptors: maximum height, height gain, addition rate, proportion of vertical pieces, latency to start and latency to restart (Table 3). Children engaged in frequent additions, whereas a lower addition rate was observed in adults. Older dyads achieved taller towers and had fewer involuntary collapses, while the rapid addition rate observed in the youngest group probably lead to a larger number of involuntary collapses.

Table 3

Addition rate (blocks per sec)					
			Est.	*C.I.*	*p*
M	*Age (F=10.74, p < 0.001, η^2=0.35)*	6YO-12YO	-0.14	-0.22 -- -0.06	*< 0.01*
		6YO-18YO	-0.11	-0.18 -- -0.03	*0.01*
F	*Age (F=8.58 p = 0.01, η^2=0.29)*	6YO-12YO	-0.16	-0.26 -- -0.06	*< 0.01*
		6YO-18YO	-0.11	-0.21 -- -0.01	*0.02*
MIX	*Age (F=6.6, p = 0.03, η^2=0.25)*	6YO-12YO	-0.15	-0.25 -- -0.05	*< 0.01*
		6YO-18YO	-0.1	-0.2 -- 0	*0.04*
Prop. of vertical pieces					
			Est.	*C.I.*	*p*
M	*Age (F=3.23, p = 0.05, η^2=0.35)*	6YO-18YO	0.11	0 -- 0.22	0.07
Height gain (cm per block added)					
			Est.	*C.I.*	*p*
M	*Age (F=12.19, p < 0.001, η^2=0.4)*	6YO-12YO	0.35	0.07 -- 0.64	0.01
		6YO-18YO	0.59	0.3 -- 0.89	*< 0.01*
MIX	*Age (F=8.45, p = 0.001, η^2=0.31)*	6YO-18YO	0.41	0.17 -- 0.66	*< 0.01*
Max height (cm)					
			Est.	*C.I.*	*p*
M	*Age (F=20.19, p < 0.001, η^2=0.39)*	6YO-12YO	52.27	31.55 -- 72.99	*< 0.01*
		12YO-18YO	35.29	15.38 -- 55.19	*< 0.01*
F	*Age (F=8.03, p = 0.001, η^2=0.1)*	6YO-18YO	30.19	11.53 -- 48.85	*< 0.01*
		12YO-18YO	18.2	0.26 -- 36.14	*0.05*
MIX	*Age (F=8.94, p = 0.001, η^2=0.31)*	6YO-12YO	19.78	0.89 -- 38.67	*0.04*
		6YO-18YO	32.57	13.68 -- 51.46	*< 0.01*
Latency to start (sec)					
			Est.	*C.I.*	*p*
M	*Age (F=3.05, p = 0.05, η^2=0.13)*	12YO-18YO	6.69	-0.038 -- 13.77	0.07
F	*Age (F=4.43, p = 0.02, η^2=0.17)*	6YO-18YO	11.87	1.45 -- 22.28	*0.02*
MIX	*Age (F=4.99, p = 0.01, η^2=0.20)*	6YO-18YO	6.06	1.34 -- 10.78	*0.01*

Latency to re-start (sec)					
			Est.	*C.I.*	*p*
M	*Age (F=4.26, p = 0.02, η^2=0.13)*	6YO-18YO	17.78	3.11 -- 32.45	*0.01*
F	*Age (F=3.26, p = 0.04, η^2=0.17)*	6YO-18YO	10.26	0.24 -- 20.29	*0.04*

When comparing the sex combinations within each age group, significant differences were found in the following behavioral descriptors: maximum height, height gain, number of collapses and color ratio (Table 4; see Tables S1 to S4 of Supplementary Material for the complete set of *post hoc* comparisons). Significant differences for the youngest group were found for the number of collapses between female and mixed dyads, while 18YO participants differed significantly in maximum height, between male and female, and male and mixed dyads. With regard to height gain, significant differences were found only between male and female 18YO dyads.

Table 4

Max height (cm)					
			Est.	*C.I.*	*p*
18YOs	*Sex (F=4.15, p = 0.02, η^2=0.16)*	MIX-F	-2.39	-27.51 -- 22.72	0.97
		M-F	25.39	0.28 -- 50.51	*0.05*
		M-MIX	27.79	1.49 -- 54.09	*0.04*
Height gain (cm per block added)					
			Est.	*C.I.*	*p*
18YOs	*Sex (F=4.01, p = 0.03, η^2=0.17)*	MIX-F	0.05	-0.25 -- 0.36	0.91
		M-F	0.34	0.03 -- 0.64	*0.03*
		M-MIX	0.29	-0.03 -- 0.61	0.09
Number of collapses					
		Est. (Std. Error)		*z*	*p*
6YOs	MIX-F	0.79 (0.3)		2.6	*0.03*
	M-F	0.52 (0.28)		1.87	0.15
	M-MIX	-0.27 (0.33)		-0.82	0.69
Color ratio					
			Est.	*C.I.*	*p*
12YOs	*Sex (F=3.8, p = 0.03, η^2=0.16)*	MIX-F	-0.12	-0.27 -- 0.02	0.12
		M-F	0.04	-0.11 -- 0.19	0.79
		M-MIX	0.16	0.01 -- 0.31	*0.03*

4. Discussion

4.1. On our aims

The current work examines risk-taking behavior on a cooperative task in children and young adults. We aimed to (i) identify sex differences on the behavioral descriptors, (ii) evaluate whether modulation occurred in mixed dyads and (iii) determine in which age group sex differences in risk-taking become more noticeable. The study yielded findings consistent with previous studies: behavioral sex differences associated with risk were found in 18YO dyads, were males showed a greater risk propensity, providing additional support for a widely accepted finding (Byrnes et al., 1999). Regarding our second aim, scores suggesting a modulatory effect were found in mixed dyads (e.g. number collapses in 6YOs and maximum tower height in 18YOs) and while it can be speculative to make claims about how one sex influences the other, we can provide some inferences based on the similarity of scores between mixed dyads and those of the same sex dyads they most resemble. Concerning when sex differences in risk-taking behavior become more evident, we found that these manifest mainly between the dyads of the oldest age group. Below, we discuss the findings pertaining to each of these aims in detail.

4.2. On sex differences

Male university dyads achieved larger maximum heights than female and mixed dyads, and yet, they did not differ in terms of addition rate, proportion of vertical pieces or number of collapses. Male 18YO dyads also scored higher than female dyads in height gain. Despite the fact that the distinction between risk taking and skillfulness may be hard to disentangle within some risk domains (e.g. recreational domain, which includes sports and outdoor activities. See Weber, Blais, & Betz [2002] for a categorization of these domains), the current results give us no reasons to think males would outperform females in the TBT due to improved skills. Moreover, mixed dyads produced towers of similar heights and had similar height gains to female dyads, suggesting males could have been influenced to some extent by their female peers. In accordance with this, risky choices were mostly represented by male dyads. This result contributes to a large body of knowledge claiming a sex disparity in relation to risk-proneness, mainly during adolescence and early adulthood (Kruger & Nesse, 2004).

4.3. On the modulatory effect of sex

We expected mixed dyads to yield intermediate scores (i.e. having same sex dyads at the opposite sides of the distribution and mixed dyads in the middle) as evidence of modulation. However, our results did not entirely support this prediction (but see early conclusion in 12YO where medians follow expected pattern). Thus, we settled to interpret modulation as cases were mixed dyads resemble those of one of the single sex dyads while being statistically different from the other same sex dyad. This was evident in mixed 18YO dyads, where interaction between sexes significantly reduced the maximum height as compared to those of male dyads. These results suggest that male risk propensity was lessened when interacting with females; we can infer males conformed to female conservative strategies that led to safer outcomes at the expense of a greater payoff. Risk-taking behaviors as cues in mate choice represent an extensive topic of research. Wilke and colleagues (2006) claim that risk proneness in particular domains such as social and recreational constitutes an attractive feature for both sexes. Additionally, anthropologists argue that male hunting represents a competitive display (Hawkes & Bird, 2002) that can enhance social status and increase mating success (Smith, Bird, & Bird, 2003). However, such studies have traditionally focused on the male displays of risky behaviors, while in the current task, the decisions are reached, at least in part, through a consensus. In such an agreement it appears that the "conservative" branch, mostly represented by females, has the upper hand.

4.4. On risk, performance and age

4.4.1. Skill

As expected, age groups differed significantly on several aspects. We found 18YO dyads achieved greater heights with lower addition rates and with fewer collapses. Given the large age differences between groups, these are likely the result of the confluence of maturation of both physical and cognitive attributes. For instance, fine motor skills involving object manipulation improve as age increases (Morris, Williams, Atwater, & Wilmore, 1982). A measure reflecting this was the large number of collapses in the younger dyads (e.g. collapses due to clumsiness were almost exclusive to the 6YO age group). Also, older participants have improved communication skills (Burleson, 1982) which are reflected by more efficiently conveying the desired building strategy, and consequently showing a heightened capacity to reach a consensus. Besides communication skills, joint attention and other social-cognitive skills that encourage cooperation, emerge early in development and continue improving as age increases (Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998). These make it easier for older dyads to coordinate their efforts towards a common goal.

4.4.2. Awareness of consequences

The more careful building approach observed in older dyads suggest a deeper awareness of consequences, which also tends to increase with age. The heightened preference for risky choices in children (even when scenarios become riskier) could reflect their unawareness of consequences, perhaps potential outcomes are not being weighted objectively, for instance: in a child-friendly economic decision-making task, children seem to weigh the winning probability-by-outcome value more heavily than the losing probability (Paulsen et al., 2011). This hypothesis may explain their higher risk propensity. In the current context, we could safely assume 18YO dyads better understand how a collapse entails a loss of time that consequently leaves less time for a further attempt, or how the value of the outcome is partially related to the time spent in the procedure, as well as the time available for a next attempt. Further evidence of an apparent lack of awareness of consequences was suggested by the fact that several of the younger dyads engaged in repeating the same structure (and achieving the same tower height) after a collapse, as if expecting an improvement.

Even though children report to understand similar hazards as adults and the elderly, seemingly as a consequence of learning from others' experience (Hermand, Mullet, &

Rompteaux, 1999), their appraisal of risk tends to be inaccurate (Little & Wyver, 2010). Children are generally less likely to fathom possible outcomes of a given action, since it may be their first time experiencing a particular situation (e.g. getting an electric shock from handling an electric socket). They simply may not be able to piece together an event and its consequence or they may lack the capacity to translate a negative outcome into a loss or damage (e.g. a collapse entails not only the loss of a tower but the loss of time spent on it as well as the shortening of the residual time for a next attempt). Later in life, appraisal of risk is influenced to some extent by higher levels of rewarding stimulation that are associated with novelty and sensation seeking, which increase dramatically at puberty (Kelley, Schochet, & Landry, 2004; Steinberg, 2004), especially among males (Cross, Cyrenne, & Brown, 2013). Certainly this behavioral sex disparity must be explained by numerous factors involving both: acquired (environmental-cultural) and inherent (biological bases) features.

4.4.3. Unwillingness to cooperate

In the current task, the best results could only be achieved by pooling resources together, which means that a cooperative interaction is required. However, a particular result arose as some of the younger dyads either did not cooperate or built separate towers at some point during the task. We think this was unlikely an issue of the task's instructions, as the experimenter was careful to quiz participants about the goal and restrictions in order to ensure their comprehension of the task. When not engaged in jointly constructing a tower, these participants built towers on their own, achieving considerable height and implicitly suggesting they understood the main goal of the task, therefore discarding attentional issues or a preference for play. Another explanation may be related to a reluctance to cooperate with their randomly assigned peer, either due to shyness or a feeling that they may underperform and be blamed for a collapse. Most of the methodologies testing cooperation in young children usually operationalize cooperation into a single decision (e.g. pull a rope or a lever simultaneously [Hamann, Warneken, Greenberg, & Tomasello, 2011]). The current task involves a highly inter-knitted cooperative scenario, were participants can gradually and repeatedly contribute to the common goal (e.g. by adding a block) as well as subtract from it (e.g. by producing a collapse). Thus, some dyad members might have preferred to engage in their own towers, either not understanding fully the benefits of cooperation or preferring to ignore them by taking their chances on their own level of skill.

Another aspect indirectly associated to cooperative behavior is the number of collapses in the youngest group. Unexpectedly, 6YO female dyads had a larger number of collapses than mixed dyads. This contradicts findings indicating that at this age, girls seem to outperform boys in terms of fine motor skills, such as balance (Morris et al., 1982), although upon closer inspection these sex differences seem to be controversial (Venetsanou & Kambas, 2016). The fact that this difference exists suggests it might be a social, rather than a motor issue. Indeed, children these ages form friendship bonds almost exclusively with same-sex peers (Sippola, Bukowski, & Noll, 1997). Thus, at this developmental stage, interaction might have been less spontaneous due to the awkward situation we created by having a mixed dyad condition, in which fewer towers were built. Supporting this idea, we also found that the 6YO mixed dyads built separate towers more often and none engaged in rehearsal behavior during the entire trial. In this sense, a familiar situation like having a same-sex peer could promote risky choices (consequently, leading to a larger number of collapses) and may also encourage collaboration.

4.5. On the tower building task

Regarding the application of a novel methodology, we encountered some unexpected findings among participants' attitudes towards tower building. First, we observed a large number of monotonous Jenga structures performed by twelve year-old dyads, which could reflect (i) their experience with the board game (unlikely, as familiarity was reported across participants of all age groups), (ii) an over-adherence to rules and goals promoted by a homogenizing education system or (iii) apathy towards the test. An explanation that involves risk assessment however, is that the traditional Jenga structure provides a considerable payoff, an above-the-mean height, while being relatively easy and safe to build. In fact, a couple of dyads added one or two vertical pieces on top of the Jenga structure, cleverly surpassing the mean while keeping the risk of collapse low.

We consider the TBT involves some advantages, such as the possibility of testing risk-taking across different age groups based on behaviorally innocuous conditions rather than self-report or indirect measures. Moreover, due to the availability and low cost of the employed material, the TBT can be easily replicated in a wide range of contexts. In spite of these favorable features, there are still a few aspects that could be improved in future testing sessions, particularly those pertaining the behavioral descriptors chosen (e.g. early conclusion) or to the materials employed (e.g. number of blocks). If the task were to end after the first collapse, early conclusion could be a more robust risk aversion indicator - the earlier dyads opted to conclude the task, the lower the risk (or it also could be interpreted as a risk-less choice for those pairs who chose to end versus those that did not). Whether the building process came to a halt because participants were satisfied with the tower height, ran out of blocks, or both, the choice to stop building despite having enough time to rebuild represents a risk averse choice. Concerning the availability of building material, when all blocks were used, dyads willing to increase height could either (i) demolish the current tower so as to engage in a new attempt or (ii) incur in deletions, that is, taking a block from the current structure with the purpose of making a further addition. Curiously, demolitions are arguably the opposite of the early conclusion, as they entail the implementation of an uncertain option. Deletions were rarely observed during trials, probably as it was ultimately perceived as a poor choice, involving a chance of collapse while representing a minor height improvement. Taken together, these risk averse (e.g. early conclusion) and risk prone (e.g. demolitions) indicators suggest that the TBT integrates some notion of *expected values*: the multiplication of the outcomes by their probabilities (Fischhoff & Kadvany, 2011). Another aspect that reflects a notion of expected values is the dimension of the Jenga blocks e.g. the vertical positon payoff implies a larger gain with a lesser contact area. While verticality suggests a riskier strategy, the expected value equals that of horizontal pieces (See expected values for the three types of block positions in the Supplementary Material). A concrete example of decisions involving expected values can be seen in people choosing 10 dollars (certain option) over a 50% chance to get 20 dollars (the uncertain and consequently the risky option). Both options have the same expected value (Kahneman &

Tversky, 1979). It is important to remember, however, that in the current task, the stability of each addition also depends on the position of previous additions.

The TBT is highly versatile and with small modifications or different experimental designs it can be used to address a wide range of risk scenarios, for instance those involving the interaction in dyads with a large age disparity, more constrained setups such as those where the task ends after the first collapse, thus making the outcome more valuable, or removing time limits to promote riskier construction attempts, etc. The task could also prove useful in evaluation of clinical contexts involving extreme and unfavorable risk preferences or prosocial behavioral issues.

4.6. Conclusion

From daily and harmless to rare and transcendental decisions, risk-taking constitutes a ubiquitous aspect of human behavior. Even though this topic has been extensively studied, further investigation can bring a better understanding of the interaction taking place in cooperative contexts, especially where decisions arise from the interplay between sexes. Our findings suggest that the TBT can provide a behavioral evaluation of risk while contemplating variables akin to those of real-life social scenarios.

Acknowledgments

We would like to thank the staff and students of Tlamatini, Ejercito Nacional and José Jesús Aguilera Palomino elementary schools for their participation in this study as well as Kevin Navarrete for assistance in the experimental sessions. We are grateful to Carlos Cordero, Arturo Becerra and Lynna Kiere for reviewing and offering comments of this manuscript. Finally, this work was supported by grants from DGAPA-PAPIIT-IA202617.

References

Apicella, C. L., Carré, J. M., & Dreber, A. (2015). Testosterone and economic risk taking: A review. *Adaptive Human Behavior and Physiology*, *1*(3), 358-385.

Apicella, C., Dreber, A., Campbell, B., Gray, P., Hoffman, M., & Little, A. (2008). Testosterone and financial risk preferences. *Evolution and Human Behavior*, *29*(6), 384–390. doi:10.1016/j.evolhumbehav.2008.07.001

Barsky, R. B., Juster, F. T., Kimball, M. S., & Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: an experimental approach in the health and retirement study. *The Quarterly Journal of Economics*, *112*(2), 537–579. doi:10.1162/003355397555280

Booth, A. L., & Nolen, P. (2012). Gender differences in risk behaviour: does nurture matter? *The Economic Journal*, *122*(558), F56–F78. doi:10.1111/j.1468-0297.2011.02480.x

Burleson, B. R. (1982). The development of comforting communication skills in childhood and adolescence. *Child Development*, *53*(6), 1578–1588. doi:10.2307/1130086

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: a meta-analysis. *Psychological Bulletin*, *125*(3), 367–383. doi:10.1037/0033-2909.125.3.367

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4), i-74.

doi:10.2307/1166214

Charness, G., Gneezy, U., & Imas, A. (2013). Experimental methods: eliciting risk preferences. *Journal of Economic Behavior & Organization*, *87*, 43–51. doi:10.1016/j.jebo.2012.12.023

Cobey, K. D., Laan, F., Stulp, G., Buunk, A. P., & Pollet, T. V. (2013). Sex differences in risk taking behavior among Dutch cyclists. *Evolutionary Psychology*, *11*(2), 350–364. doi:10.1177/147470491301100206

Cross, C. P., Cyrenne, D. L. M., & Brown, G. R. (2013). Sex differences in sensation-seeking: a meta-analysis. *Scientific Reports*, *3*, 2486–2491. doi:10.1038/srep02486

Donkers, B., Melenberg, B., & Van Soest, A. (2001). Estimating risk attitudes using lotteries: a large sample approach. *Journal of Risk and Uncertainty, 22*(2), 165–195.

Dreber, A., Apicella, C. L., & Eisenberg, D. T. A. (2009). The 7R polymorphism in the dopamine receptor D4 gene (DRD4) is associated with financial risk taking in men. *Evolution and Human Behavior, 30*(2)*, 85-92.*

Fan, J. X., & Xiao, J. J. (2005). A cross-cultural study in risk tolerance: comparing Chinese and Americans. *Journal of Personal Finance, 5*(3), 54-75.

Fischhoff, B., & Kadvany, J. (2011). *Risk: A Very Short Introduction* (1st ed). New York, NY, USA: Oxford University Press.

Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution, 7*(11), 1325-1330.

Furby, L., & Beyth-Marom, R. (1992). Risk taking in adolescence: a decision-making perspective. *Developmental Review*, *12*(1), 1-44.

Ginsburg, H. J., & Miller, S. M. (1982). Sex differences in children's risk-taking behavior. *Child Development*, *53*(2), 426. doi:10.2307/1128985

Gneezy, U., Leonard, K. L., & List, J. A. (2009). Gender differences in competition: evidence from a matrilineal and a patriarchal society. *Econometrica, 77*(5), 1637–1664.

Gneezy, U., & Potters, J. (1997). An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics*, *112*(2), 631–645.

Grable, J. E. (2000). Financial risk tolerance and additional factors that affect risk taking in everyday money matters. *Journal of Business and Psychology, 14*(4), 625–630.

Grable, J., & Joo, S.H. (2004). Environmental and biopsychosocial factors associated with financial risk tolerance. *Journal of Financial Counseling and Planning, 15*(1), 73-82.

Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature*, *476*(7360), 328–331. doi:10.1038/nature10278

Hawkes, K., & Bird, R. B. (2002). Showing off, handicap signaling, and the evolution of men's work. *Evolutionary Anthropology: Issues, News, and Reviews, 11*(2), 58–67.

Hermand, D., Mullet, E., & Rompteaux, L. (1999). Societal risk perception among children, adolescents, adults, and elderly people. *Journal of Adult Development, 2*(6), 137–143.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica, 47*, 263–292.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*(4), 341–350.

Kaplan, S., & Garrick, B. J. (1981). On the quantitative definition of risk. *Risk Analysis, 1*(1), 11–27.

Keller, F. S., & Schoenfeld, W. N. (1950). *Principles of Psychology: A Systematic text in the Science of Behavior* (pp. 357–358). New York, NY, USA: Appleton-Century-Crofts.

Kelley, A. E., Schochet, T., & Landry, C. F. (2004). Risk taking and novelty seeking in adolescence: introduction to part I. *Annals of the New York Academy of Sciences*, *1021*, 27–32. doi:10.1196/annals.1308.003

Keltikangas-Järvinen, L., Räikkönen, K., Ekelund, J., & Peltonen, L. (2004). Nature and nurture in novelty seeking. *Molecular Psychiatry*, *9*(3), 308–311. doi:10.1038/sj.mp.4001433

Kruger, D. J., & Nesse, R. M. (2004). Sexual selection and the male: female mortality ratio. *Evolutionary Psychology, 2*(1), 147470490400200130.

Lenth, R.V. (2016). Least-squares means: the R package lsmeans. *Journal of Statistical Software, 69*(1), 1-33. doi:10.18637/jss.v069.i01

Little, H., & Wyver, S. (2010). Individual differences in children's risk perception and appraisals in outdoor play environments. *International Journal of Early Years Education, 18*(4), 297–313.

Lüdecke, D. (2018). sjstats: Statistical Functions for Regression Models. R package version 0.14.1. Retrieved from https://CRAN.R-project.org/package=sjstats

Mehta, P. H., Welker, K. M., Zilioli, S., & Carré, J. M. (2015). Testosterone and cortisol jointly modulate risk-taking. *Psychoneuroendocrinology*, *56*, 88–99. doi:10.1016/j.psyneuen.2015.02.023

Moffitt, T. E., Caspi, A., Rutter, M., & Silva, P. A. (2001). *Sex Differences in Antisocial Behaviour: Conduct Disorder, Delinquency, and Violence in the Dunedin Longitudinal Study*. Cambridge, UK: *Cambridge University Press.*

Morris, A. M., Williams, J. M., Atwater, A. E., & Wilmore, J. H. (1982). Age and sex differences in motor performance of 3 through 6 year old children. *Research Quarterly for Exercise and Sport*, *53*(3), 214–221. doi:10.1080/02701367.1982.10609342

Paulsen, D. J., Platt, M. L., Huettel, S. A., & Brannon, E. M. (2011). Decision-making under risk in children, adolescents, and young adults. *Frontiers in Psychology*, *2*, 72. doi:10.3389/fpsyg.2011.00072

Pawlowski, B., & Atwal, R. (2008). Sex differences in everyday risk-taking behavior in humans. *Evolutionary Psychology, 6*(1), 29-42.

Pfefferbaum, B., & Wood, P. B. (1994). Self-report study of impulsive and delinquent behavior in college students. *The Journal of Adolescent Health*, *15*(4), 295–302. doi:10.1016/1054-139X(94)90602-5

R Core Team (2017). *R: A Language and Environment For Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Sippola, L. K., Bukowski, W. M., & Noll, R. B. (1997). Dimensions of Liking and Disliking Underlying the Same-Sex Preference in Childhood and Early Adolescence. *Merrill-Palmer Quarterly (1982-)*, 591-609.

Slovic, P. (1966). Risk-taking in children: age and sex differences. *Child Development*, *37*(1), 169. doi:10.2307/1126437

Smith, E. A., Bird, R. B., & Bird, D. W. (2003). The benefits of costly signaling: Meriam turtle hunters. *Behavioral Ecology, 14*(1), 116–126.

Steinberg, L. (2004). Risk taking in adolescence: what changes, and why? *Annals of the New York Academy of Sciences*, *1021*, 51–58. doi:10.1196/annals.1308.005

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review, 28*(1), 78–106. doi:10.1016/j.dr.2007.08.002

Venetsanou, F., & Kambas, A. (2016). Motor proficiency in young children. *SAGE Open*, *6*(1), 1-10. doi:10.1177/2158244015626226

Weber, E. U., Blais, A. R., & Betz, N. E. (2002). A domain- specific risk- attitude scale: measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making, 15*(4), 263–290.

Wickham, H. (2009) *ggplot2: elegant graphics for data analysis*. New York, NY, USA: Springer-Verlag.

Wilke, A., Hutchinson, J. M. C., & Todd, P. M. (2006). Is risk taking used as a cue in mate choice? *Evolutionary Psychology, 4*(1), 367-393. doi/abs/10.1177/147470490600400130

Zuckerman, M. (2007). The sensation seeking scale V (SSS-V): still reliable and valid. *Personality and Individual Differences*, *43*(5), 1303–1305. doi:10.1016/j.paid.2007.03.021

Zuckerman, M., Kolin, E. A., Price, L., & Zoob, I. (1964). Development of a sensation-seeking scale. *Journal of Consulting Psychology*, *28*(6), 477.

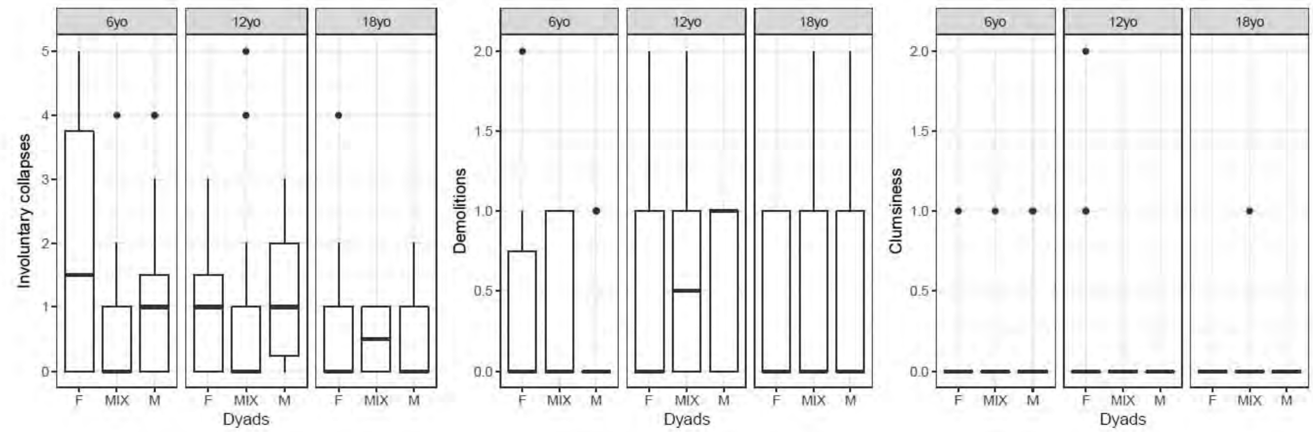
Supplementary Material
Figure Legends

Fig. S1. Breakdown of collapse types by sex and age. Boxplots of the number of involuntary collapses, demolitions and collapses due to clumsiness. Horizontal line at the middle of the box marks median values, box limits represent the 1^{st} and 3^{rd} quartiles. Whiskers extend from the limits of the box to the smallest and largest value no further than 1.5 times the inter-quartile range. Outliers (dark circles) extend past this range.

Fig. S2. Complete set of construction sequences. Y axis represents height achieved (cm) and X axis represents time elapsed (sec). Horizontal line on top of plot shows sequence of additions: each color represents additions made by one member of the dyad, blue additions are above the line whereas red are below. Short lines represent horizontal additions and long lines are for vertical additions. Solid black circles mark the time of a collapse. Plots are shown for every dyad.

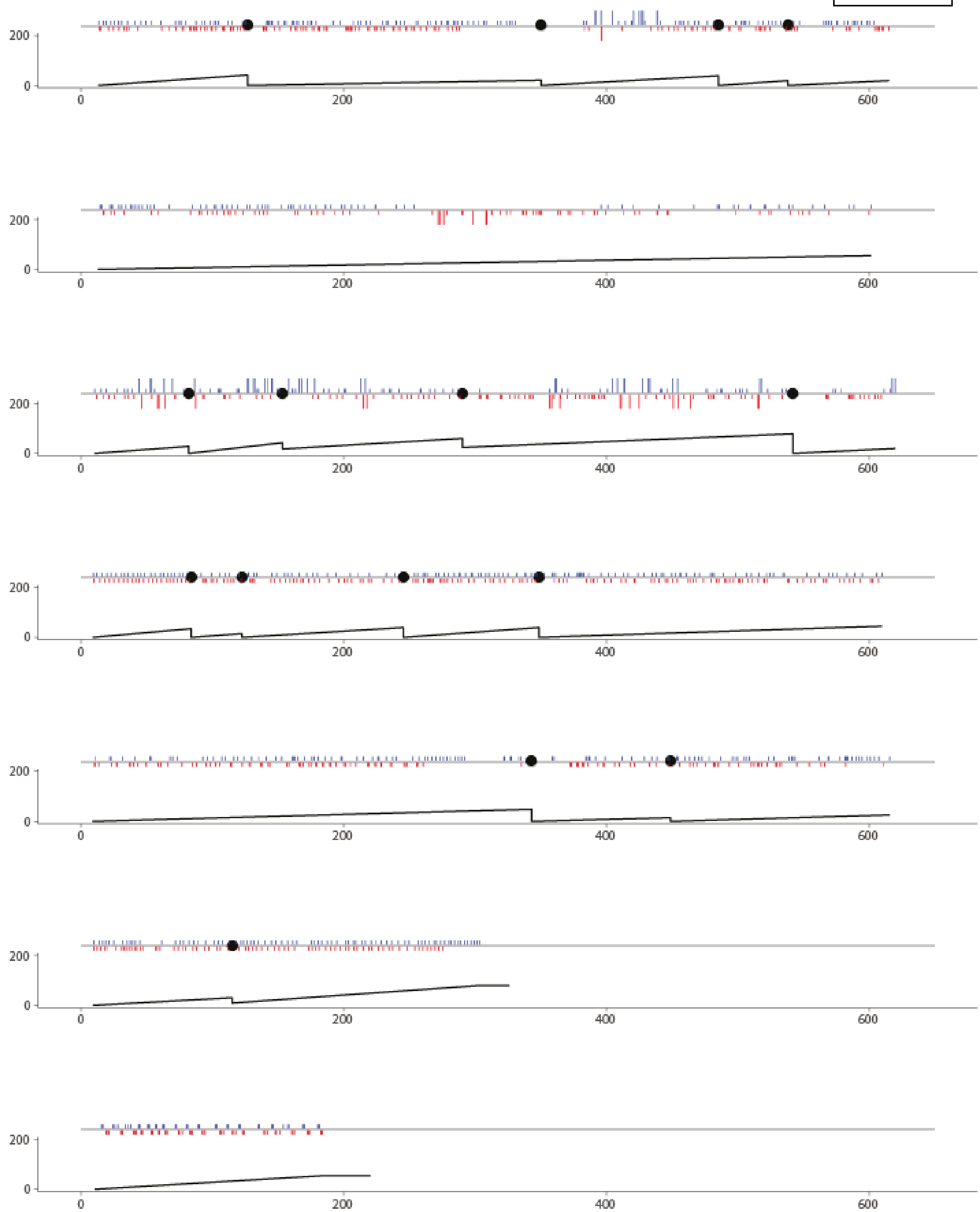
Fig. S3. Sample of building arrangements. a) Tower made from adjacent structures, b) replication of Jenga structure, c) vertical tunnels, d) arches.

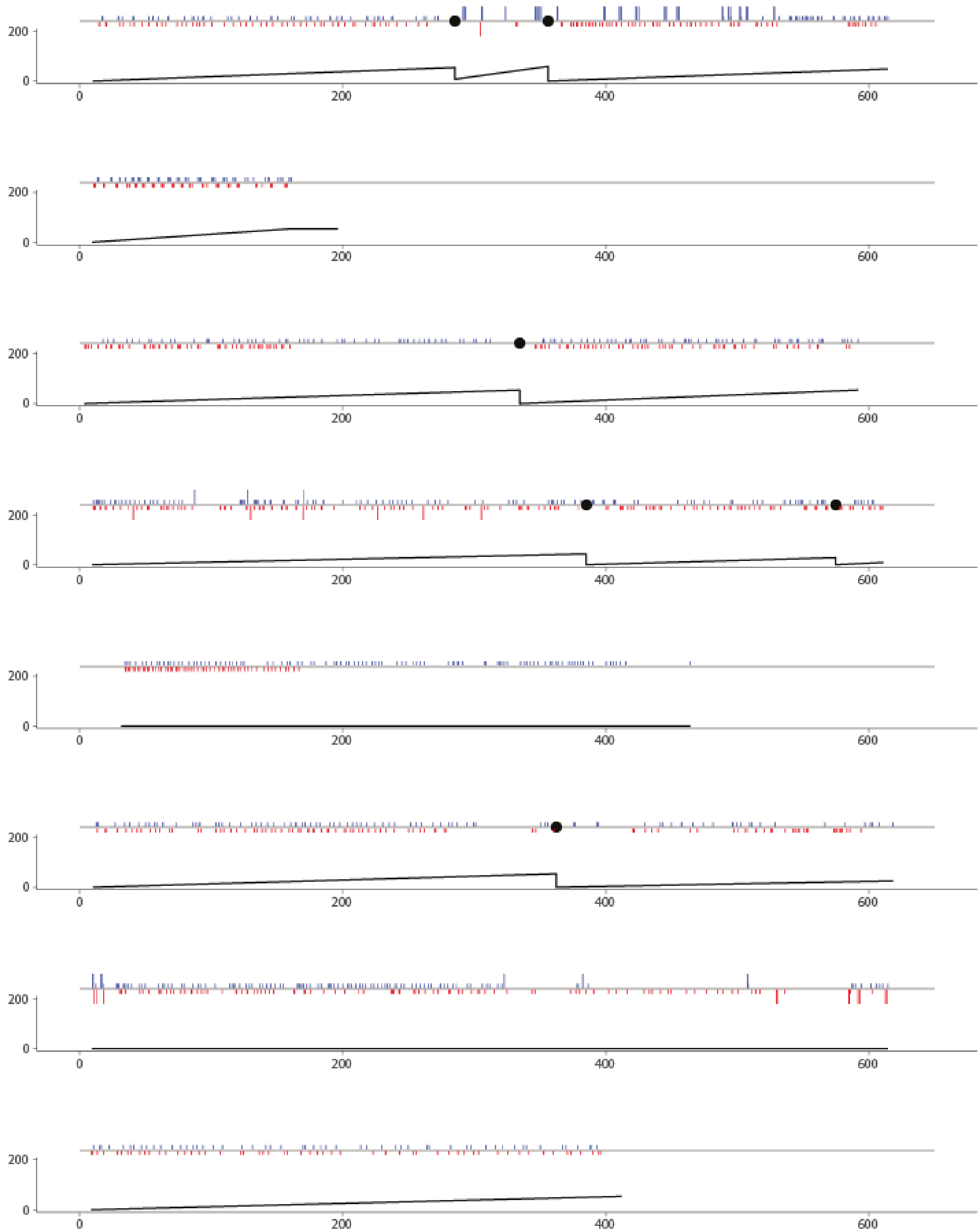
Fig. S1


6yo
12yo
18yo
Involuntary collapses
Dyads
F
MIX
M
Demolitions
Clumsiness

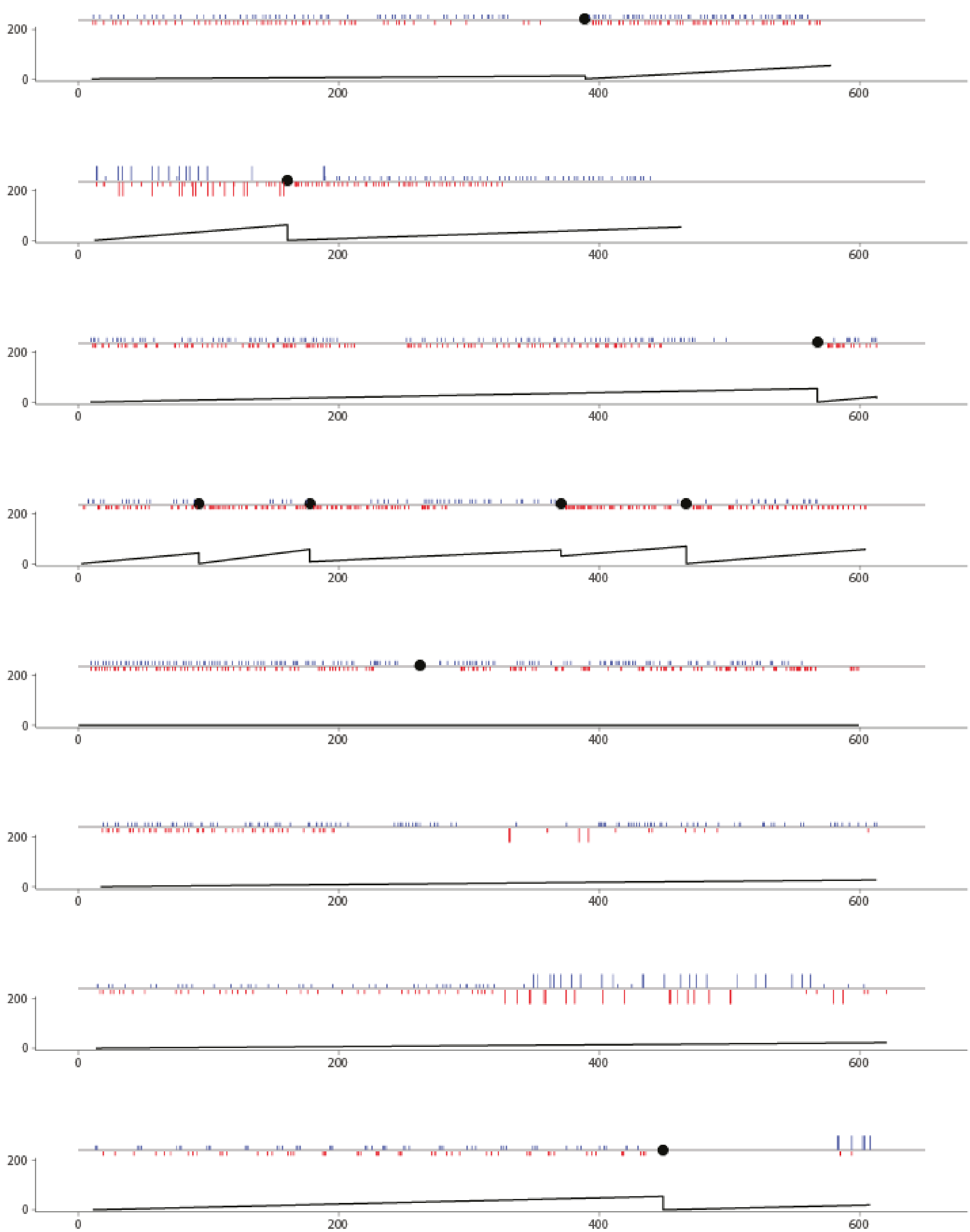
6YO Male Dyads

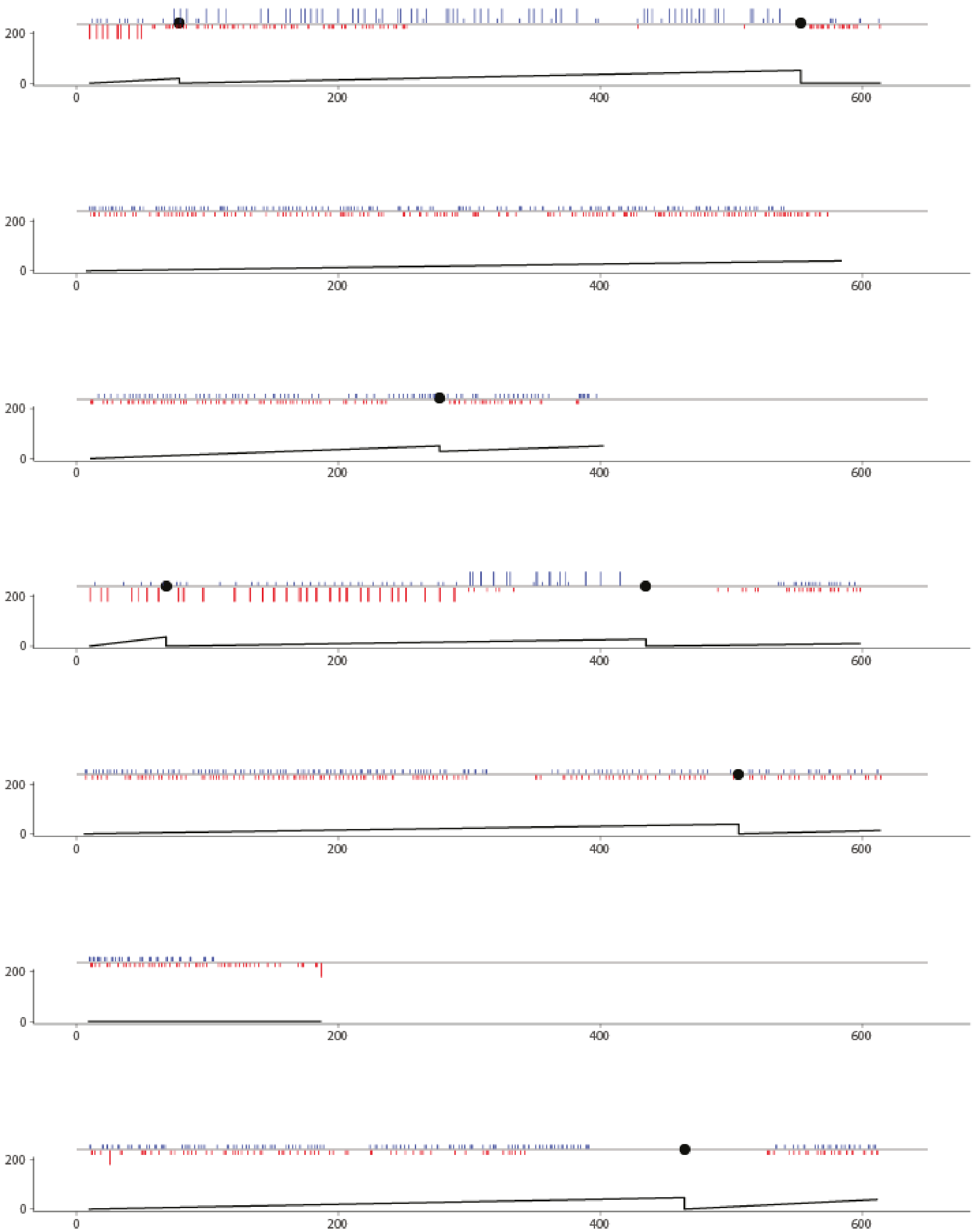
Fig. S2



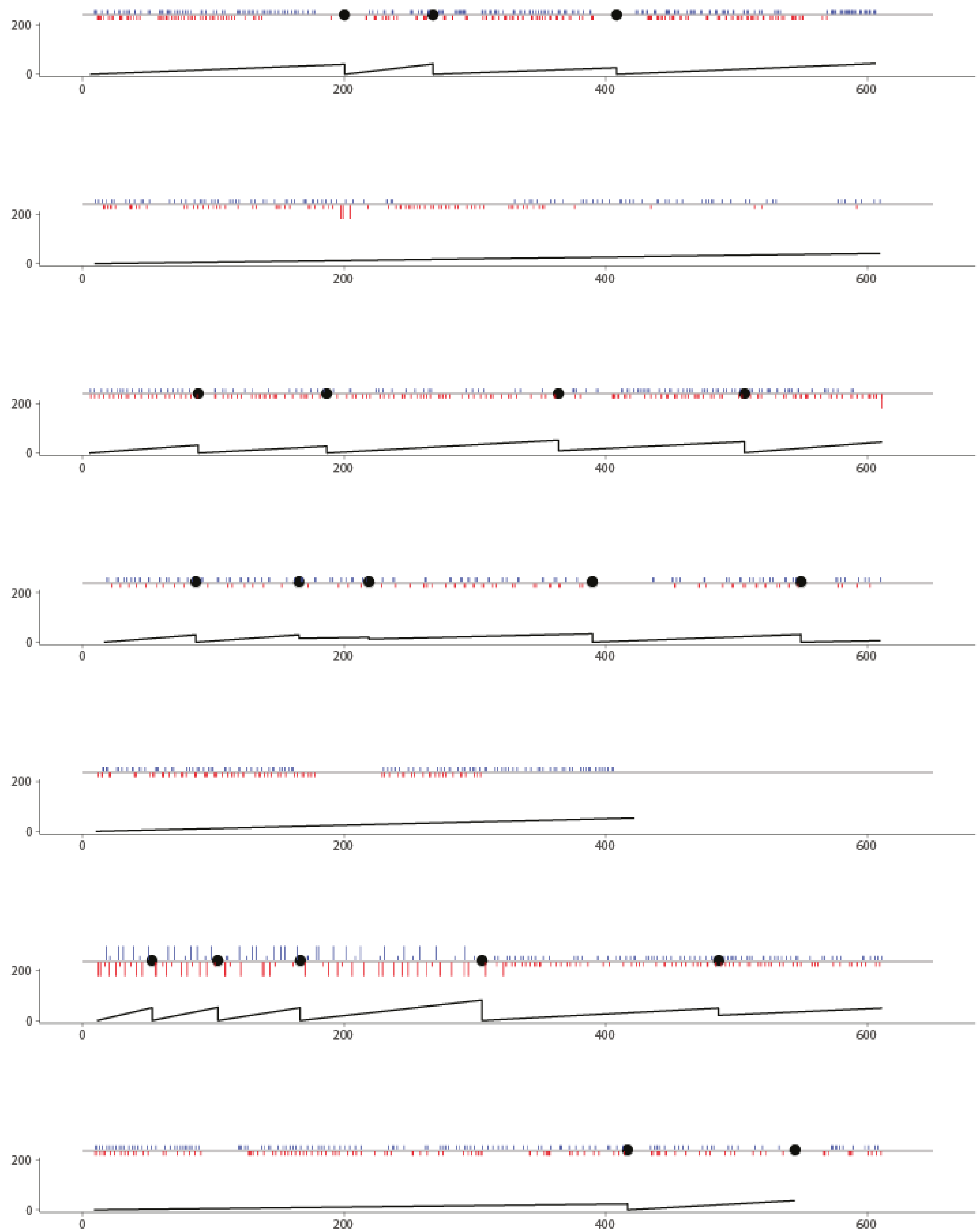

200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600

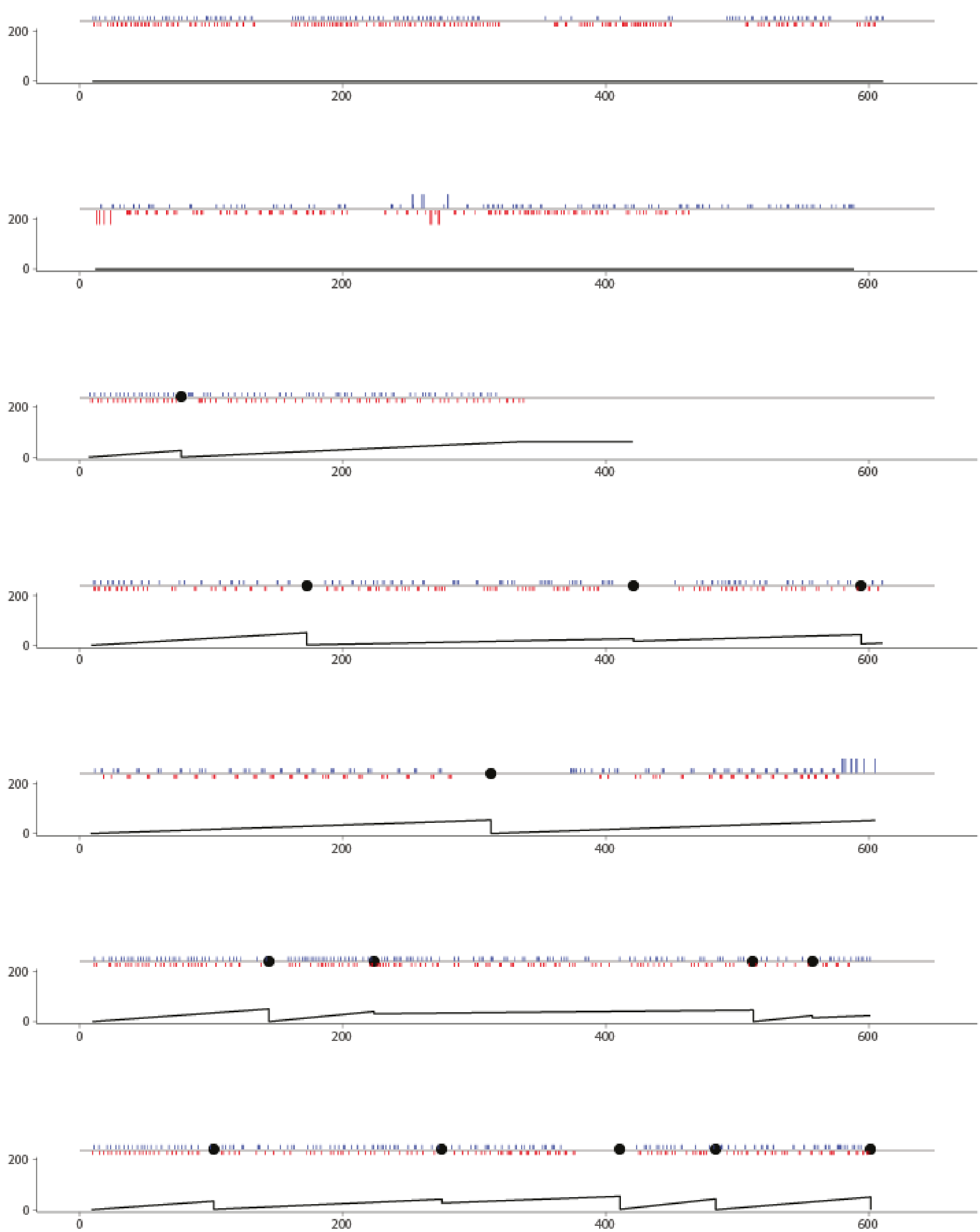
6YO Mix Dyads



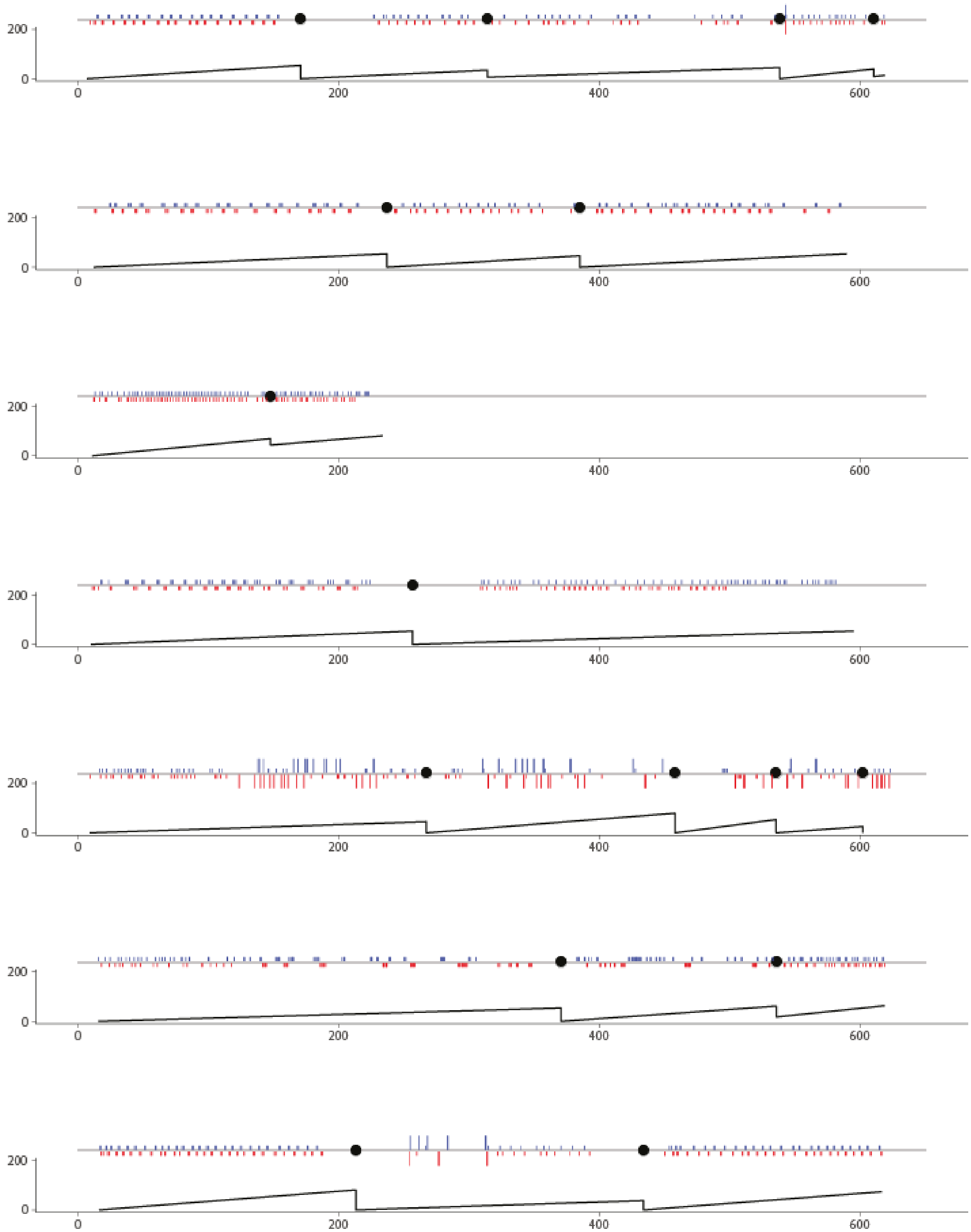

200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600

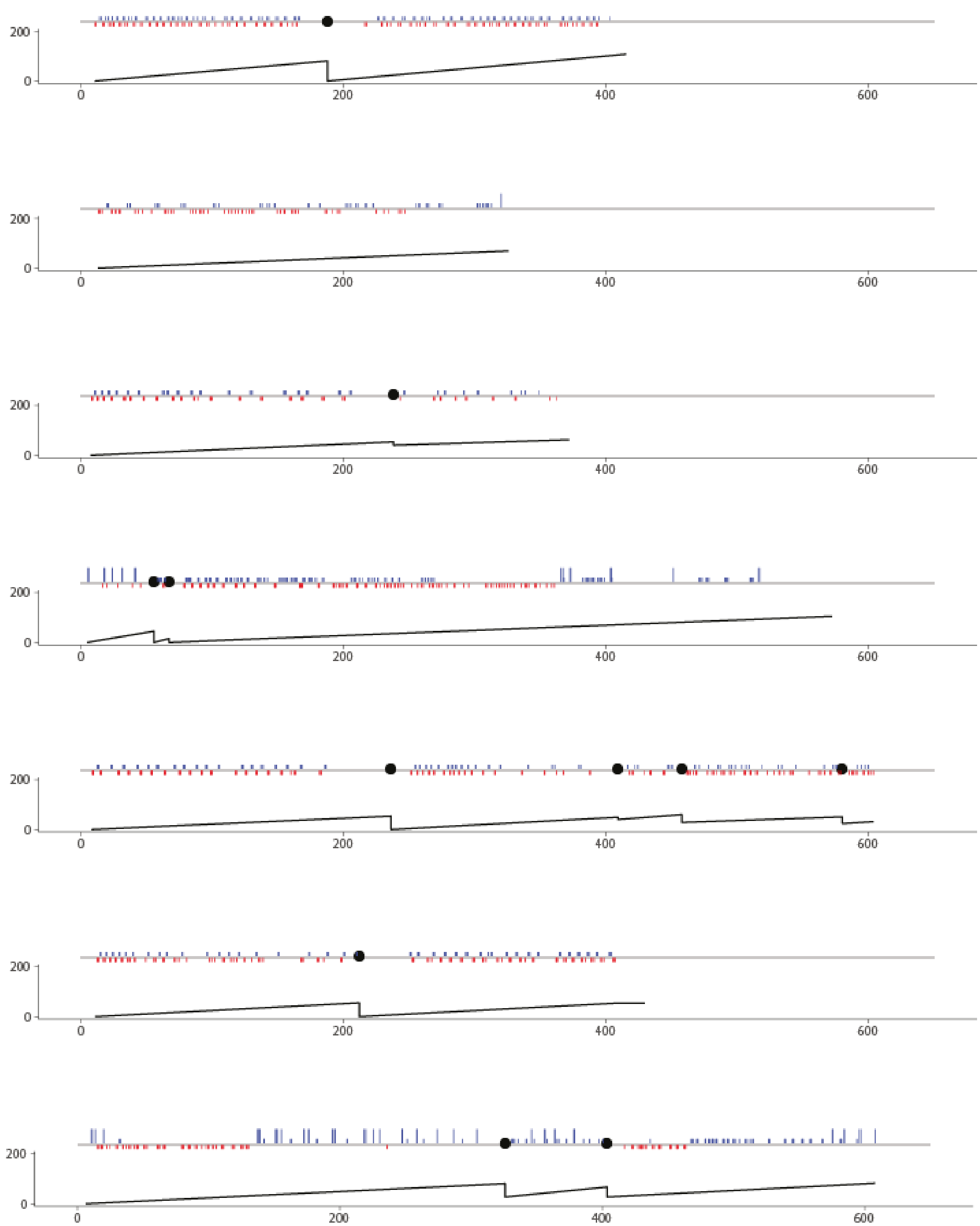
6YO Female Dyads



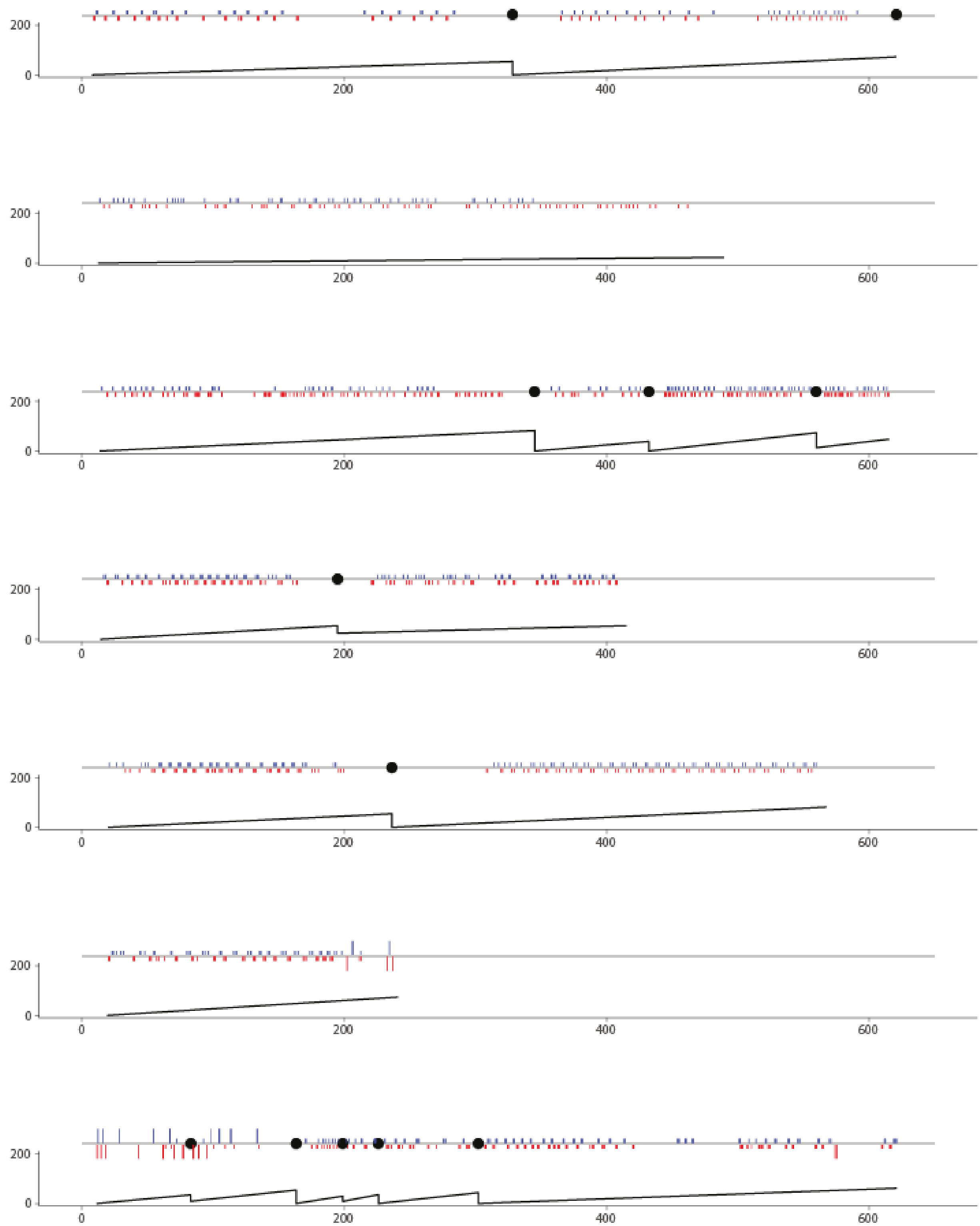

200
0
0
200
400
600

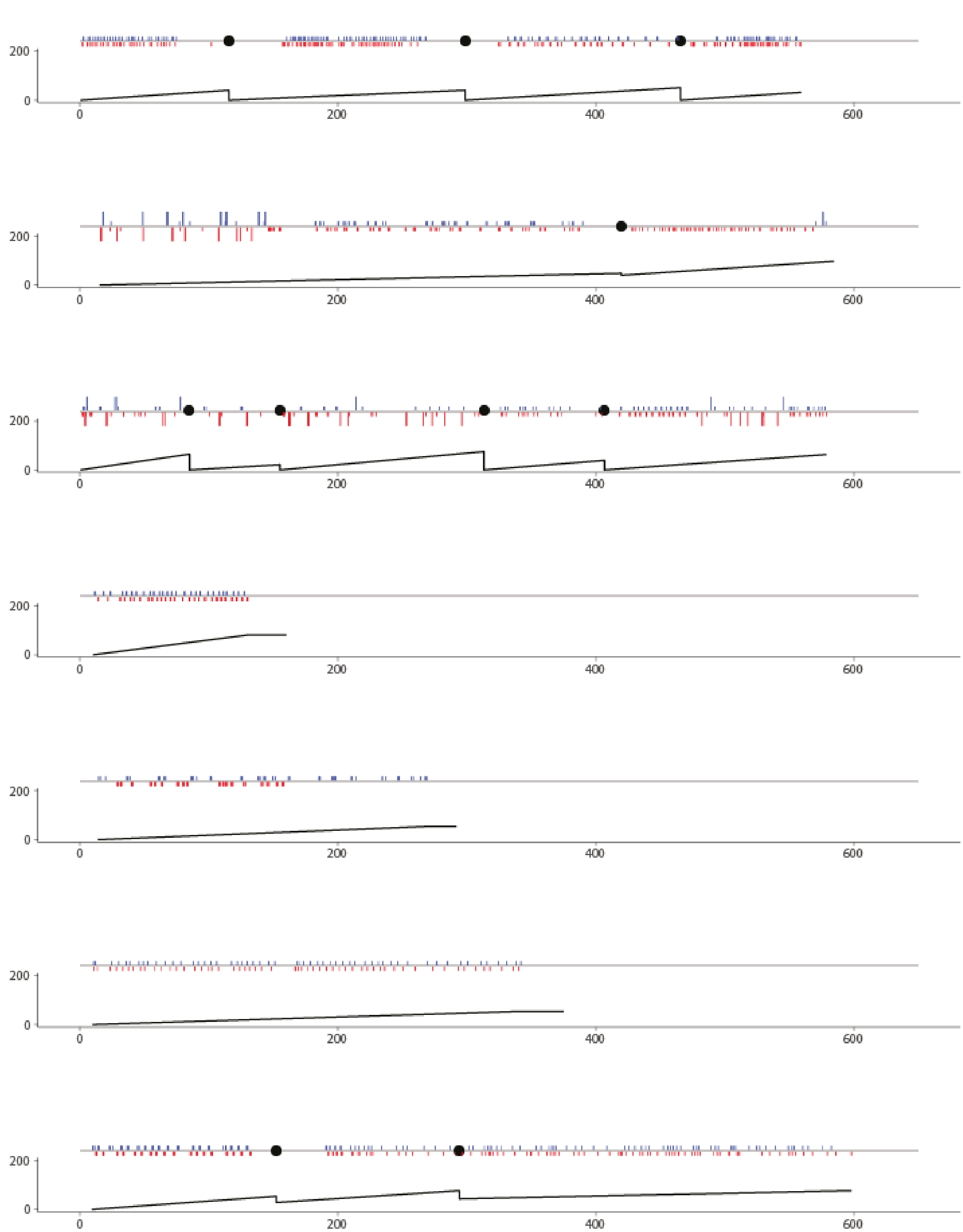
12YO Male Dyads



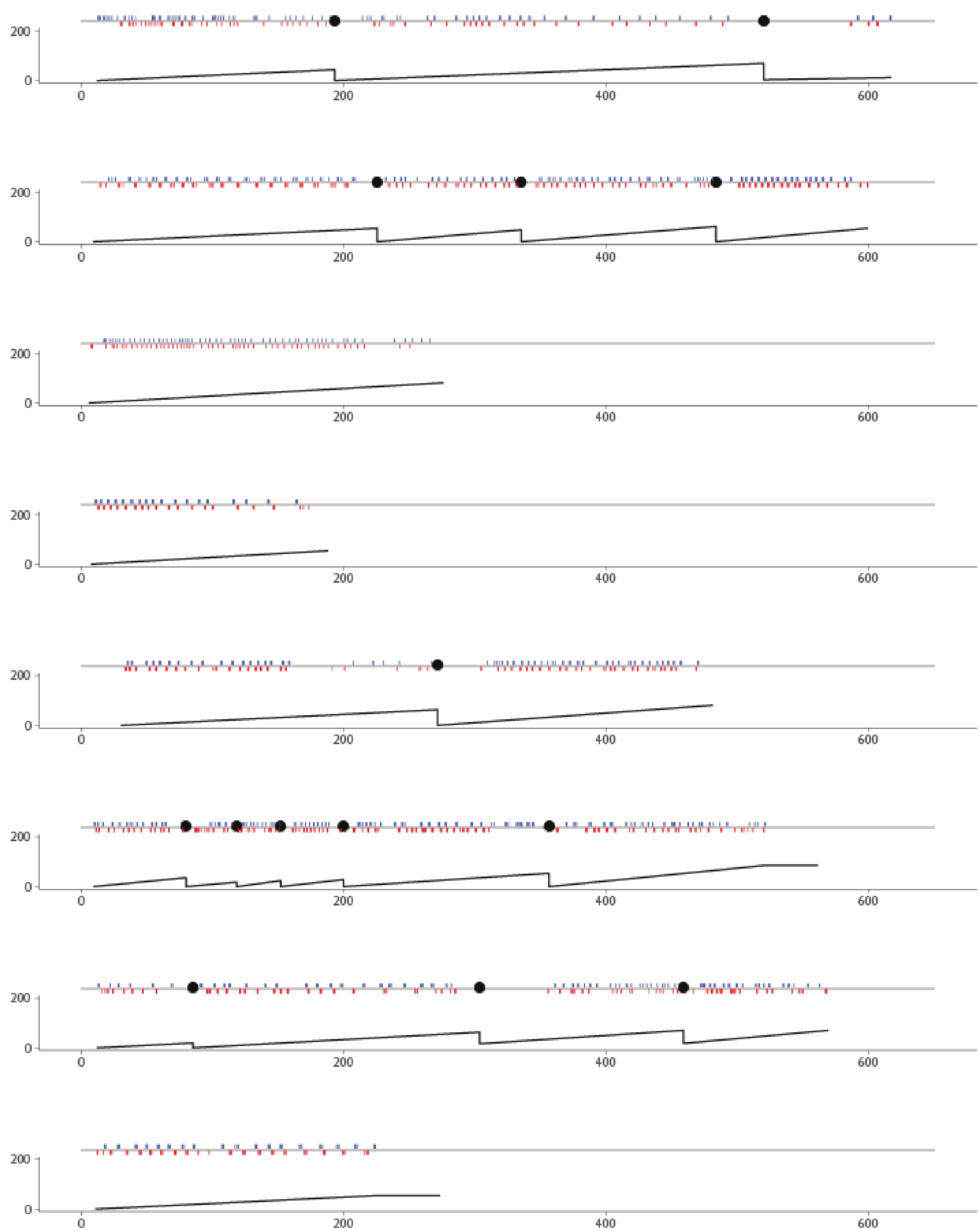

200
0
0
200
400
600

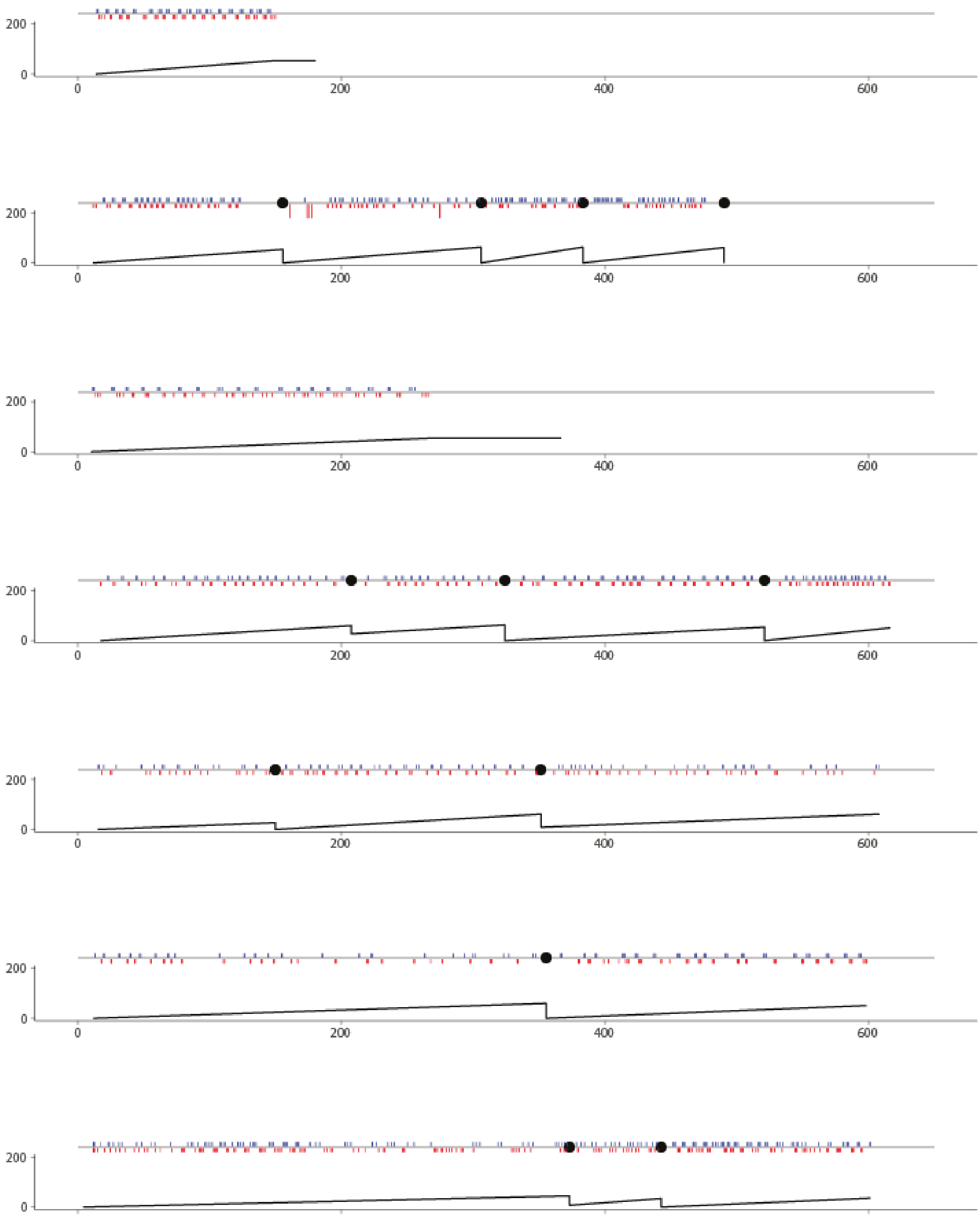
12YO Mix Dyads



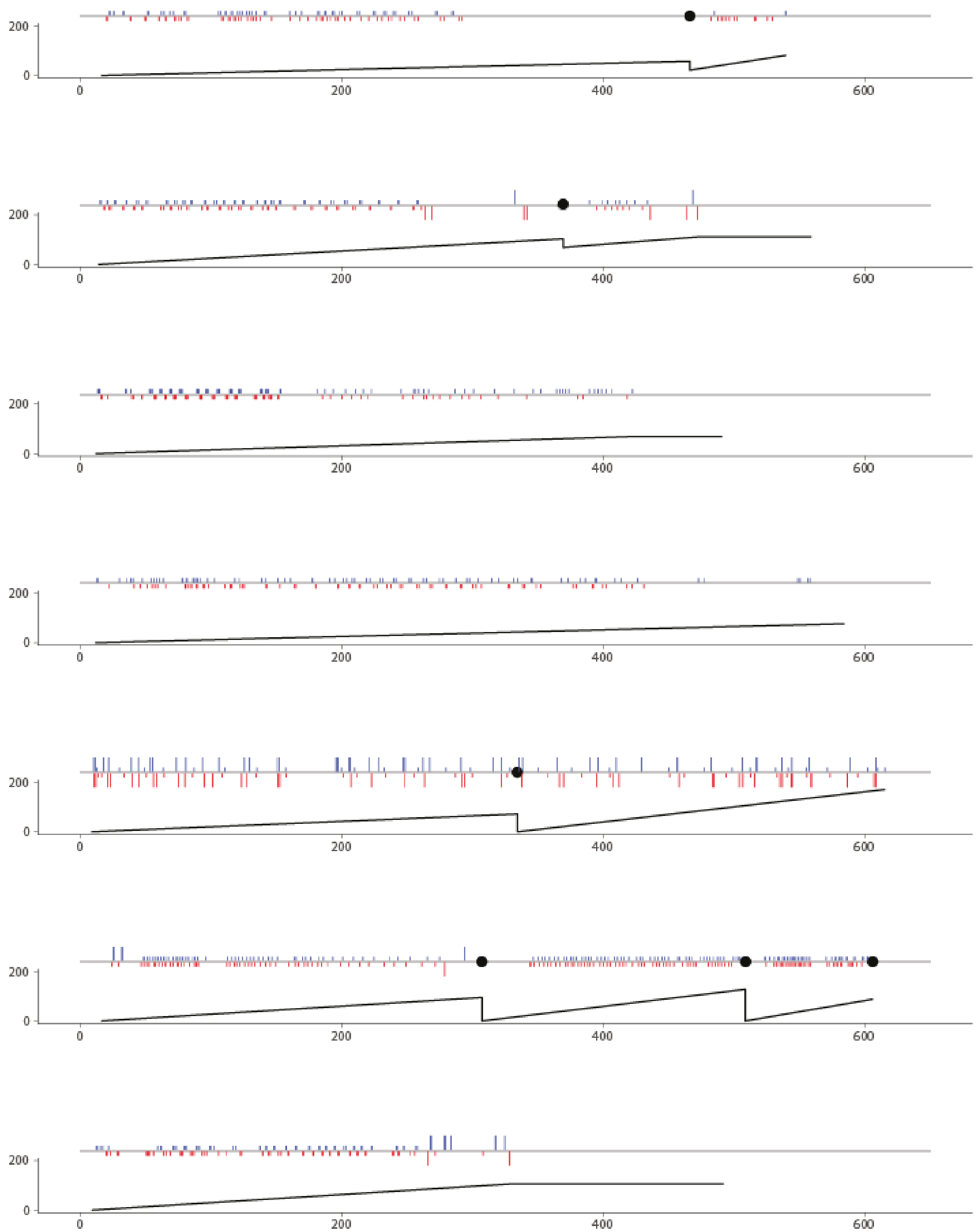

200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600

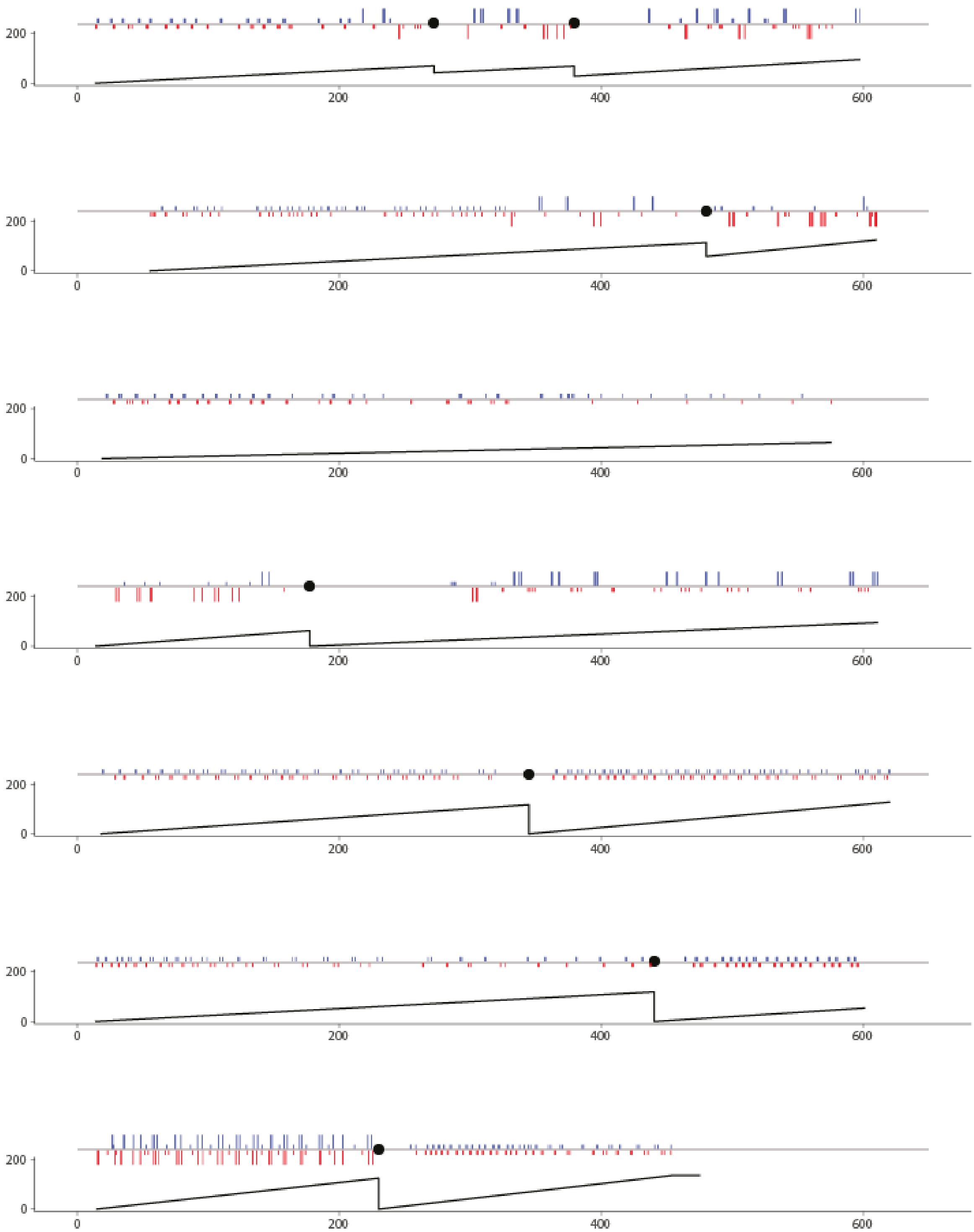
12YO Female Dyads




200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600

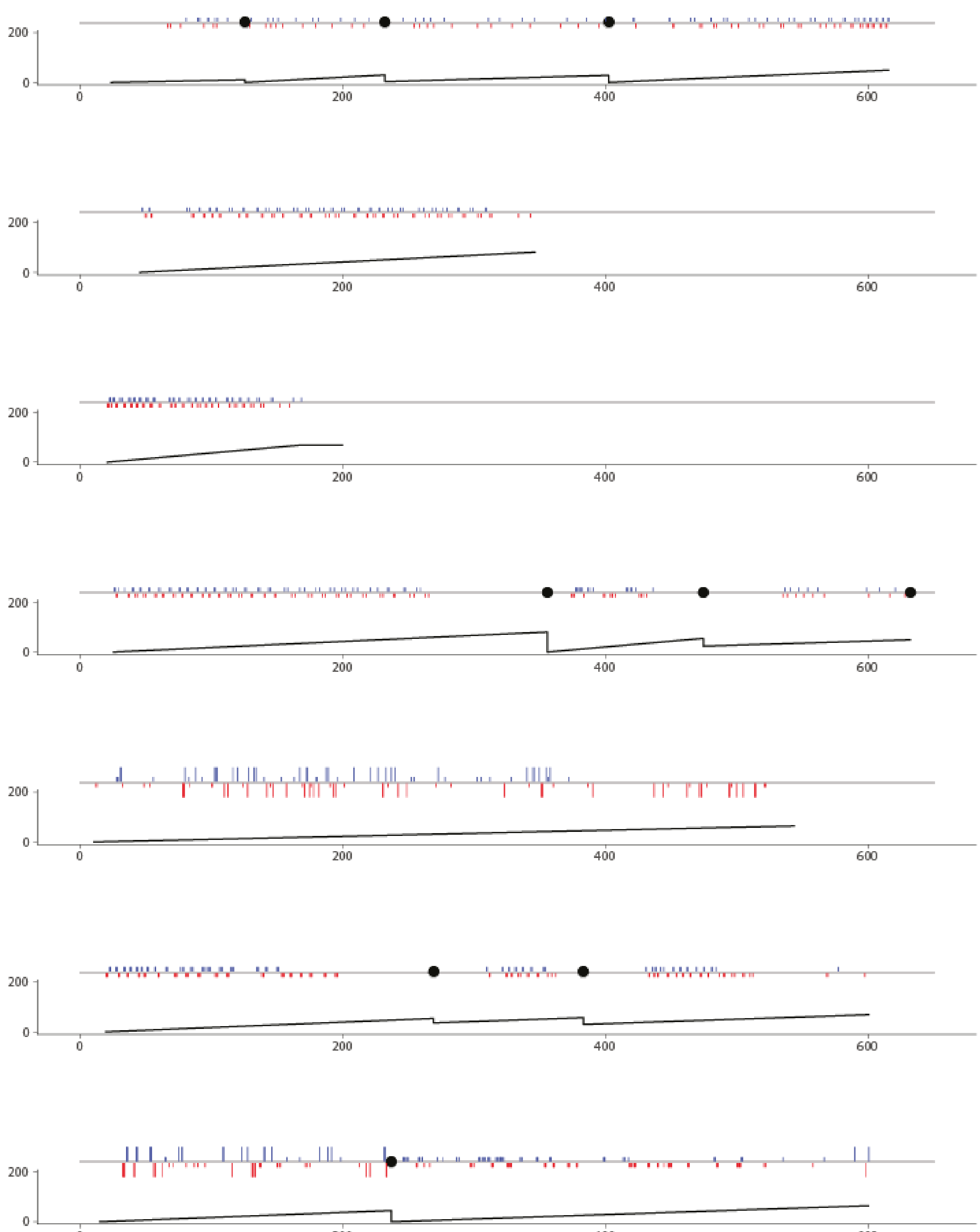
18YO Male Dyads

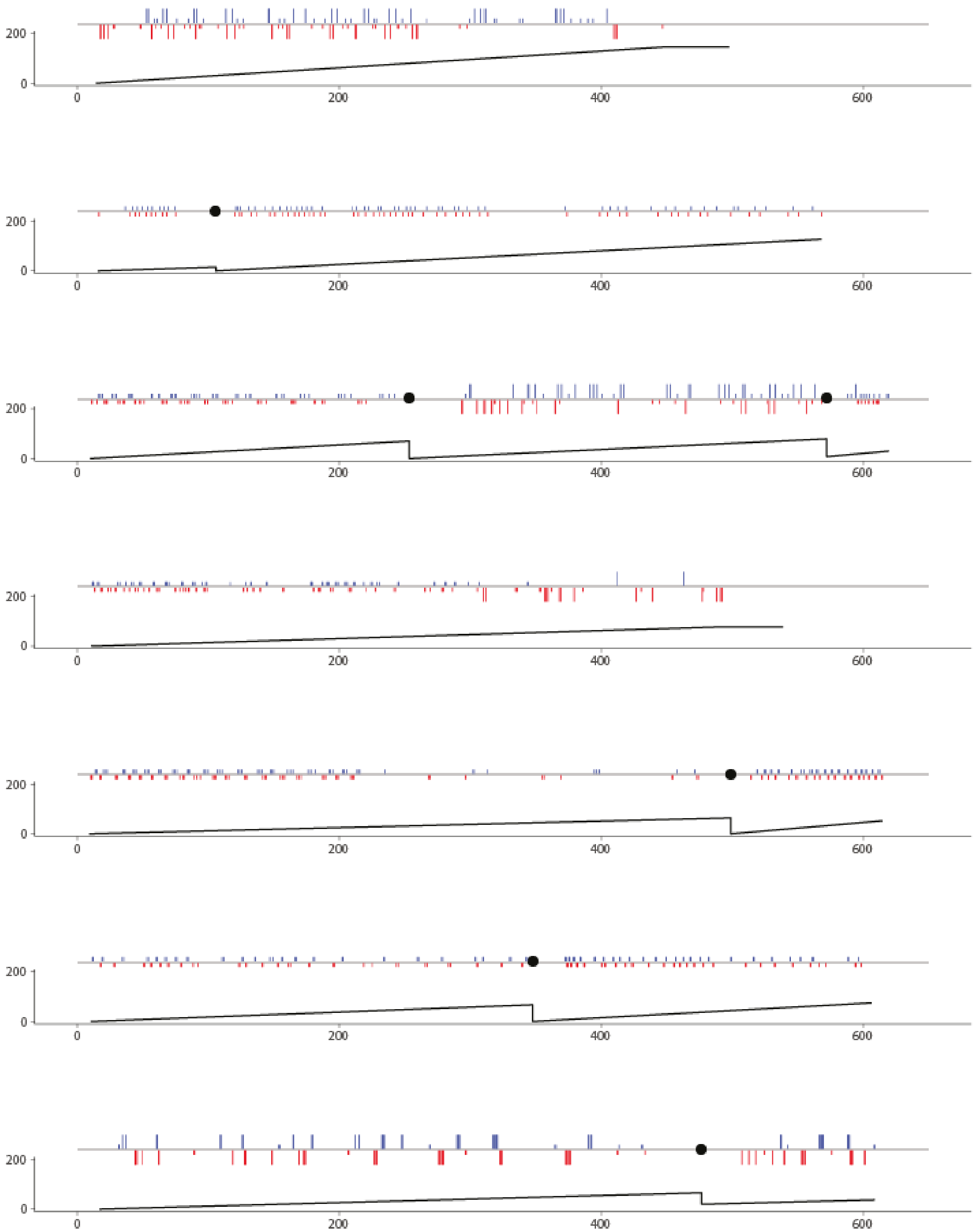




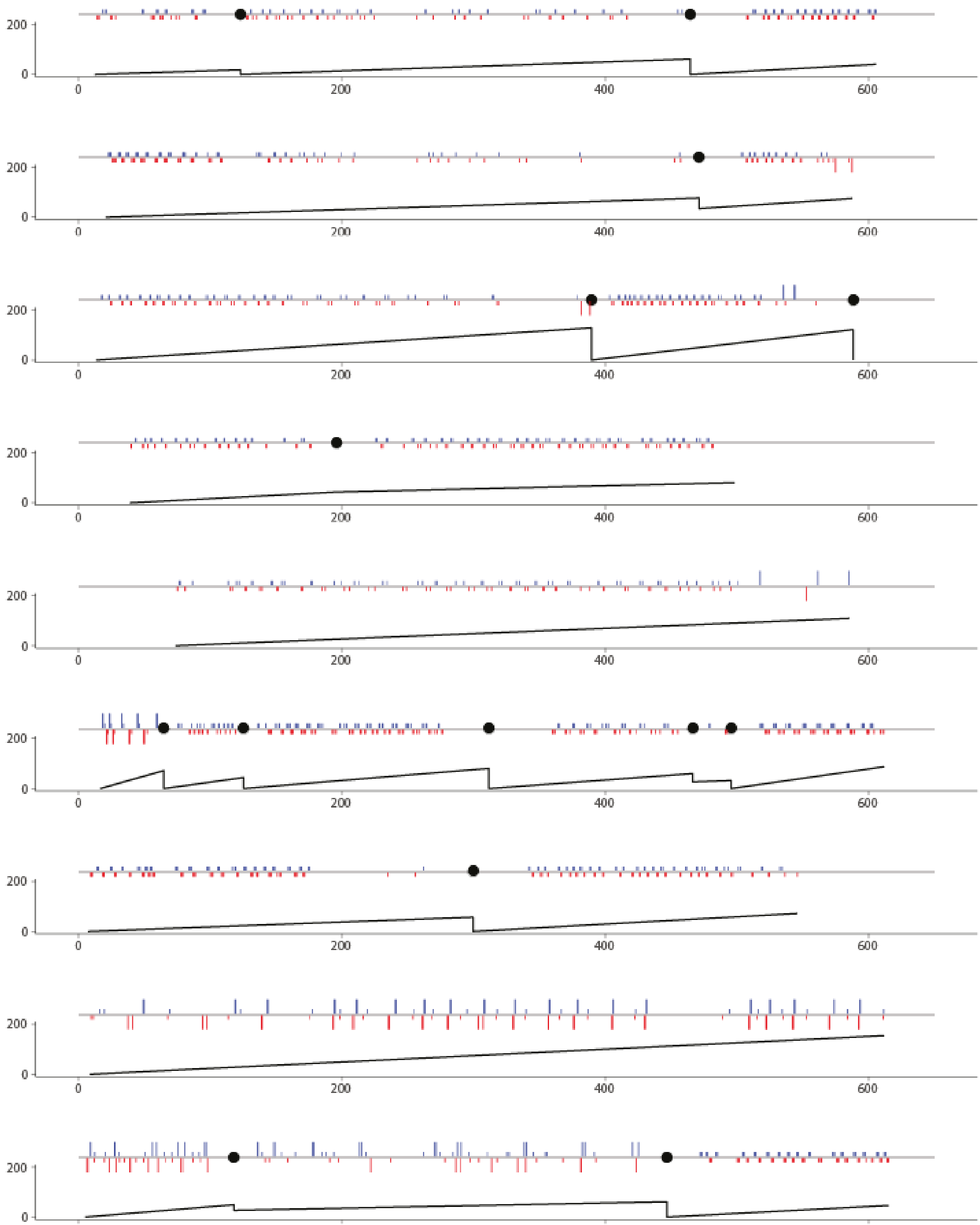
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600

18YO Mix Dyads




200
0
0
200
400
600

18YO Female Dyads



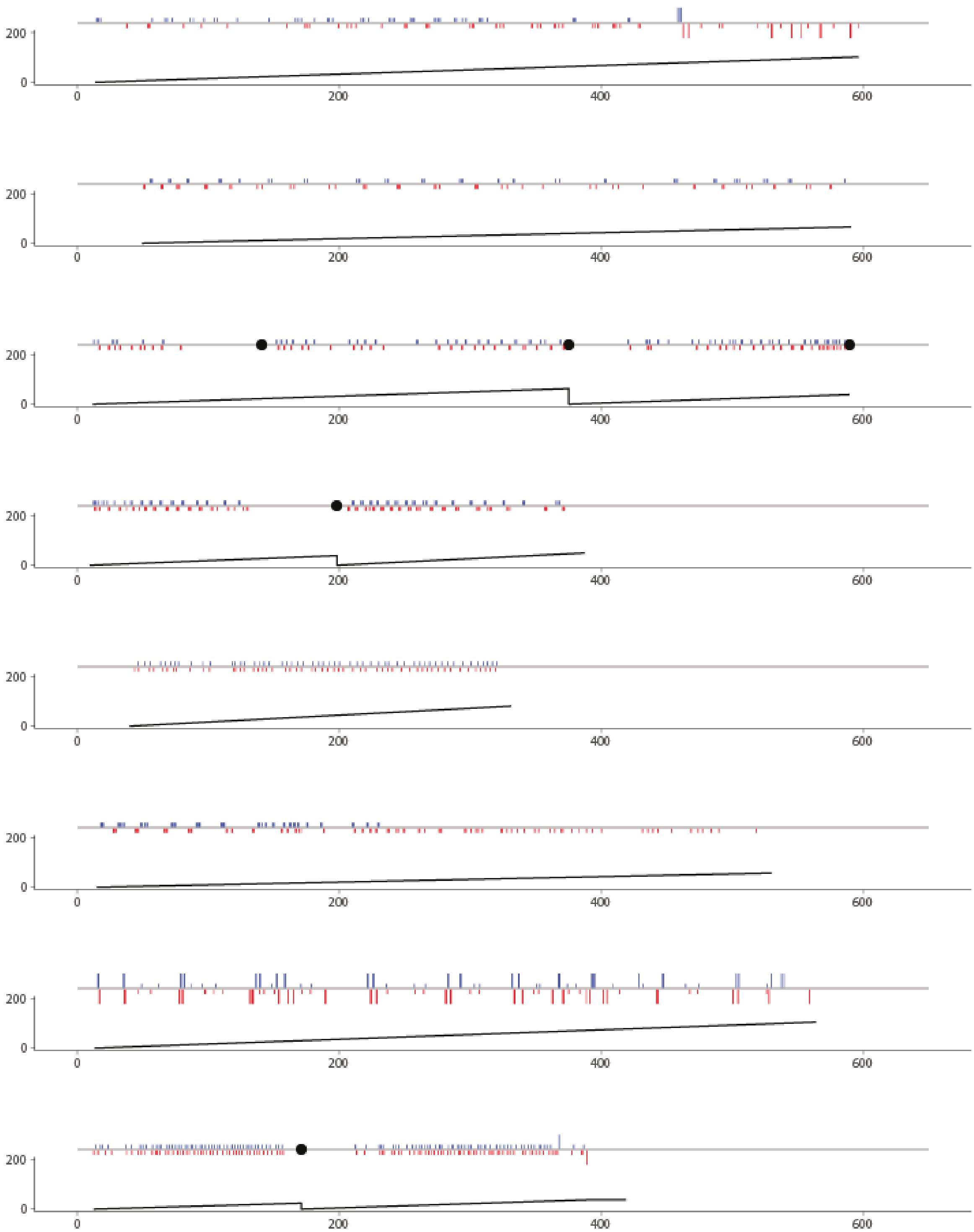

200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600
200
0
0
200
400
600

Fig. S3

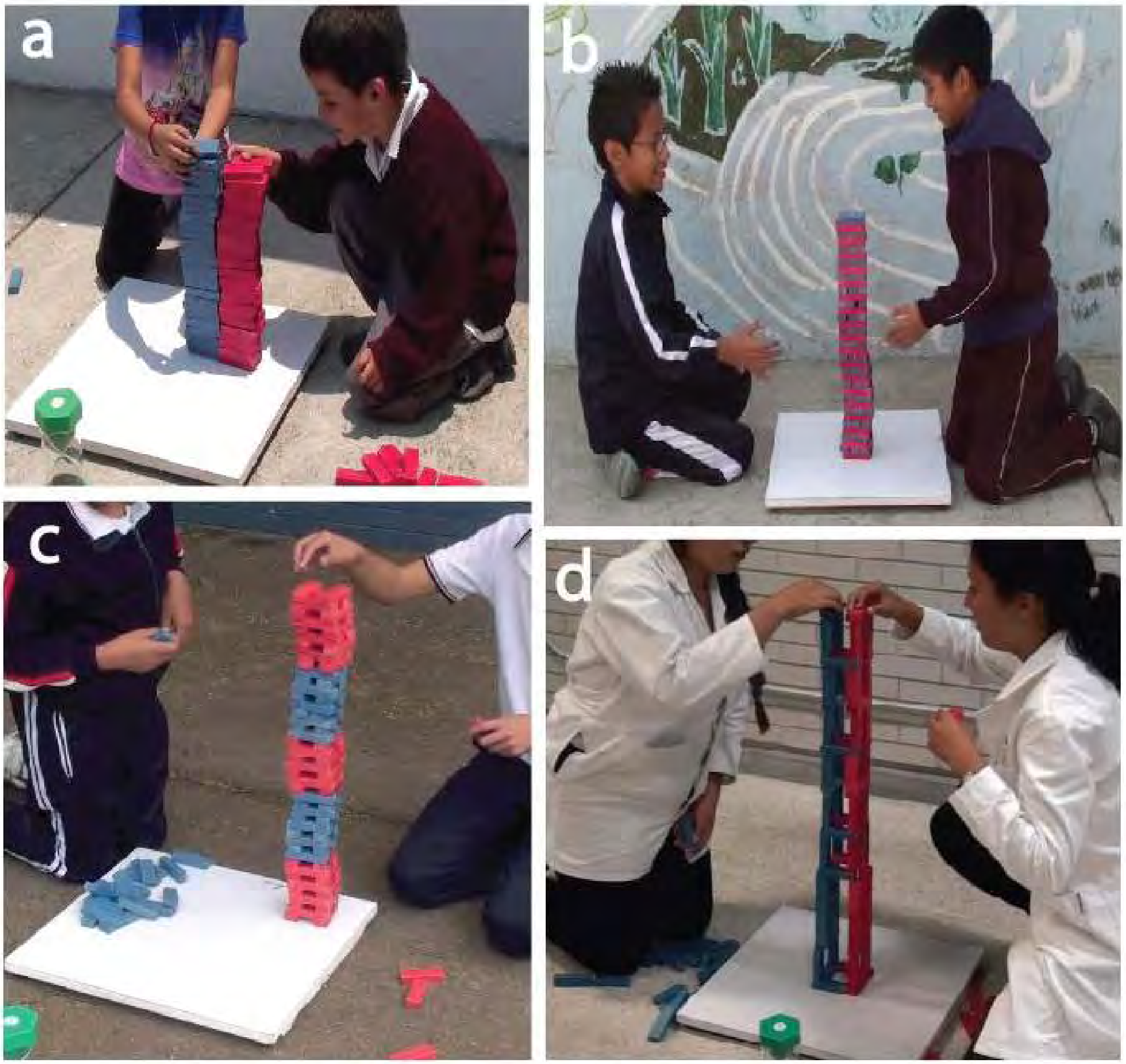


Table S1a. Results of ANOVA for each age category and *post hoc* comparisons between sex combinations for maximum height and height gain.

Max height (cm)					
			Diff.	*C.I.*	*p*
6YOs	*Sex (F=1.28, p = 0.29, η^2=0.07)*	MIX-F	-4.77	-16.99 -- 7.46	0.61
		M-F	3.32	-9.16 -- 15.79	0.79
		M-MIX	8.09	-4.39 -- 20.56	0.26
12YOs	*Sex (F=4.15, p = 0.02, η^2=0.12)*	MIX-F	3.02	-11.77 -- 17.82	0.87
		M-F	8.31	-6.49 -- 23.1	0.37
		M-MIX	5.29	-9.76 -- 20.33	0.67
18YOs	*Sex (F=4.15, p = 0.02, η^2=0.16)*	MIX-F	-2.39	-27.51 -- 22.72	0.97
		M-F	25.39	0.28 -- 50.51	*0.05*
		M-MIX	27.79	1.49 -- 54.09	*0.04*
Height gain (cm per block added)					
			Diff.	*C.I.*	*p*
6YOs	*Sex (F=1.35, p = 0.27, η^2=0.07)*	MIX-F	-0.17	-0.44 -- 0.09	0.24
		M-F	-0.07	-0.34 -- 0.2	0.79
		M-MIX	0.1	-0.16 -- 0.37	0.61
12YOs	*Sex (F=2.62, p = 0.09, η^2=0.05)*	MIX-F	0.04	-0.14 -- 0.22	0.85
		M-F	0.16	-0.02 -- 0.34	0.08
		M-MIX	0.12	-0.06 -- 0.3	0.24
18YOs	*Sex (F=4.01, p = 0.03, η^2=0.17)*	MIX-F	0.05	-0.25 -- 0.36	0.91
		M-F	0.34	0.03 -- 0.64	*0.03*
		M-MIX	0.29	-0.03 -- 0.61	0.09

Table S1b. Results of ANOVA for each sex combination and *post hoc* comparisons between age categories for maximum height and height gain.

Max height (cm)					
			Est.	*C.I.*	*p*
M	*Age (F = 20.19, p = 0, η^2 = 0.4)*	6YO-12YO	16.98	-3.74 -- 37.7	0.13
		6YO-18YO	52.27	31.55 -- 72.99	0
		12YO-18YO	35.29	15.38 -- 55.19	0
F	*Age (F = 8.03, p = 0.001, η^2 = 0.1)*	6YO-12YO	11.99	-7.2 -- 31.18	0.29
		6YO-18YO	30.19	11.53 -- 48.85	0
		12YO-18YO	18.2	0.26 -- 36.14	0.05
MIX	*Age (F = 8.94, p = 0.001, η^2 = 0.31)*	6YO-12YO	19.78	0.89 -- 38.67	0.04
		6YO-18YO	32.57	13.68 -- 51.46	0
		12YO-18YO	12.79	-5.75 -- 31.32	0.23
Height gain (cm per block added)					
			Est.	*C.I.*	*p*
M	*Age (F = 12.19, p = 0, η^2 = 0.4)*	6YO-12YO	0.24	-0.05 -- 0.54	0.13
		6YO-18YO	0.59	0.3 -- 0.89	0
		12YO-18YO	0.35	0.07 -- 0.64	0.01
F	*Age (F = 2.3, p = 0.113, η^2 = 0.1)*	6YO-12YO	0.01	-0.24 -- 0.26	0.99
		6YO-18YO	0.19	-0.06 -- 0.43	0.17
		12YO-18YO	0.17	-0.06 -- 0.41	0.18
MIX	*Age (F = 8.45, p = 0.001, η^2 = 0.31)*	6YO-12YO	0.23	-0.02 -- 0.47	0.07
		6YO-18YO	0.41	0.17 -- 0.66	0
		12YO-18YO	0.18	-0.06 -- 0.42	0.16

Table S2. Results of Poisson-fitted GLM for age and sex combinations and *post hoc* comparisons for total number of collapses and number of towers.

Number of collapses				
		Est. (Std. Error)	*z*	*p*
6YOs	MIX-F	0.79 (0.3)	2.6	*0.03*
	M-F	0.52 (0.28)	1.87	0.15
	M-MIX	-0.27 (0.33)	-0.82	0.69
12YOs	MIX-F	0.1 (0.29)	0.34	0.94
	M-F	-0.11 (0.27)	-0.39	0.92
	M-MIX	-0.2 (0.29)	-0.71	0.76
18YOs	MIX-F	0.04 (0.35)	0.12	0.99
	M-F	0.19 (0.36)	0.51	0.86
	M-MIX	0.14 (0.38)	0.38	0.92
Number of towers				
		Est. (Std. Error)	*z*	*p*
6YOs	MIX-F	0.54 (0.24)	2.26	0.06
	M-F	0.31 (0.23)	1.35	0.37
	M-MIX	0.22 (0.26)	0.85	0.67
12YOs	MIX-F	0.06 (0.23)	0.28	0.96
	M-F	-0.07 (0.22)	-0.31	0.95
	M-MIX	0.13 (0.23)	0.58	0.83
18YOs	MIX-F	-0.03 (0.26)	-0.12	0.99
	M-F	0.04 (0.26)	0.17	0.98
	M-MIX	-0.07 (0.27)	-0.27	0.96

Table S3a. Results of ANOVA for each age category and *post hoc* comparisons between sex combinations for addition rate, color ratio and proportion of vertical pieces.

Addition rate (pieces per sec)					
			Diff.	*C.I.*	*p*
6YOs	*Sex (F=0.56, p = 0.57, η^2=0.03)*	MIX-F	-0.04	-0.13 -- 0.06	0.61
		M-F	-0.03	-0.12 -- 0.06	0.65
		M-MIX	0	-0.09 -- 0.09	1
12YOs	*Sex (F=0.61, p = 0.55, η^2=0.03)*	MIX-F	-0.04	-0.14 -- 0.07	0.66
		M-F	-0.04	-0.15 -- 0.06	0.57
		M-MIX	-0.01	-0.11 -- 0.1	0.99
18YOs	*Sex (F=0.47, p = 0.63, η^2=0.02)*	MIX-F	-0.03	-0.11 -- 0.05	0.69
		M-F	-0.03	-0.11 -- 0.05	0.69
		M-MIX	0	-0.08 -- 0.08	1
Color ratio					
			Diff.	*C.I.*	*p*
6YOs	*Sex (F=0.33, p = 0.72, η^2=0.02)*	MIX-F	-0.07	-0.31 -- 0.16	0.72
		M-F	-0.02	-0.25 -- 0.21	0.98
		M-MIX	0.06	-0.17 -- 0.28	0.82
12YOs	*Sex (F=3.8, p = 0.03, η^2=0.16)*	MIX-F	-0.12	-0.27 -- 0.02	0.12
		M-F	0.04	-0.11 -- 0.19	0.79
		M-MIX	0.16	0.01 -- 0.31	0.03
18YOs	*Sex (F=0.79, p = 0.46, η^2=0.04)*	MIX-F	-0.02	-0.09 -- 0.05	0.76
		M-F	0.02	-0.05 -- 0.09	0.81
		M-MIX	0.04	-0.04 -- 0.11	0.43

Proportion of vertical pieces					
			Diff.	*C.I.*	*p*
6YOs	*Sex ($F=1.3$, $p = 0.28$, $\eta^2=0.06$)*	MIX-F	0.06	-0.05 -- 0.17	0.38
		M-F	0	-0.11 -- 0.11	1
		M-MIX	-0.06	-0.17 -- 0.04	0.33
12YOs	*Sex ($F=1$, $p = 0.38$, $\eta^2=0.05$)*	MIX-F	0.04	-0.03 -- 0.11	0.38
		M-F	0.03	-0.04 -- 0.1	0.55
		M-MIX	-0.01	-0.08 -- 0.07	0.96
18YOs	*Sex ($F=0.45$, $p = 0.64$, $\eta^2=0.02$)*	MIX-F	0.08	-0.12 -- 0.28	0.62
		M-F	0.03	-0.17 -- 0.23	0.92
		M-MIX	-0.05	-0.26 -- 0.17	0.85

Table S3b. Results of ANOVA for each sex combination and *post hoc* comparisons between age categories for addition rate, color ratio and proportion of vertical pieces.

Addition rate (pieces per sec)					
			Est.	*C.I.*	*p*
M	*Age ($F = 10.74$, $p = 0$, $\eta^2 = 0.35$)*	6YO-12YO	0.04	-0.04 -- 0.11	0.49
		6YO-18YO	-0.11	-0.18 -- -0.03	0.01
		12YO-18YO	-0.14	-0.22 -- -0.06	0
F	*Age ($F = 8.58$, $p = 0.001$, $\eta^2 = 0.29$)*	6YO-12YO	0.05	-0.05 -- 0.15	0.5
		6YO-18YO	-0.11	-0.21 -- -0.01	0.02
		12YO-18YO	-0.16	-0.26 -- -0.06	0
MIX	*Age ($F = 6.6$, $p = 0.003$, $\eta^2 = 0.25$)*	6YO-12YO	0.05	-0.06 -- 0.15	0.53
		6YO-18YO	-0.1	-0.2 -- 0	0.04
		12YO-18YO	-0.15	-0.25 -- -0.05	0
Color ratio					
			Est.	*C.I.*	*p*
M	*Age ($F = 0.49$, $p = 0.615$, $\eta^2 = 0.35$)*	6YO-12YO	0.01	-0.16 -- 0.19	0.98
		6YO-18YO	-0.06	-0.23 -- 0.12	0.72
		12YO-18YO	-0.07	-0.25 -- 0.11	0.63
F	*Age ($F = 1.69$, $p = 0.197$, $\eta^2 = 0.29$)*	6YO-12YO	-0.05	-0.17 -- 0.08	0.66
		6YO-18YO	-0.09	-0.21 -- 0.03	0.17
		12YO-18YO	-0.05	-0.17 -- 0.07	0.61
MIX	*Age ($F = 0.74$, $p = 0.484$, $\eta^2 = 0.25$)*	6YO-12YO	-0.09	-0.28 -- 0.09	0.45
		6YO-18YO	-0.04	-0.23 -- 0.15	0.88
		12YO-18YO	0.06	-0.14 -- 0.25	0.76

Proportion of vertical pieces					
			Est.	*C.I.*	*p*
M	*Age ($F = 3.23, p = 0.05, \eta^2 = 0.35$)*	6YO-12YO	0.01	-0.11 -- 0.12	0.99
		6YO-18YO	0.11	-0.01 -- 0.23	0.07
		12YO-18YO	0.1	-0.02 -- 0.22	0.1
F	*Age ($F = 2.43, p = 0.1, \eta^2 = 0.29$)*	6YO-12YO	-0.03	-0.15 -- 0.1	0.85
		6YO-18YO	0.08	-0.05 -- 0.2	0.29
		12YO-18YO	0.1	-0.02 -- 0.22	0.1
MIX	*Age ($F = 1.84, p = 0.172, \eta^2 = 0.25$)*	6YO-12YO	-0.05	-0.23 -- 0.13	0.79
		6YO-18YO	0.09	-0.09 -- 0.28	0.42
		12YO-18YO	0.14	-0.04 -- 0.33	0.16

Table S4a. Results of ANOVA for each age category and *post hoc* comparisons between sex combinations for latency to start, latency to restart and early conclusion.

Latency to start (sec)					
			Diff.	*C.I.*	*p*
6YOs	*Sex (F=0.74, p = 0.48, η^2=0.03)*	MIX-F	-0.18	-4.32 -- 3.96	0.99
		M-F	1.67	-2.47 -- 5.82	0.59
		M-MIX	1.85	-2.22 -- 5.92	0.52
12YOs	*Sex (F=0.49, p = 0.61, η^2=0.0)*	MIX-F	-0.53	-5.25 -- 4.19	0.96
		M-F	-1.88	-6.6 -- 2.84	0.6
		M-MIX	-1.35	-6.15 -- 3.45	0.77
18YOs	*Sex (F=0.83, p = 0.44, η^2=0.0)*	MIX-F	-5.99	-18.11 -- 6.14	0.46
		M-F	-4.65	-16.52 -- 7.23	0.61
		M-MIX	1.34	-11.34 -- 14.01	0.96
Latency to restart (sec)					
			Diff.	*C.I.*	*p*
6YOs	*Sex (F=1.71, p = 0.19, η^2=0.03)*	MIX-F	9.05	-5.19 -- 23.3	0.29
		M-F	-2.37	-15.42 -- 10.68	0.9
		M-MIX	-11.43	-26.94 -- 4.08	0.19
12YOs	*Sex (F=0.25, p = 0.78, η^2=0.02)*	MIX-F	2.96	-8.21 -- 14.13	0.8
		M-F	2.54	-8.05 -- 13.13	0.83
		M-MIX	-0.42	-11.49 -- 10.65	1
18YOs	*Sex (F=0.25, p = 0.78, η^2=0.04)*	MIX-F	1.55	-15.58 -- 18.68	0.97
		M-F	5.14	-12.71 -- 23	0.77
		M-MIX	3.59	-15.2 -- 22.38	0.89

Early conclusion (sec)					
			Diff.	*C.I.*	*p*
6YOs	*Sex (F=0.49, p = 0.63, η^2=0.29)*	MIX-F	-70.4	-386.29 -- 245.49	0.79
		M-F	-97.36	-386.88 -- 192.15	0.61
		M-MIX	-26.97	-279.68 -- 225.74	0.95
12YOs	*Sex (F=0.41, p = 0.67, η^2=0.12)*	MIX-F	33.06	-131.9 -- 198.03	0.86
		M-F	59.53	-113.49 -- 232.55	0.65
		M-MIX	26.47	-146.55 -- 199.49	0.92
18YOs	*Sex (F=1.2, p = 0.36, η^2=0.05)*	MIX-F	-59.88	-323.34 -- 203.59	0.77
		M-F	77.28	-155.08 -- 309.63	0.59
		M-MIX	137.15	-140.57 -- 414.87	0.35

Table S4b. Results of ANOVA for each sex combination, and *post hoc* comparisons between age categories for latency to start, latency to restart and early conclusion.

Latency to start (sec)					
			Est.	*C.I.*	*p*
M	*Age ($F = 3.05, p = 0.058, \eta^2 = 0.13$)*	6YO-12YO	-1.15	-8.1 -- 5.81	0.92
		6YO-18YO	5.55	-1.41 -- 12.5	0.14
		12YO-18YO	6.69	-0.38 -- 13.77	0.07
F	*Age ($F = 4.43, p = 0.018, \eta^2 = 0.17$)*	6YO-12YO	2.4	-8.32 -- 13.13	0.85
		6YO-18YO	11.87	1.45 -- 22.28	0.02
		12YO-18YO	9.46	-0.76 -- 19.68	0.07
MIX	*Age ($F = 5, p = 0.012, \eta^2 = 0.2$)*	6YO-12YO	2.06	-2.57 -- 6.68	0.53
		6YO-18YO	6.06	1.34 -- 10.78	0.01
		12YO-18YO	4	-0.79 -- 8.8	0.12
Latency to restart (sec)					
			Est.	*C.I.*	*p*
M	*Age ($F = 4.26, p = 0.019, \eta^2 = 0.13$)*	6YO-12YO	6.15	-5.94 -- 18.25	0.44
		6YO-18YO	17.78	3.11 -- 32.45	0.01
		12YO-18YO	11.63	-2.4 -- 25.66	0.12
F	*Age ($F = 3.26, p = 0.044, \eta^2 = 0.17$)*	6YO-12YO	1.24	-7.89 -- 10.37	0.94
		6YO-18YO	10.26	0.24 -- 20.29	0.04
		12YO-18YO	9.02	-1.48 -- 19.53	0.11
MIX	*Age ($F = 0.47, p = 0.628, \eta^2 = 0.2$)*	6YO-12YO	-4.85	-24.08 -- 14.37	0.82
		6YO-18YO	2.76	-18.27 -- 23.8	0.95
		12YO-18YO	7.62	-11.98 -- 27.21	0.62

Early conclusion (sec)					
			Est.	*C.I.*	*p*
M	*Age (F = 3.27, p = 0.081, η^2 = 0.13)*	6YO-12YO	31.34	-123.47 -- 186.15	0.85
		6YO-18YO	162.32	-16.44 -- 341.07	0.08
		12YO-18YO	130.98	-47.78 -- 309.73	0.16
F	*Age (F = 2.24, p = 0.162, η^2 = 0.17)*	6YO-12YO	-125.55	-345 -- 93.89	0.3
		6YO-18YO	-12.33	-245.08 -- 220.43	0.99
		12YO-18YO	113.23	-60.26 -- 286.71	0.22
MIX	*Age (F = 0.03, p = 0.969, η^2 = 0.2)*	6YO-12YO	-22.09	-307.26 -- 263.08	0.97
		6YO-18YO	-1.8	-369.95 -- 366.35	1
		12YO-18YO	20.29	-309 -- 349.58	0.98

Expected values for the three types of block positions

In the context of the Tower Building Task, the probability of success may be understood as the stability of the block placed, which is mainly given by the surface contact whereas the value of the reward (outcome) is represented by height gain. In three cases the expected value is identical; however, the level of risk changes among options, being the horizontal 1 the safer.

Position	Height gain (cm)	Surface contact (cm2)	Expected value
Vertical	7.5	3.75	28.12
Horizontal 1	1.5	18.75	28.12
Horizontal 2	2.5	11.25	28.12